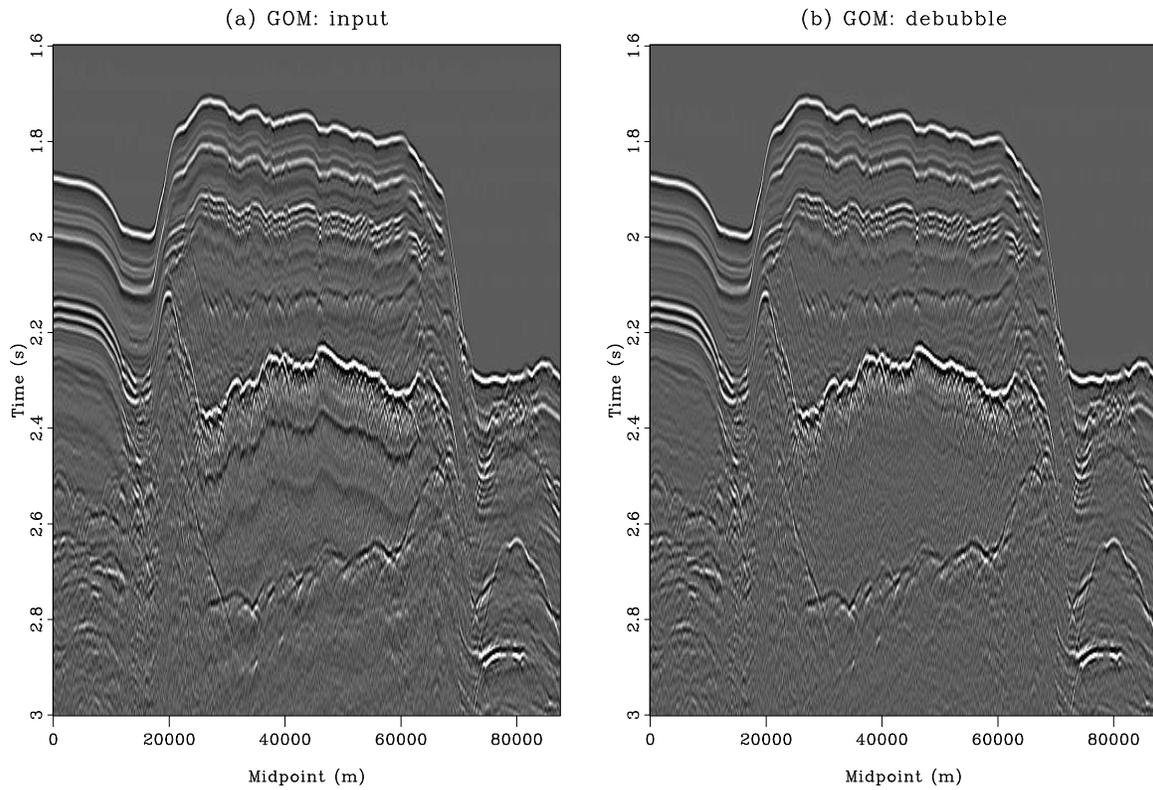


# STANFORD EXPLORATION PROJECT

*Ali Almomin, Humberto Arevalo-Lopez, Biondo Biondi, Jason Chang, Jon Claerbout, Robert Clapp, Taylor Dahlke, Marine Denolle, Antoine Guitton, Adam Halpert, Huy Le, Chris Leader, Stewart Levin, Stewart Levin Yunyue Li, Musa Maharramov, Eileen Martin, Gary Mavko, Dave Nichols, Sjoerd de Ridder, Shuki Ronen, Xukai Shen, and Mandy Wong*

Report Number 150, September 2013



*Copyright © 2013*

*by the Board of Trustees of the Leland Stanford Junior University*

*Copying permitted for all internal purposes of the Sponsors of Stanford Exploration Project*

## Preface

The electronic version of this report<sup>1</sup> makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR] are promises by the author about the reproducibility of each figure result. Reproducibility is a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

**ER** denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library<sup>2</sup>. We assume you have a UNIX workstation with Fortran, Fortran90, C, C++, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

**CR** denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. The primary reasons for the CR designation is that the processing requires 20 minutes or more, MPI or CUDA based code, or commercial packages such as Matlab or Mathematica.

**NR** denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.6 (using the Intel compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. More information on reproducing SEP's electronic documents is available online<sup>3</sup>.

---

<sup>1</sup><http://sepwww.stanford.edu/private/docs/sep150>

<sup>2</sup><http://sepwww.stanford.edu/public/docs/sepdatalib/toc.html>

<sup>3</sup><http://sepwww.stanford.edu/research/redoc/>



## SEP150 — TABLE OF CONTENTS

### Deconvolution and signal processing

<i>Jon Claerbout and Antoine Guitton</i> , Ricker-compliant deconvolution . . . .	1
<i>Stewart A. Levin, Jon Claerbout, and Eileen R. Martin</i> , Shortest path to whiteness . . . . .	13
<i>Stewart A. Levin</i> , OpenCPS experience in my summer seismic preprocessing workshop . . . . .	17

### Velocity estimation

<i>Biondo Biondi and Ali Almomin</i> , Simultaneous inversion of full data bandwidth by tomographic full waveform inversion (TFWI) . . . . .	21
<i>Xukai Shen</i> , Near-surface velocity estimation for a realistic 3D synthetic model . . . . .	43
<i>Ali Almomin and Biondo Biondi</i> , Hessian analysis of tomographic full waveform inversion operators . . . . .	53
<i>Musa Maharramov and Biondo Biondi</i> , Simultaneous time-lapse full waveform inversion . . . . .	63

### Imaging through inversion

<i>Mandy Wong, Biondo Biondi, and Shuki Ronen</i> , Joint-LSRTM in practice with the Deimos ocean bottom field data set . . . . .	71
<i>Chris Leader and Robert Clapp</i> , Designing an object-oriented library for large scale iterative inversion . . . . .	95

### Surface-waves and passive seismic

<i>Jason P. Chang</i> , Velocity dispersion at Long Beach . . . . .	101
<i>Sjoerd de Ridder, Dave Nichols, and Biondo Biondi</i> , Scholte-wave azimuthal-anisotropic phase-velocity images in the near surface at Ekofisk from seismic noise correlations . . . . .	109
<i>Marine Denolle, Sjoerd de Ridder, Jason P. Chang, Eileen R. Martin, Taylor Dahlke, Humberto Arevalo-Lopez, Sr., and Stewart A. Levin</i> , Scholte-wave excitation . . . . .	121

### Modeling and Anisotropy

<i>Huy Le, Robert G. Clapp, and Stewart A. Levin</i> , Equivalent accuracy at a fraction of the cost: Tackling spatial dispersion . . . . .	131
---	-----

*Yunyue (Elita) Li, Mandy Wong, and Robert Clapp*, Equivalent accuracy at a fraction of the cost: Overcoming temporal dispersion ... 141

*Yunyue (Elita) Li, Dave Nichols, and Gary Mavko*, Stochastic rock physics modeling for seismic anisotropy with two different shale models ..... 153

*Yunyue Li, Biondo Biondi, Robert Clapp, and Dave Nichols*, Wave equation migration velocity analysis for VTI models footnote[1]..... 161

**Image segmentation**

*Adam Halpert*, Integrated seismic image segmentation and efficient model evaluation: field data example..... 175

*Taylor Dahlke*, Distance Regularized Level Set Salt Body Segmentation . 187

SEP phone directory ..... 189

Research personnel ..... 191

(‘SEP article published or in press, 2013’,)..... 198

## Ricker-compliant deconvolution

*Jon Claerbout and Antoine Guitton*

### ABSTRACT

Ricker compliant deconvolution spikes at the center lobe of the Ricker wavelet. It enables deconvolution to preserve and enhance seismogram polarities. Expressing the phase spectrum as a function of lag, it works by suppressing the phase at small lags. A byproduct of this decon is a pseudo-unitary (very clean) debubble filter where bubbles are lifted off the data while onset waveforms (usually Ricker) are untouched.

### INTRODUCTION

A long-standing assumption in reflection seismology is that seismic sources are minimum phase (Robinson and Treitel, 1964). This means the source wavelet and its inverse (the decon filter) are both causal. The decon filter is derived from the data spectrum. Its inverse should be the source wavelet. These assumptions lead to the non-plausible source waveforms in Figure 2. We should be seeing something closer to the symmetrical Ricker wavelet (Ricker, 1953): let  $Z = e^{i\omega dt}$  where  $dt$  is the two way travel time from the gun (or the receiver) to the surface and back. Since the surface reflection coefficient is -1, the frequency response at the gun is  $1 - Z$  and at the hydrophone is also  $1 - Z$  making the composite response  $1 - 2Z + Z^2$ . Therefore, the Ricker wavelet results from a water-surface ghost at the marine gun convolved with another at the hydrophone (Lindsey, 1960). In all four of the deep-marine regions we tested and show in Figure 2, the minimum-phase source wavelets estimated with the Kolmogoroff method (see below) have the expected three lobes, but they are not symmetric. The first lobe is more than double the third. This is a long-standing issue in seismic processing (Levy and Oldenburg, 1982) and many methods based, for instance, on homomorphic transformations (Jin and Rogers, 1983), frequency domain filtering (Ghosh, 2000), higher order statistics (Sacchi and Ulrych, 2000) or stochastic methods (Velis, 2008) have been proposed to handle non minimum-phase wavelets.

In the conflict between the Ricker idea and the minimum-phase idea we take it here that the Ricker idea is closer to the truth (Rice, 1962). The main contribution of this paper is a method for finding phase that respects Ricker symmetry. The new time origin is at the center of the Ricker wavelet. With the new seismic wavelet, seismogram polarity will be seen more apparent. A byproduct of this approach is a debubble process giving results of outstanding clarity.

Our proposed method is a simple addition to the Kolmogoroff method in the lag-log domain (also known as the “cepstrum”). Parameterizing the logarithm of the spectrum in the time domain lays out parameter choices in a natural way along the “quefrequency” axis. Wavelets presented here and computed in  $N \log_2 N$  time from field data examples show the expected Ricker shape. We regard them as a final analytical stage before invoking iterative inverse theory for the wavelet estimation problem.

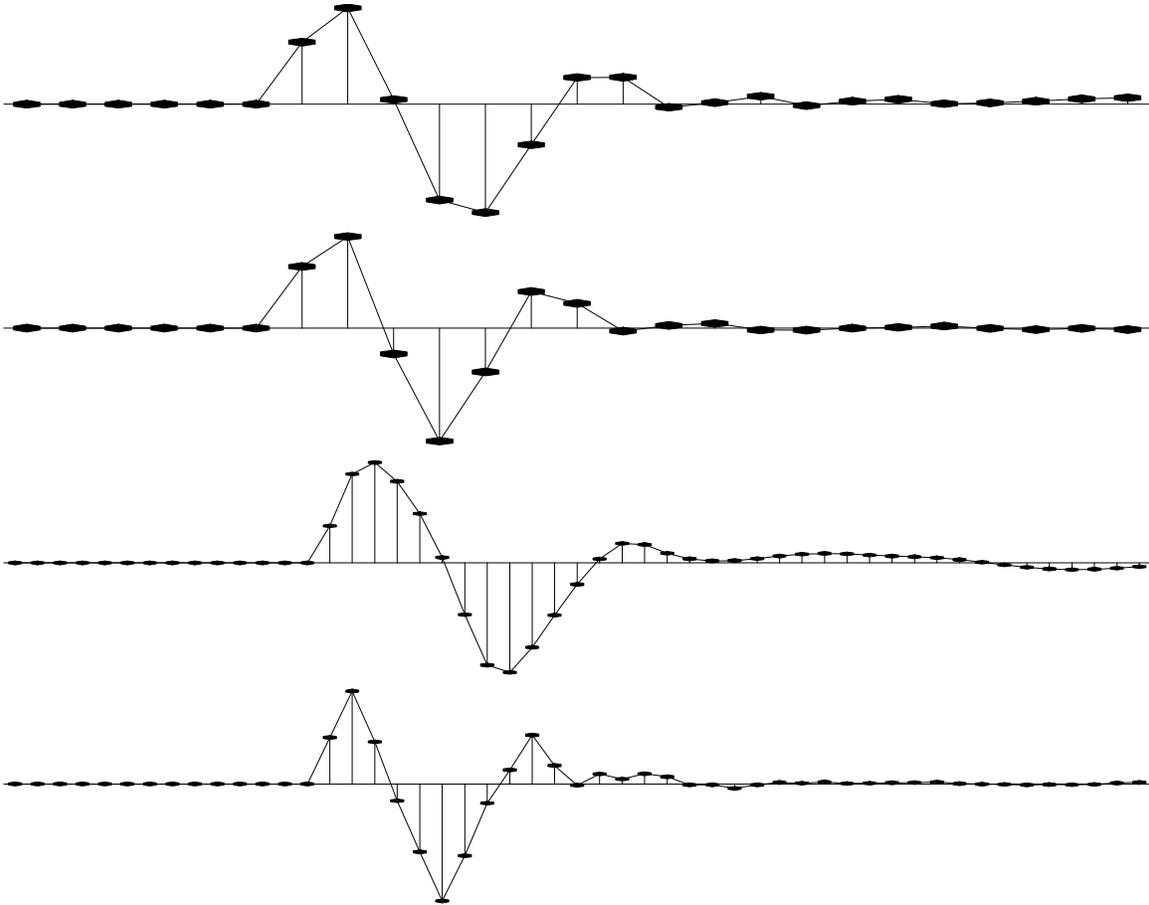


Figure 1: Estimated shot wavelets from four deep-marine regions when the minimum-phase assumption is made. They are not time symmetric about the middle lobe. Top two wavelets are from 4 ms data and the bottom two from 2 ms data. These results are based on data corrected for divergence. Spectra were averaged over many hundreds of inner-offset seismograms. These spectra were not smoothed. [NR] [jon1/. sources](#)

## BASICS OF LAG-LOG SPACE

Let  $F(\omega)$  be a filter in frequency domain. Let  $\ln F(\omega) = U(\omega)$  so the filter is  $F(\omega) = e^{U(\omega)}$ . With the definition  $Z = e^{i\omega\Delta t}$  Fourier transforms become polynomials ( $Z$ -transforms). Thus  $U(\omega)$  relates to the time function  $u_\tau$  by a Fourier sum  $U(\omega) = \sum_{\tau=0}^{\tau_{\max}} u_\tau Z^\tau$ . The  $u_\tau$  values will be our parameterization of the filter. Historically, the  $\tau$  axis is called the “quefrequency” axis though we sometimes refer to it as the “lag-log” axis.

The property of exponentials that  $e^{A+B+C} = e^A e^B e^C$  has an interesting meaning when we exponentiate a  $Z$ -transform  $\exp(A+B+C) = \exp(\sum_{\tau=1}^{\tau_{\max}} u_\tau Z^\tau)$ . The  $Z$ -transform sum may be split up into small lags, medium lags, and large lags. This decomposes a filter (or waveform) into a sequence of three filters, each with its own meaning where for example in common marine seismology:

$$\begin{aligned} e^{A+B+C} &= e^A e^B e^C & (1) \\ e^{\sum_{\tau=1}^{\tau_{\max}} u_\tau Z^\tau} &= e^{\sum_1^c} e^{\sum_{c+1}^r} e^{\sum_{r+1}^{\tau_{\max}}} & (2) \\ \text{(wavelet)} &= \text{(continuity)}(\text{Ricker})(\text{bubble}) & (3) \end{aligned}$$

where  $c$  and  $r$  define the boundaries of the continuity and Ricker regions, respectively. Equation (2) defines the boundaries of the three regions abruptly although in practice we blend them smoothly with sine-squared weighting. Changing the sign of  $(A+B+C)$  changes a filter to its inverse. Both are parameterized by the same  $A$ ,  $B$ , and  $C$ .

### Exploring $e^A$ : band limiting filter

Theoretically, the output of a prediction-error filter tends to be white. In practice, energy near the Nyquist frequency will be nearly all noise, so nobody wants to see a white output. For this reason deconvolution is generally followed by band-pass filtering. Equation 2 integrates the two processes, deconvolution and the band pass. The parameter  $c$  is analogous to the bandpass cutoff frequency, but it is a cutoff in lag hence resembles inverse frequency. Although we may hope to resolve some very short lags, we do not want our deconvolved data dominated by short lags. We have made no attempt to achieve a sharp cutoff, instead using a sine squared taper. Typically, we experiment with  $c$  and have taken  $c$  to correspond to about 7ms.

### Exploring $e^C$ : debubbling filter

We may specify  $u_\tau$  from prior knowledge, or from knowledge gained from various kinds of data averaging, or from some mixture of the two. Commonly we begin from Kolmogoroff spectral factorization (next section) giving us all the  $u_\tau$ . We may design a filter  $e^{A+B+C}$  by over-riding Kolmogoroff with  $A = 0$  and  $B = 0$ . Such a filter would do nothing to its inputs at small and intermediate lags but would affect longer lags. To see what happens, consider the filter  $e^C = 1 + C + C^2/2! + \dots \approx 1 + C$ . Examine its leading coefficients. They are  $(1, 0, 0, \dots, 0, u_{r+1}, u_{r+2}, u_{r+3}, \dots)$ . Figure 2 shows the application of such a filter with  $r = 15$ . This operation on the data is called “debubbling”. Debubbling in this manner seems to leave first arrivals untouched. The 15 interval gap on 4 ms data is 60 ms, a number

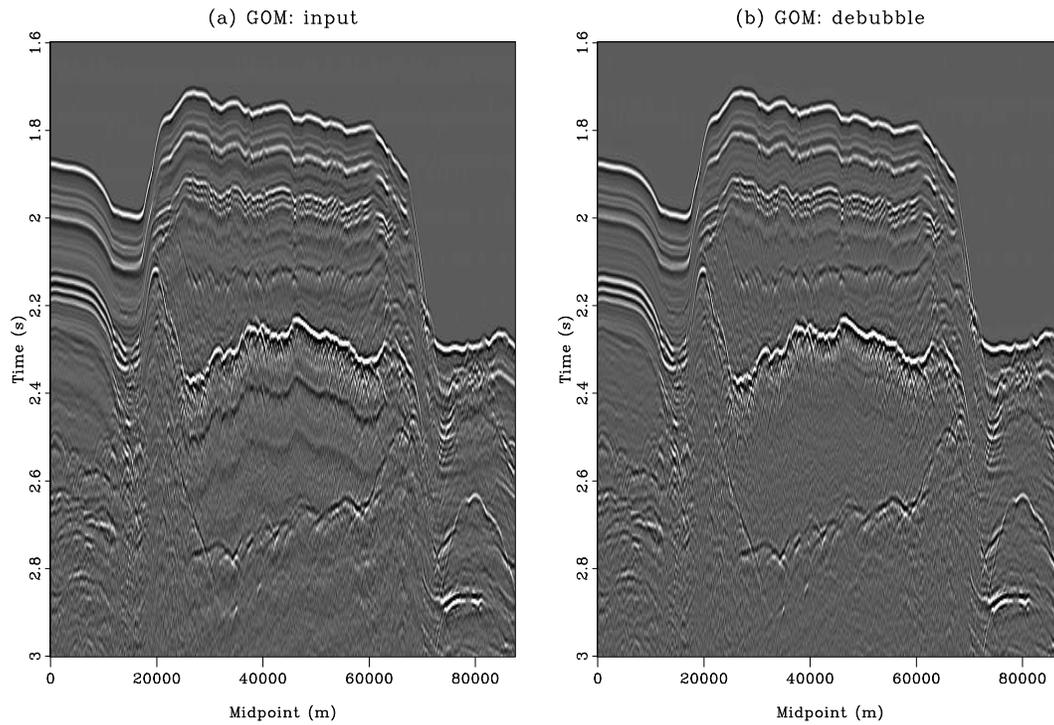


Figure 2: Gulf of Mexico data before (a) and after debubble (b). This process preserves the wave onset while it lifts off the bubbles. Here the effect of the process is visible nearly everywhere after 2.4 sec, but also visible around 1.85 sec. On blinking displays it is easy to see bubble removed nearly everywhere. [NR] [jon1/. GOM-unitary](#)

half way between the “end of the ghosts” and the “onset of the bubbles”. This result may be described as “textbook quality” (meaning it is the best we have ever produced).

Before we go on to attack the middle-lag terms we review the starting point, Kolmogoroff spectral factorization.

### Kolmogoroff spectral factorization

When a time function such as  $b_t$  vanishes at all negative time lags it is said to be causal. Its  $Z$  transform is  $B(Z) = b_0 + b_1Z + b_2Z^2 + b_3Z^3 + \dots$ . Observe that  $B(Z)^2$  is also causal because it has no negative powers of  $Z$ , alternately, because the convolution of a causal with a causal is causal. Likewise  $e^B$  is causal because it is a sum of causals.

$$e^B = 1 + B + B^2/2! + B^3/3! + \dots \quad (4)$$

Happily, this infinite series always converges because of the strong influence of the denominator factorials. The time-domain coefficients for  $e^B$  could be computed the hard way, putting polynomials into power series, or  $e^B$  may be computed by Fourier transforms. To do so, we would evaluate  $B(Z)$  for many real  $\omega$ , and then invoke an inverse Fourier transform program to uncover the time-domain coefficients.

Let  $r = r(\omega)$ ,  $\phi = \phi(\omega)$ , and  $Z^\tau = e^{i\omega\tau}$ . Let us investigate the consequence of exponentiating a causal filter.

$$|r|e^{i\phi} = e^{\ln|r|+i\phi} = e^{\sum_\tau b_\tau Z^\tau} \quad (5)$$

Notice a pair of filters, both causal and inverse to each other.

$$|r|e^{i\phi} = e^{+\sum_\tau b_\tau Z^\tau} \quad (6)$$

$$|r|^{-1}e^{-i\phi} = e^{-\sum_\tau b_\tau Z^\tau} \quad (7)$$

A filter from any such pair is said to be “minimum phase”. Many filters are not minimum phase because they have no causal inverse. For example the delay filter  $Z$ . Its inverse,  $Z^{-1}$  is not causal. Such filters do not relate to a causal complex logarithm. If they have a logarithm, it must be non causal.

Given a spectrum  $r(\omega)$  we can construct a minimum-phase filter with that spectrum. Since  $r(\omega)$  is a real even function of  $\omega$ , the same may be said of its logarithm. Let the inverse Fourier transform of  $\ln|r(\omega)|$  be  $e_\tau$ , where  $e_\tau$  is a real even function of time. Imagine a real odd function of time  $o_\tau$ .

$$|r|e^{i\phi} = e^{\ln|r|+i\phi} = e^{\sum_\tau (e_\tau + o_\tau)Z^\tau} \quad (8)$$

The phase  $\phi(\omega)$  transforms to  $o_\tau$ . We can assert causality by choosing  $o_\tau$  so that  $e_\tau + o_\tau = 0$  for all negative  $\tau$ . This defines  $o_\tau$  at negative  $\tau$ . Since  $o_\tau$  is odd, we also know its values at positive lags. This creates a causal exponent which creates a causal minimum-phase filter with the specified spectrum. Therefore the causal minimum-phase filter is simply obtained by multiplying  $e_\tau$  by a step function of height 2 to preserve the real part. This computation is called Kolmogoroff spectral factoring. The word “factoring” enters because in applications one begins with an energy spectrum  $|r|^2$  and factors it into an  $re^{i\phi}$  times its conjugate (time reverse).

### Ricker compliant decon

Start with the  $b_\tau$  resulting from a Kolmogoroff factorization described above. Split it into even and odd parts,  $u_\tau^{\text{odd}} = (b_\tau - b_{-\tau})/2$  and  $u_\tau^{\text{even}} = (b_\tau + b_{-\tau})/2$  whose sum is  $b_\tau$ . The even part Fourier transforms to the logarithm of the amplitude spectrum (i.e., equation 8). The odd part Fourier transforms to the phase spectrum. Here we monkey with the phase while not changing the amplitude. We simply taper  $u_\tau^{\text{odd}}$  towards zero for small lags. Figure 3a shows an illustration of an odd function together with the weighting function going smoothly to zero at negative lags in Figure 3b. The parameter  $\tau_a$  controls the amount of anticausality in the wavelet. Figure 4 examines the consequences of various numerical choices of  $\tau_a$  for the data in Figure 2a. In practice, the cepstrum of all traces is averaged before applying the decon. As we increase the anticausality, the time function of the wavelet  $e^B$  increases in symmetry near  $t = 0$ . Our favored choice is  $\tau_a = 64$  ms: it is larger than the Ricker width, about 20 ms, but not as large as the bubble delay, about 150 ms.

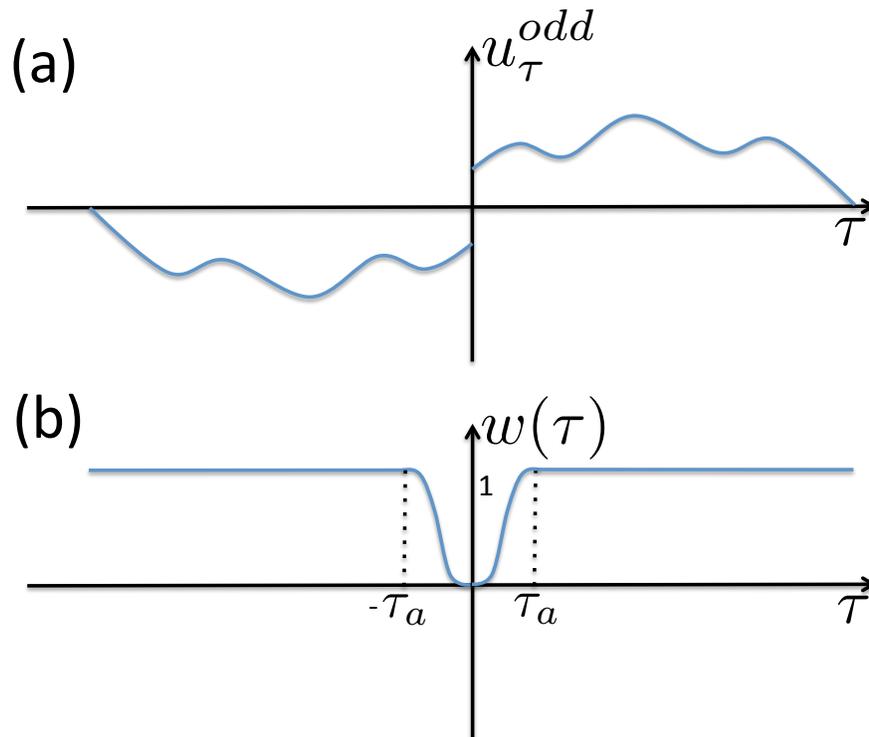


Figure 3: The odd part of  $b_\tau$  (a) and the weighting function (b). The parameter  $\tau_a$  determines the amount of anticausality in the wavelet. [NR] jon1/. anticausal

Removing all the phase of any wavelet makes it symmetric in time. Any deconvolution filter can be made Ricker-compliant simply by transforming its phase function from frequency to time lag. Suppressing the phase only near zero lag makes the wavelet symmetric only near zero lag (Ricker like).

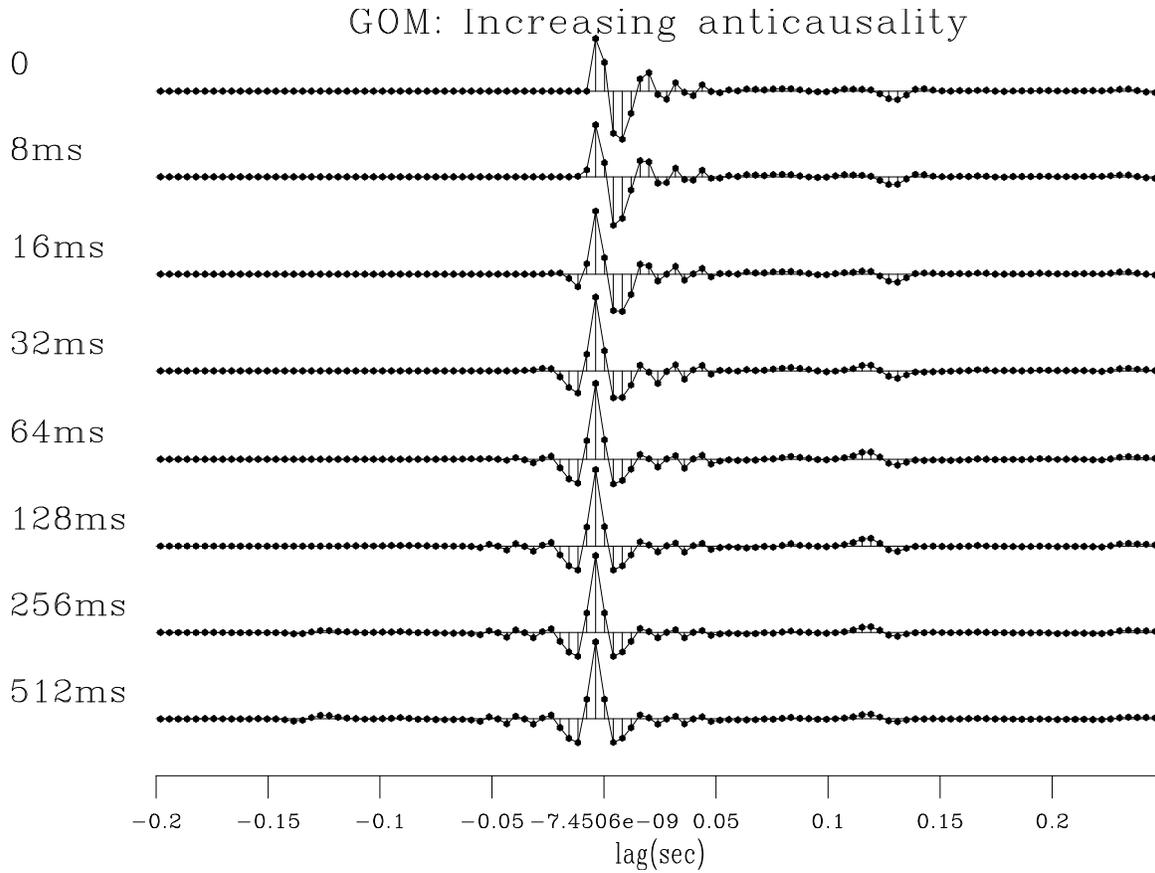


Figure 4: Gulf of Mexico  $\Delta t = 4$  ms: Increasing the anticausality  $\tau_a$  in Ricker compliant decon. Too much ( $\tau_a > 128$  ms) causes half the bubble to appear before the shot.  $\tau_a = 0$  ms corresponds to the minimum-phase wavelet. In this manuscript, our favored choice is  $\tau_a = 64$  ms. [NR] jon1/. GOM-laggard

The next section presents Ricker deconvolutions and debubbling examples on three field datasets. For all these examples, we use portions of the near-offset sections compensated for spherical divergence and attenuation (applying a  $t^2$  correction). All these results demonstrate the effectiveness of the proposed method for estimating Ricker-like, zero-phase wavelets as well as debubbling seismic data.

## RICKER DECONVOLUTION AND DEBUBBLING RESULTS

We first start by presenting results of the debubbling process on a dataset from offshore Baja California (also called “Cabo” in this manuscript) shown in Figure 5a. Debubbling (Figure 5b) is very effective as the bubbles are lifted off everywhere in the section (the first bubble arrives 150 ms after the main spike.)

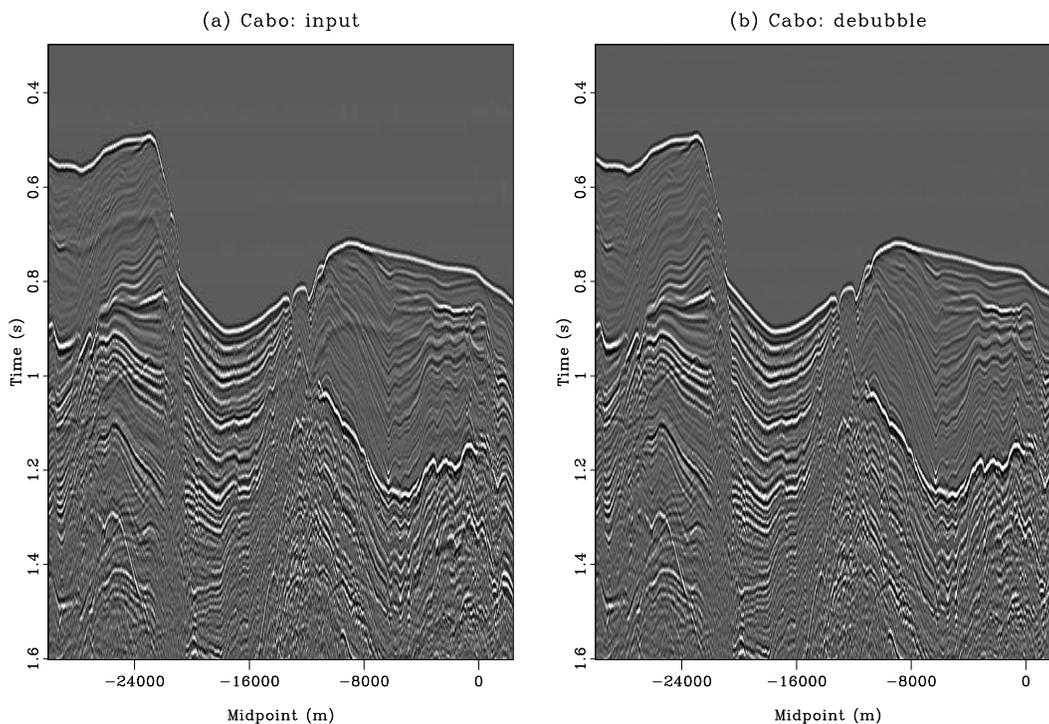


Figure 5: Cabo data before (a) and after debubble (b). This process preserves the wave onset while it lifts off the bubbles. [NR] [jon1/. Cabo-unitary](#)

Then, the Ricker deconvolution applied to the Gulf of Mexico dataset of Figure 2b is shown in Figure 6a with  $c = 0$  ms and in Figure 6b with  $c = 10$  ms. Increasing  $c$  acts as a bandpass filter by attenuating high-frequencies. The same filtering effect is seen with the Cabo dataset in Figure 7b. Notice in both Figures 6 and 7 how the deconvolution is able to remove the Ricker-like behavior of the main reflections. Figure 8 displays the different waveforms extracted from the Gulf of Mexico and Baja California datasets. Figures 8a and 8b show the Ricker wavelets (including the bubbles) extracted with our process. Figures 8c and 8d show the bubbles only ( $e^C$  in equation 2). The bubbles, deconvolved from the ghosts, show uniform polarity.

Finally, we apply the Ricker deconvolution to a near-offset section from offshore Aus-

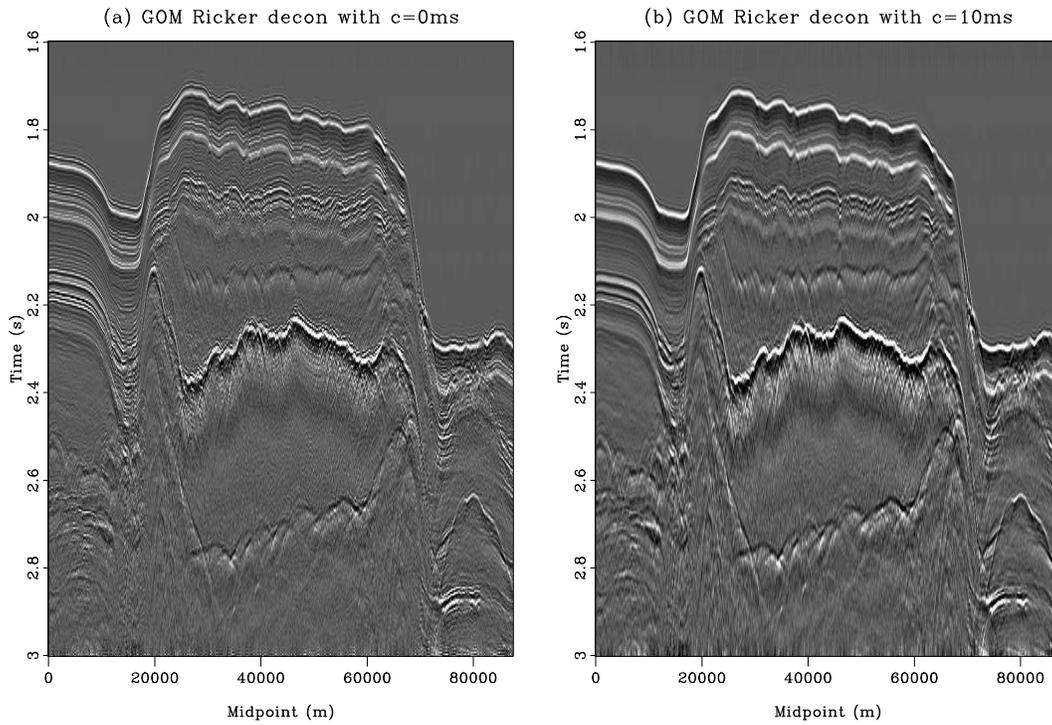


Figure 6: Ricker deconvolution results applied to the Gulf of Mexico dataset in Figure 2 for (a)  $c = 0$  ms and (b)  $c = 10$  ms. [NR] [jon1/. GOM-tresol2](#)

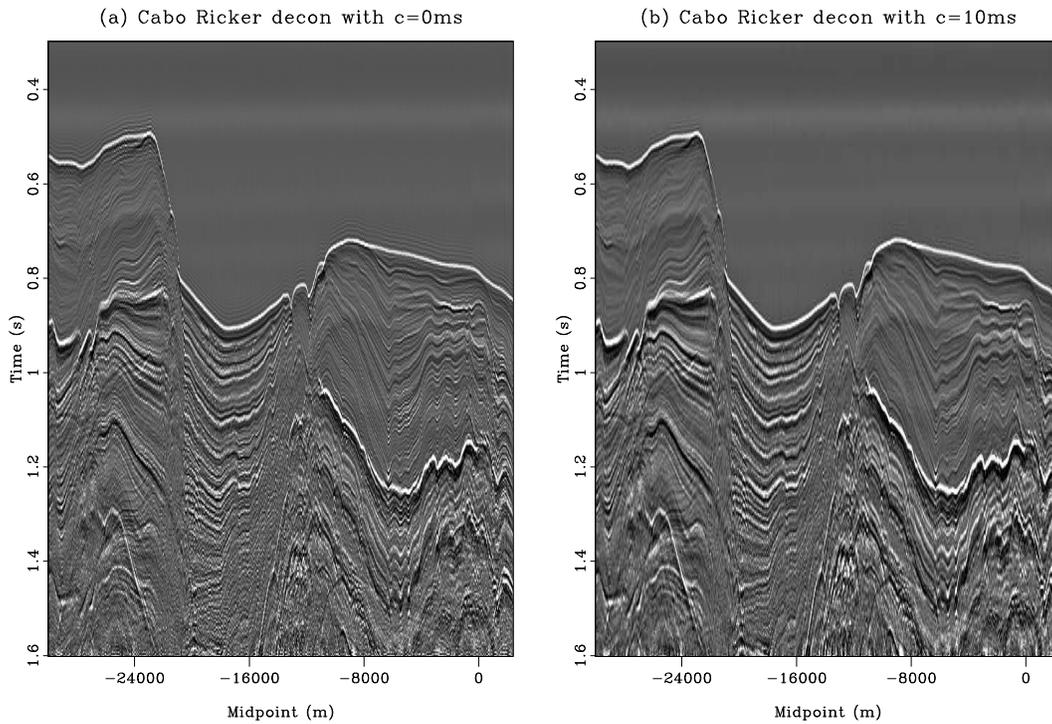


Figure 7: Ricker deconvolution results applied to the Cabo dataset for (a)  $c = 0$  ms and (b)  $c = 10$  ms. [NR] [jon1/. Cabo-tresol2](#)

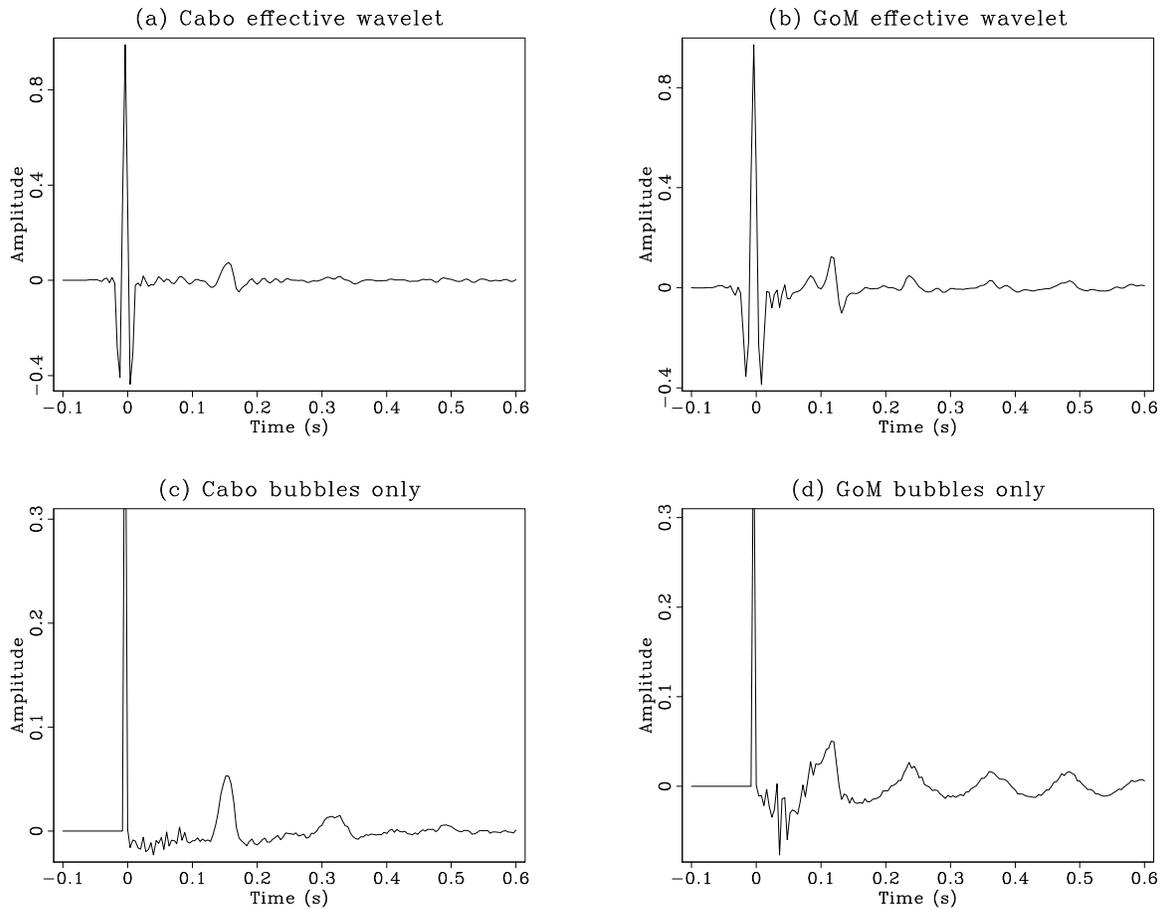


Figure 8: Bottom row shows the seismic source function after the ghosts have been removed (bubble only), top row before. Removing the ghosts reveals the uniform polarity of the bubbles. [NR] jon1/. derick

tralia. Figure 9a shows the input data, sampled at 2 ms. Figure 9b shows the deconvolution result (using our default  $c = 10$  ms). The Ricker wavelet and bubbles extracted are shown in Figure 10. Note how weak the bubbles are in this case, showing that the gun array was very well tuned (contrary to the other two field data examples, because the bubbles are so small, we decided against showing the debubbled section).

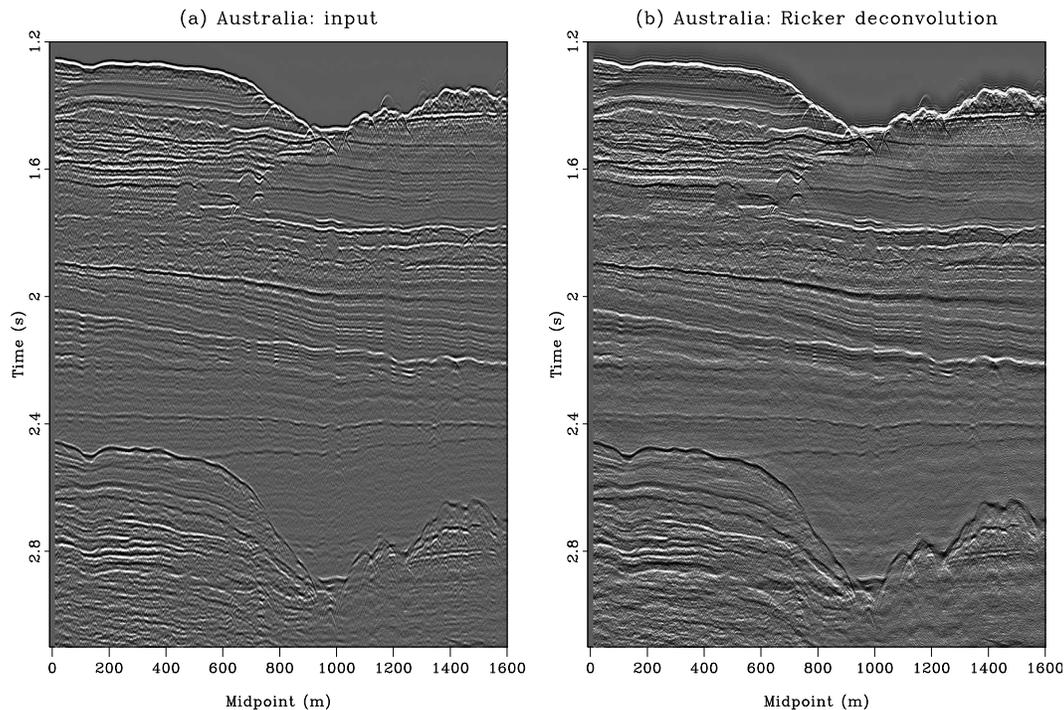


Figure 9: Dataset from offshore Australia (a) before and (b) after Ricker deconvolution.

[NR] `jon1/. cvx-decon`

## CONCLUSION

The proposed analytical method, a simple addition to the Kolmogoroff factorization, works well on the three data sets tested. Simply stated, phase, normally a function of frequency, is brought into the time lag domain and altered (suppressed) near zero lag. Additionally, it might provide an initialization and provide a regularization that should be helpful in fitting more complicated data models.

## REFERENCES

- Ghosh, S., 2000, Deconvolving the ghost effect of the water surface in marine seismics: *Geophysics*, **65**, 1831–1836.
- Jin, D. and J. Rogers, 1983, Homomorphic deconvolution: *Geophysics*, **48**, 1014–1016.
- Levy, S. and D. Oldenburg, 1982, The deconvolution of phase-shifted wavelets: *Geophysics*, **47**, 1285–1294.
- Lindsey, J., 1960, Elimination of seismic ghost reflections by means of a linear filter: *Geophysics*, **25**, 130–140.

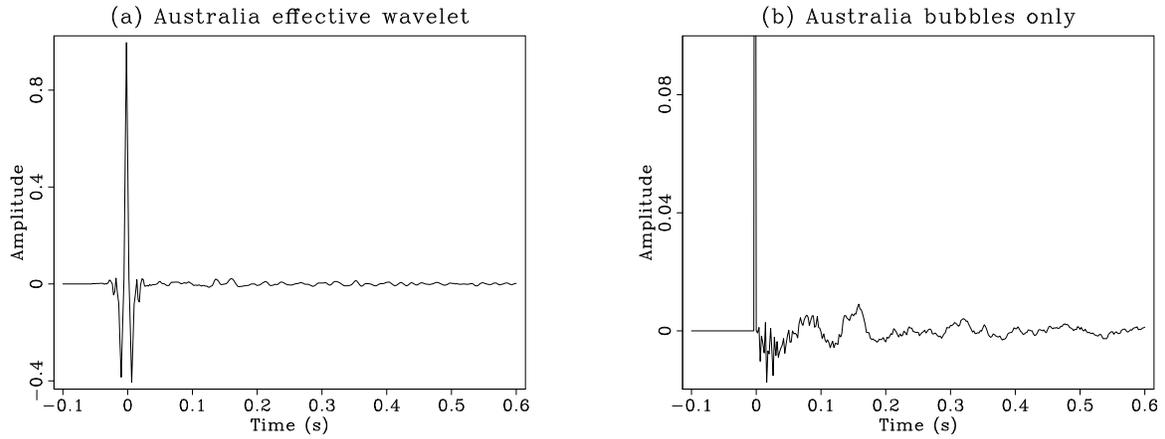


Figure 10: Waveforms extracted from the Australia datasets: (a) effective Ricker wavelet and (b) estimated bubbles. In this case, the bubbles are very weak. [NR] jon1/. derickcvx

Rice, R., 1962, Inverse convolution filters: *Geophysics*, **27**, 4–18.

Ricker, N., 1953, The form and laws of propagation of seismic wavelets: *Geophysics*, **18**, 10–40.

Robinson, E. and S. Treitel, 1964, Principles of digital filtering: *Geophysics*, **29**, 395–404.

Sacchi, M. and T. Ulrych, 2000, Nonminimum-phase wavelet estimation using higher order statistics: *The Leading Edge*, **19**, 80–83.

Velis, D., 2008, Stochastic sparse-spike deconvolution: *Geophysics*, **73**, R1–R9.

## Shortest path to whiteness

Stewart A. Levin, Jon Claerbout, and Eileen R. Martin

### ABSTRACT

*The output of a prediction error filter is white.* Easy to state, annoyingly hard for students to understand. We provide here two short, clean paths to that understanding.

### INTRODUCTION

The basic idea of least-squares fitting is that the residual is orthogonal to each of the fitting functions. Applied to the prediction error filter (PEF) this idea means that the output of the PEF is orthogonal to lagged inputs. The orthogonality applies only for lags in the past, because prediction knows only the past while it aims to the future. What we soon see here is different, namely, that the output is uncorrelated with *itself* (as opposed to the input) for lags in *both* directions; hence the output spectrum is white. An explication of these facts had been included in Image Estimation by Example (Claerbout and Fomel, 2013), but was found to confuse many readers. Here we revise that explanation, starting with the infinite case, for which we provide two simple proofs, one using  $Z$ -transforms and the other the cepstral lag-log domain.

#### Hilbert space proof

Let  $\mathbf{d}$  be a vector whose components contain a time-dependent function. Let  $Z^n \mathbf{d}$  represent shifting the components to delay the data in  $\mathbf{d}$  by  $n$  samples. The definition of a prediction-error filter (PEF) is that it minimizes  $\|\mathbf{r}\|_2^2$  by adjusting filter coefficients  $a_m$  in the residual

$$\mathbf{r} = \mathbf{d} + a_1 Z^1 \mathbf{d} + a_2 Z^2 \mathbf{d} + \dots \quad (1)$$

We set out to choose the best  $a_m$  by setting to zero the derivative of  $\frac{1}{2}\|\mathbf{r}\|_2^2 = \frac{1}{2}\mathbf{r} \cdot \mathbf{r}$  by  $a_m$ . After the best  $a_m$  are chosen, the residual is perpendicular to each of the fitting functions:

$$\begin{aligned} 0 &= \frac{1}{2} \frac{d}{da_m} (\mathbf{r} \cdot \mathbf{r}) \\ &= \mathbf{r} \cdot \frac{d\mathbf{r}}{da_m} = \mathbf{r} \cdot Z^m \mathbf{d} \quad . \end{aligned}$$

Thus, for any fixed  $k$  in  $\mathbb{Z}^+$  (the set of positive integers)

$$\begin{aligned} \mathbf{r} \cdot Z^k \mathbf{r} &= \mathbf{r} \cdot (Z^k \mathbf{d} + a_1 Z^{k+1} \mathbf{d} + a_2 Z^{k+2} \mathbf{d} + \dots) \\ &= \mathbf{r} \cdot Z^k \mathbf{d} + a_1 \mathbf{r} \cdot Z^{k+1} \mathbf{d} + a_2 \mathbf{r} \cdot Z^{k+2} \mathbf{d} + \dots \\ &= 0 + a_1 0 + a_2 0 + \dots \\ &= 0 \quad . \end{aligned}$$

Since the autocorrelation is symmetric,  $\mathbf{r} \cdot Z^{-k}\mathbf{r}$  is also zero for  $k \in \mathbb{Z}^+$ , so the autocorrelation of  $\mathbf{r}$  is an impulse. In other words, the spectrum of the time function  $r_t$  is white. Thus  $\mathbf{d}$  and  $\mathbf{a}$  have mutually inverse spectra.

Since the output of a PEF is white, the PEF itself has a spectrum inverse to its input.

Many applications are found in Claerbout's online books<sup>1</sup>. There are one dimensional examples with both synthetic data and field data in "EARTH SOUNDINGS ANALYSIS: Processing versus Inversion (PVI)," including the `gap` parameter (dead space between the initial impulse and the adjustable coefficients). Multidimensional examples are found in his online book "IMAGE ESTIMATION BY EXAMPLE: Geophysical Soundings Image Construction."

But what happens if we only solve for a finite number of terms for our prediction error filter? Obviously, we can't guarantee a perfectly impulsive residual autocorrelation. Instead we have some terms that aren't guaranteed by the least squares fit to be orthogonal to the residual. In most applications such terms tend to be small. The reason is in most applications predictions tend to degrade with time lag. There are exceptions, however. To predict unemployment next month, it helps a lot to know the unemployment this month. On the other hand, because of seasonal effects, the unemployment from a year before next month (11 months back) might provide even better prediction. But mostly, older data has diminishing ability to enhance prediction.

Finite difference equations resemble PEFs, and they use only a short range of lags, for example, a wave equation containing only the three lags intrinsic to  $\partial^2/\partial t^2$ . So, short PEFs are often powerful.

## Phase space proof

Here we specialize the arguments in an earlier paper (with a more complicated model) by Claerbout et al. (2012a) (also SEP-147 (2012b)) to supply an alternate PEF whiteness proof.

The data is again  $d_t$  while adjustable model parameters are  $u_\tau$ , initially  $u_\tau = 0$ . The forward modeling operation acts upon data  $d_t$  (in the Fourier domain  $D(Z)$  where  $Z = e^{i\omega}$ ) producing deconvolved data  $r_t$  (the residual).

$$r_t = \text{FT}^{-1} D(Z) e^{\dots+u_2Z^2+u_3Z^3+u_4Z^4+\dots} \quad (2)$$

$$\frac{dr_t}{du_\tau} = \text{FT}^{-1} D(Z) Z^\tau e^{\dots+u_2Z^2+u_3Z^3+u_4Z^4+\dots} \quad (3)$$

$$\frac{dr_t}{du_\tau} = r_{t+\tau} \quad (4)$$

The last step follows because  $Z^\tau$  simply shifts the data  $D(Z)$  by  $\tau$  units which shifts the residual the same. An output formerly at time  $t$  gets moved to time  $t + \tau$ . This result may look familiar, but it is not. The familiar result is that the derivative of a filter output with

---

<sup>1</sup><http://sep.stanford.edu/sep/prof/>

respect to the filter coefficient at lag  $\tau$  is the shifted input  $d_{t+\tau}$ , not the shifted output  $r_{t+\tau}$  we see above.

The power series definition of our exponential tells us constraining  $u_0 = 0$  assures the PEF begins with a “1”. Hence  $\Delta u_0 = 0$ . To find the update direction at nonzero lags  $\Delta \mathbf{u} = (\Delta \mathbf{u}_\tau)$  take the derivative of  $\sum_t r_t^2/2$  by  $u_\tau$ .

$$\Delta \mathbf{u} = \sum_t \frac{1}{2} \frac{dr_t^2}{du_\tau} \quad \tau \neq 0 \quad (5)$$

$$= \sum_t \frac{dr_t}{du_\tau} r_t \quad \tau \neq 0 \quad (6)$$

$$\Delta \mathbf{u} = \sum_t r_{t+\tau} r_t \quad \tau \neq 0 \quad (7)$$

At the end of the iteration, the gradient of  $\|\mathbf{r}\|_2^2 \rightarrow \mathbf{0}$ . Thus  $\Delta \mathbf{u}$  vanishes and we see that the residual is orthogonal to itself shifted, i.e. the residual is white.

The SEG expanded abstract and SEP-147 article mentioned above generalize this result to hyperbolic penalty functions. It is also generalized there to echo data with gain increasing with time.

## REFERENCES

- Claerbout, J., Q. Fu, and A. Guitton, 2012a, Decon in the log domain with variable gain: SEG Technical Program Expanded Abstracts, **31**, 1–5.
- Claerbout, J., A. Guitton, and Q. Fu, 2012b, Decon in the log domain with variable gain.
- Claerbout, J. F., 2004, Earth Soundings Analysis: Processing versus Inversion: accessed 4 Jun. 2010; [sep.stanford.edu/data/media/public/sep/prof/pvi.pdf](http://sep.stanford.edu/data/media/public/sep/prof/pvi.pdf).
- Claerbout, J. F., and S. Fomel, 2013, Image Estimation by Example: Geophysical Soundings Image Construction: accessed 7 Apr. 2013; [sep.stanford.edu/data/media/public/sep/prof/geeClass3-2013.pdf](http://sep.stanford.edu/data/media/public/sep/prof/geeClass3-2013.pdf).



## OpenCPS experience in my summer seismic preprocessing workshop

*Stewart A. Levin*

### ABSTRACT

Many students (and faculty) throughout the Earth Sciences have at least an occasional need to deal with seismic data. To help them handle some basic data import, preparation and, as is common here at Stanford, export, I ran a seven session workshop over the summer. During the course of the workshop, I exposed students to multiple packages: Seismic Unix, ProMAX/SeisSpace, and, the very recently arrived, OpenCPS. In this report, I discuss our experience with that last package, both as a vehicle for allowing students to easily focus on the geophysics and for the capabilities it provides the casual user.

### INTRODUCTION

Ever since I returned to Stanford two years ago, I have been helping students and faculty import and prepare seismic data for use in their research. With many students around over the summer and not tied up with classes, I decided to offer a summer hands-on workshop (Fig. 1) for those who might be interested. About ten students showed up for some or all of the sessions, originating from SEP, borehole seismology, crustal seismology and even petroleum (um Energy Resource) engineering.

During the course of planning and preparation, one of Simon Klemperer's students informed me that he had recently obtained a license for Open Geophysical Inc.'s OpenCPS package and had found it easy to use and very helpful in understanding a seismic dataset he had recently recorded in the field. This package was designed and developed by a group headed up by Roland Gunther, a former Stanford student who had done undergraduate work with Simon Klemperer and graduate work at the Stanford Exploration Project. We were able to transfer the license to a machine that the students could all access, allowing me to use the software as part of the workshop exercises and instructions.

As detailed writeups of the sessions are available from my SEP website under [sep.stanford.edu/oldsep/stew/Summer2013Workshop/](http://sep.stanford.edu/oldsep/stew/Summer2013Workshop/), I will limit my discussion to the highlights or lowlights of that experience and refer the interested reader to my website for fuller details.

### OPENCPS

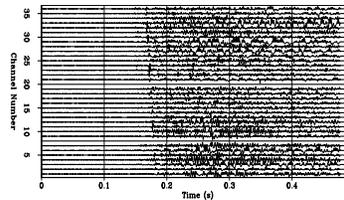
The "CPS" in OpenCPS refers to the former ConocoPhillips seismic processing system subsequently converted to an open-source version at [sourceforge.net/projects/cpseis/](http://sourceforge.net/projects/cpseis/). These modules formed the initial basis for the core geophysical methods of OpenCPS, though the

## Summer Workshop on Seismic Data Preprocessing

Instructor: Stewart A. Levin

Where: Mitchell A-65

When: Tuesdays 10 AM



Purpose: Assist students and faculty with decoding and preprocessing of seismic datasets, including exposure to various software tools available to the School of Earth Sciences.

Organizational Meeting: June 25<sup>th</sup>

Flexible syllabus includes:

- Basic terms and units (sample rate, stacking, coordinate reference system, velocity, ...)
- U CN DCD SSMC DAT (how to dump/scan and inspect binary seismic formats)
- Geometry (importing and applying auxiliary files)
- Visualization
- Filters and Deconvolution
- Velocity Analysis
- Stacking and Signal Enhancement

Figure 1: Workshop announcement [stew2/. SummerPreprocessingWorkshop](http://stew2/. SummerPreprocessingWorkshop)

user parameter interface to them has, thankfully, been almost totally replaced by Open Geophysical, which built its user interface framework on top of the Eclipse<sup>1</sup> open source development and application framework.

## Installation

Installation of OpenCPS requires two things: a compressed tar file and a license file. The compressed tar is downloaded from the [opengeophysical.com](http://opengeophysical.com) website and expanded into any chosen installation directory. The website download is updated quite frequently to include any new modules and bug fixes they've vetted. At our site, the system administrator created a symbolic link from `/usr/local/bin/opencps` to the startup script of that same name in the installation directory. After placing the license file into the installation directory, no further configuration is required to invoke the `opencps` command, although supplemental batch queue configuration would be useful for large scale processing such as migration of 3D prestack marine data.

## Initialization

When the user interface first comes up, there are a few preferences and defaults that can usefully be set up, the most important being the default location for where seismic projects should be created. Since those projects will contain volumes of seismic data, the initial choice of the user's home directory should be overridden. (When a new project is created, you can, of course, override the default.) In my summer workshop, several students initially got stuck because of storage quota limits on their home directories. I got stuck later on because the package sometimes creates large temporary work files under `$HOME/.eclipse`. (That was fixed by moving `.eclipse` elsewhere and making a symbolic link from my home directory to the new spot.)

## Overall Impressions

OpenCPS advertises itself as *Interactive Seismic Processing* and we all found it quite easy to set up a project, use the canned flow wizards to quickly run through most basic preprocessing steps, and view the results with just a right click on the relevant seismic dataset or table.

When we ran into problems, mostly the result of our academic system configuration, the Open Geophysical developers were quick to address them, either making quick bug fixes or giving us workarounds. They were also very responsive to suggestions for clarifying user interface verbiage that confused us novices.

The development environment was sufficient for us to write a deconvolution tool or two with only minor glitches in the installation process. (You clone one of their examples, tell the user interface where your code and libraries reside, and hit Refresh on the Tools list.)

The OpenCPS system wears two hats—Batch and Interactive. The difference is that Batch runs a flow and writes output, Interactive displays data instead of writing output and repeatedly runs the flow in response to changes in flow tool parameters. In addition,

---

<sup>1</sup>[www.eclipse.org](http://www.eclipse.org)

the Interactive view provides quick ways to bounce around through even a massive dataset, allowing you to fine tune parameters and tables before running the whole input through the flow. I found this quite useful in playing around with the parameters in my port of Jon's Ricker-compliant decon. I haven't exercised the parallel capabilities of the Batch view, in particular, the cluster job replication template mechanism, so I cannot comment on that area. I do note, though, that OpenCPS definitely takes advantage of multiple cores to very nicely utilize available I/O bandwidth.

One area I found annoying was seismic data gain control. All too often when I switched the display parameters (exposed by clicking on an arrow at the top right of the plot window) from even something simple like smooth interpolation to nearest neighbor interpolation, I had to manually regain by as much as a factor of 5.

Another negative is the limited number of tools, presumably because their open source business model incorporates upselling more sophisticated modules, and the buggy/limited interface to the Seismic Unix tool list they provide. For us these are not really an issue for the main purpose of importing and cleaning up data before export to SEG-Y, SEP, or, perhaps, Matlab.

## CONCLUSIONS

Overall, OpenCPS works quite nicely to provide the limited list of capabilities and the flexible interactivity that most users here might desire. The folks who have been using ProMAX (but not SeisSpace, I rush to clarify) found their work moving along more rapidly under OpenCPS. It still has some room for production hardening, but overall, it nicely fills the basic preprocessing niche that I targeted in my workshop.

# Simultaneous inversion of full data bandwidth by tomographic full waveform inversion (TFWI)

*Biondo Biondi and Ali Almomin*

## ABSTRACT

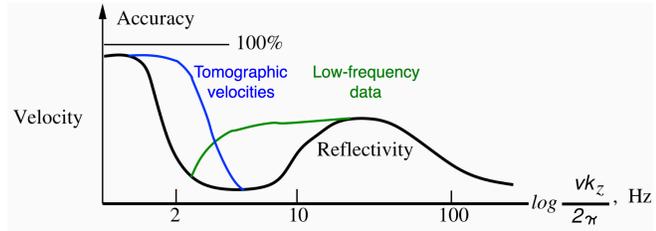
Convergence of full waveform inversion can be improved by extending the velocity model along either the subsurface-offset axis or the time-lag axis. The extension of the velocity model along the time-lag axis enables us to linearly model large time shifts caused by velocity perturbations. This linear modeling is based on a new linearization of the scalar wave equation where perturbation of the extended slowness-squared is convolved in time with the second time derivative of the background wavefield. The linearization is accurate for both reflected events and transmitted events. We show that it can effectively model both conventional reflection data as well as modern long-offset data containing diving waves. It also enables the simultaneous inversion of reflections and diving waves, even when the starting velocity model is far from being accurate. We solve the optimization problem related to the inversion with a nested algorithm. The inner iterations are based on the proposed linearization and on a mixing of scales between the short and long wavelength components of the velocity model. Numerical tests performed on synthetic data modeled on the Marmousi model and on the “Caspian Sea” portion of the well-known BP model demonstrate the global-convergence properties as well the high-resolution potential of the proposed method.

## INTRODUCTION

Conventional seismic imaging relies on a separation of scales between the migration velocity model (long-wavelength components) and the reflectivity (short-wavelength components). Figure 1 shows a simplified 1D graphical representation of the separation of scales concept. The black line represents the two disjointed wavelength ranges (mapped into corresponding temporal-frequency bands) and the consequent gap in information between long wavelengths and short wavelengths. This conceptual understanding leads naturally to a sequential approach for seismic imaging; the velocity model is estimated first, and then it is used as input to migration for imaging reflectivity. In current velocity-estimation practice, reflectivity is used only indirectly to measure the focusing power of the velocity model. The only important exceptions occur when migrated volumes are used to interpret boundaries of geobodies (e.g. salt bodies) and to estimate predominant dips in the geologic layering that are then used to constraint a tomographic velocity update.

The sequential imaging process is slowly being undermined by three long-standing trends in the industry: 1) acquisition of lower-frequency data, 2) imaging under a complex overburden which requires higher-resolution velocity models to focus and correctly position reflectors, and 3) acquisition of longer-offset data. As the industry strives to widen the data

Figure 1: Simplified 1D graphical representation of the separation of scales in seismic imaging (black line) and how current industry trends are narrowing the gap between the estimation of long wavelengths and short wavelengths (blue and green lines). (Adapted from Jon Claerbout's *Imaging the Earth Interior*. [NR] [biondi1/.iei-rely-new-ppt](http://biondi1/.iei-rely-new-ppt))



frequency band at both the low and high end, the reflectivity band is extended at the low end, as graphically represented by the green line in Figure 1. The high end of the velocity band is also pushed upward (blue line in Figure 1) by the application of sophisticated tomographic methods that enable the estimation of the high-resolution velocity models needed to focus reflectors located under complex overburden. However, tomography (either ray or wave-equation based) is a more challenging task than migration, and thus often it falls short of providing the accuracy and resolution necessary to satisfactorily image the high frequencies in the data. The acquisition of longer-offset data enables the recording of diving waves and refracted arrivals that provide a complementary illumination of the velocity components in the crucial scale gap and blur the distinction between migration and tomography since they contain forward-scattering perturbations to transmission events.

As the information gap narrows, imaging methods that simultaneously estimate the velocity and reflectivity model by taking advantage of all the information in the data are becoming more attractive. The renewed interest in full waveform inversion (FWI) (Bamberger et al., 1982; Tarantola, 1984) could be mostly explained as an attempt to overcome the limitations imposed by the sequential imaging approach, as well as the availability of the computational power sufficient for practical applications of FWI. FWI has been the most successful when applied to the low frequencies in the data (green line in Figure 1) to improve the velocity-model estimation needed for imaging the high frequencies in the data under complex overburdens. FWI has been less successful in using the high-frequencies in the data to tomographically estimate the long-wavelengths in the model.

Since the 1980s it has been recognized that FWI has both a migration component and a tomographic component (Mora, 1989). However, to ensure convergence of the tomographic component the recorded and modeled data must be almost in phase with each other; the rule of thumb being that the residual time-shifts must be shorter than the half period of the dominant frequency in the data. Bootstrapping the inversion by starting from the low frequencies may ameliorate the convergence problems, but it still depends on conventional velocity estimation methods to deliver starting models sufficiently accurate to satisfy the convergence criterion for the FWI tomographic component. It also undermines the goal of simultaneous estimation because the high frequencies in the data contain the high-resolution tomographic information that facilitates the estimation of the velocity components represented by the blue line in Figure 1.

To perform true simultaneous and synergistic inversion for all the model scales we must

address the convergence problems of the tomographic term in FWI. These convergence problems are related to the non-linearity of the wave-equation with respect to perturbations in the long wavelengths of the velocity model. Long-wavelength perturbations cause substantial time shifts of the propagating wavefields that are poorly approximated by the linearization of the wave equation based on the first order Born approximation. In this paper we introduce a linearization of the wave equation based on an extension of the velocity model along the time lag axis ( $\tau$ ). This extension enables the linear modeling of large time shifts in the propagating wavefields, and consequently in the data. Based on this extension we define an objective function that has a model-focusing term in addition to the conventional FWI data-fitting term. Numerical examples with realistically complex velocity models demonstrate that this objective function has excellent convergence behavior, although currently convergence is unsatisfactorily slow.

The usefulness of extending the reflectivity model (as prestack images in the angle or offset domain) to manage the non-linearities in wave-equation velocity analysis was demonstrated in the context of differential semblance optimization (DSO) (Symes and Carazzone, 1991; Shen and Symes, 2008) and wave-equation migration velocity analysis (WEMVA) (Biondi and Sava, 1999; Sava and Biondi, 2004a,b). The generalization of this idea to the extension of the propagation component of the velocity model (long wavelength) was first introduced by Symes (2008) and more recently successfully applied by Sun and Symes (2012); Almomin and Biondi (2012); Biondi and Almomin (2012). These methods are based on a velocity extension along the subsurface-offset or plane-wave ray-parameter axes. In this paper, we introduce an extension along the time lag axis ( $\tau$ ) because it is better suited to describe the large time shifts in wave propagation that are at the root of FWI convergence problems. Furthermore, extending the velocity along the time lag axis can easily handle forward-scattered events recorded at long offsets as well as the reflections recorded at near and intermediate offsets. We believe that the time-lag formulation has the potential to deliver high-quality results also for modern long-offset data sets. Furthermore, a one-dimensional extension along time is computationally more efficient than a two-dimensional extension along horizontal subsurface offsets. Yang and Sava (2009, 2010) have discussed the use and the computational advantages of time-lag gathers to perform WEMVA for reflected events.

## TOMOGRAPHIC FULL WAVEFORM INVERSION (TFWI)

Conventional full waveform inversion is performed by solving the following optimization problem

$$\min_{\mathbf{s}^2} J_{\text{FWI}}(\mathbf{s}^2), \quad (1)$$

where:

$$J_{\text{FWI}}(\mathbf{s}^2) = \frac{1}{2} \|\mathcal{L}(\mathbf{s}^2) - \mathbf{d}\|_2^2, \quad (2)$$

$\mathbf{s} = s(\vec{x})$  is the slowness vector,  $\mathcal{L}$  is a wave-equation operator nonlinear with respect to slowness perturbations. The data vector  $\mathbf{d}$  is the pressure field  $\mathbf{P} = P(t, \vec{x})$  measured at the surface through a sampling linear operator  $\mathbf{S}$ ; such as  $\mathbf{d} = \mathbf{S}\mathbf{P}$

The wave-equation operator is usually evaluated by recursively solving the following finite difference equation

$$[\mathbf{s}^2 \mathbf{D}_2 - \nabla^2] \mathbf{P} = \mathbf{f}, \quad (3)$$

where  $\mathbf{D}_2$  is a finite-difference representation of the second derivative in time,  $\nabla^2$  is a finite-difference representation of the Laplacian, and  $\mathbf{f}$  is the source function.

The most efficient solution of the optimization problem expressed in equation 1 is performed by gradient based methods, and thus requires the evaluation of the linear operator  $\mathbf{L}$ , which is the linearization of  $\mathcal{L}$  with respect to slowness perturbations  $\delta\mathbf{s}^2$ . This linear operator can be derived by perturbing equation C-1 as follows

$$[(\mathbf{s}_o^2 + \delta\mathbf{s}^2) \mathbf{D}_2 - \nabla^2] (\mathbf{P}_o + \delta\mathbf{P}) = \mathbf{f}, \quad (4)$$

where  $\mathbf{P}_o$  and  $\mathbf{s}_o$  are the background wavefield and slowness, respectively, and  $\delta\mathbf{P}$  is the scattered wavefield.

Equation 4 can be rewritten as the following coupled equations:

$$[\mathbf{s}_o^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_o = \mathbf{f}, \quad (5)$$

$$[\mathbf{s}_o^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P} = \delta\mathbf{s}^2 \mathbf{D}_2 (\mathbf{P}_o + \delta\mathbf{P}), \quad (6)$$

which represents a nonlinear relationship between the slowness perturbations and the scattered wavefield. In conventional FWI, to linearize this relationship we drop the term multiplying the perturbations with each other; that is, we drop the scattered wavefield from the right-hand-side of equation 6 and obtain the following coupled equations:

$$[\mathbf{s}_o^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_o = \mathbf{f}, \quad (7)$$

$$[\mathbf{s}_o^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P} = \delta\mathbf{s}^2 \mathbf{D}_2 \mathbf{P}_o. \quad (8)$$

The linear operator  $\mathbf{L}$  is evaluated by recursively propagating the background wavefield  $\mathbf{P}_o$  and the scattered wavefield  $\delta\mathbf{P}$  by solving equations 7–8. Its adjoint operator  $\mathbf{L}'$ , which is needed to compute the gradient of the FWI objective function 1, is evaluated by backward propagating the scattered wavefield solving equation 8 using the data residuals as boundary conditions, and evaluating the zero time lag of the cross-correlation between the background and scattered wavefields.

Equations 7–8 define a linear relationship between  $\delta\mathbf{P}$  and  $\delta\mathbf{s}^2$ ; however, they cannot model large time shifts between the background wavefield and the scattered wavefield. These large time shifts are correctly modeled by the nonlinear equations 5–6 through multiple scattering, that is, by the accumulation of time shifts into the scattered wavefield which enter in the expression of the virtual sources injected by the right-hand-side of equation 6. When we drop the scattered wavefield from equation 6, we prevent this accumulation of large time shifts into the scattered wavefield.

These observations suggest that a simple method to improve the capability of the linearization to model large time shifts is to introduce time shifts directly into the slowness-perturbations term in the right-hand-side of equation 8. We extend the slowness model along the time lag axis  $\tau$  and convolve its perturbations  $\delta\tilde{\mathbf{s}}^2(\tau)$  with the second time derivative of the background wavefield; we rewrite equation 8 as:

$$[\tilde{\mathbf{s}}_o^2(\tau=0) \mathbf{D}_2 - \nabla^2] \delta\mathbf{P} = \delta\tilde{\mathbf{s}}^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \mathbf{P}_o, \quad (9)$$

where  $\overset{\tau}{*}$  denotes convolution in  $\tau$ . Onwards we use the tilde sign above operators and model vectors to denote their extension along the time-lag axis.

Equations 7 and 9 define the operator  $\tilde{\mathbf{L}}(\tilde{\mathbf{s}}(\tau = 0))$  which is linear with respect to  $\delta\tilde{\mathbf{s}}^2(\tau)$ , but nonlinear with respect to  $\tilde{\mathbf{s}}^2(\tau = 0)$ . The combination of the wave equation operator  $\mathcal{L}(\mathbf{s})$  and of  $\tilde{\mathbf{L}}$  defines the extended nonlinear operator

$$\tilde{\mathcal{L}}(\tilde{\mathbf{s}}) = \mathcal{L}(\tilde{\mathbf{s}}_o(\tau = 0)) + \tilde{\mathbf{L}}(\tilde{\mathbf{s}}_o(\tau = 0)) \delta\tilde{\mathbf{s}}^2. \quad (10)$$

The modeling equation 10 is used to define the TFWI objective function as

$$J_{\text{TFWI}}(\tilde{\mathbf{s}}) = \frac{1}{2} \left\| \tilde{\mathcal{L}}(\tilde{\mathbf{s}}) - \mathbf{d} \right\|_2^2 + \epsilon \left\| |\tau| \tilde{\mathbf{s}}^2 \right\|_2^2. \quad (11)$$

The second term in the equation 11 rewards focusing of the data around zero time lag. It introduces a strong tomographic component, which is necessary to constrain the optimization problem because the slowness extension relaxes the constraints on the modeled data kinematics imposed by the data-fitting term (first term) in equation 11. This objective function can be minimized using a nested optimization algorithm with scale mixing, as discussed in the next section.

## 1D modeling example

We will use a simple 1D numerical example to analyze some of the characteristics of the TFWI method we introduced above. Figure 2 shows the difference between the background wavefield propagated with  $\mathbf{v}=1.2$  km/s and the wavefield propagated with the true velocity of  $\mathbf{v}=1.13$  km/s. The source function is a zero-phase wavelet bandlimited between 5 and 20 Hz. The wavefield difference is displayed as a function of propagation distance and travelttime. The velocity error is sufficiently high that the wavefields are completely out of phase after propagating for a couple of kilometers. This is therefore a situation like the ones described where the first order Born linearization (equation 8) would fail to model the data residuals and conventional FWI would have difficulties to converge, even though the problem is extremely simple.

Figure 3 shows the conventional FWI objective functions when the data are recorded with a single receiver located at 7 km for a total of 4 km offset from the source. The plot shows the value of the initial value of the objective function for several 1D transmission problems sharing the same starting velocity (1.2 km/s) and with different true velocities. If the true velocity is lower than  $\approx 1.18$  km/s or larger than  $\approx 1.22$  km/s a gradient based method starting from a velocity of 1.2 km/s will not converge to the right solution. On the contrary, the linearized modeling equation defined in equation 9 would have no troubles to model the data residual. For example, we can easily reproduce the wavefield difference shown in Figure 2 by setting the extended-velocity perturbation to be a delta function along the  $\tau$  axis, where the shift of the delta function linearly increases with the distance from the origin. This linear shift is computed by integrating the difference in slowness between the background model and the true model. The extended-velocity perturbation is shown in Figure 4. Figure 5 shows the result of solving equation 9 with the model shown in Figure 4. The approximation of the scattered wavefield  $\delta\mathbf{P}$  is almost identical to the wavefield difference shown in Figure 2.

Figure 2: Difference between the background wavefield computed with the starting velocity (1.2 km/s) and the wavefield propagated with the true velocity (1.13 km/s). [CR]

biondo1/. Init-Residual

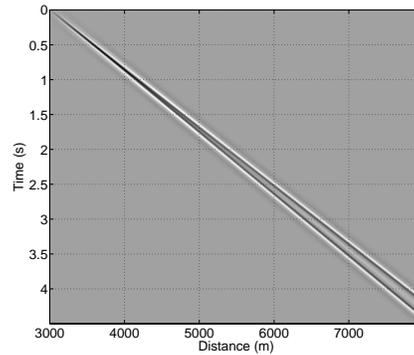


Figure 3: FWI norm as a function of the true velocity, when the starting velocity is equal to 1.2 km/s. [CR]

biondo1/. FWI-Norm-new

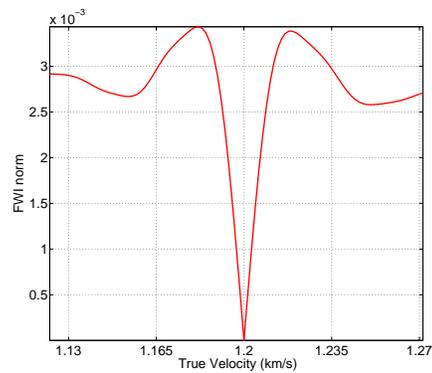


Figure 4: Extended velocity perturbation chosen to approximately model the wavefield difference shown in Figure 2. [CR]

biondo1/. Delta-Vel-Extend

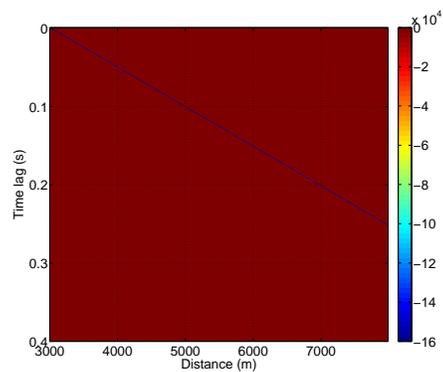
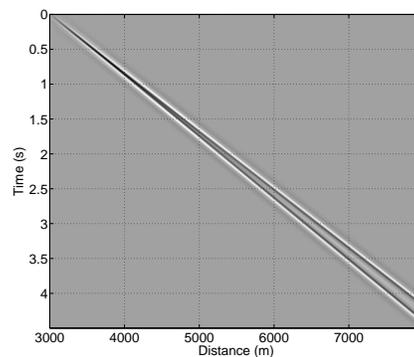


Figure 5: Perturbed wavefield computed by solving equation 9 with the model shown in Figure 4. [CR]

biondo1/. Lin-Forw-Ext-new



## Diving-wave modeling example

One of the advantages of extending the velocity model along the time-lag axis  $\tau$  is the capability to model with a linear operator large time shifts in the diving waves recorded by modern long-offset acquisition geometries. The capability of modeling time shifts in these events enables robust convergence of the inversion even when the starting velocity model is far from the correct one. To show these modeling capabilities we use one long-offset shot profile recorded over a half space with a vertical velocity gradient. The starting velocity model is assumed to be uniform and equal to the velocity at the surface. Figure 6a shows the data residual; both the recorded diving wave as well as the data modeled with the starting velocity are clearly visible.

The backprojection of the data residuals shown in Figure 6a, by the application of  $\tilde{\mathbf{L}}'$ , generates the velocity perturbation cube shown in Figure 7. The front panel of the cube shown in Figure 7 displays the zero time lag of the velocity perturbations. A substantial amount of the energy in the residual has been backprojected away from the zero time-lag panel.

Figure 6c displays the result of forward modeling the data residuals by the application of  $\tilde{\mathbf{L}}$  to the extended velocity perturbation shown in Figure 7. Although a “squaring” of the wavelet is evident in Figure 6c, the kinematics of these modeled residuals are very close to the kinematics of the true residuals shown Figure 6a. In contrast, when we use only the zero time lag of the velocity perturbation (front panel of the cube shown in Figure 7) to model the data residuals, we obtain the seismograms displayed in Figure 6b. The diving wave is totally missing from these modeled residuals because the background wavefield propagates with constant velocity along the horizontal direction.

Figure 8 shows the velocity perturbation cube when the velocity is extended along the horizontal subsurface offset axis. The front panel of the cube displays the zero subsurface offset, and thus it is identical to the front panel of the cube shown in Figure 7. Figure 6d displays the result of forward modeling the data residuals starting from the extended velocity perturbation shown in Figure 8. The diving wave event is present in these modeled residuals. However, it dies out at larger offsets, starting at about 8 kilometers. A subsurface offset extension of the velocity has difficulties in modeling large time shifts in transmitted events propagating in directions orthogonal to the subsurface offset axes. Consequently, as the propagation paths of the diving waves have longer vertical components, the less accurate the modeled residuals are. To address this limitation we could use the vertical subsurface offset in addition to the horizontal ones (Biondi and Symes, 2004), but the computational cost would increase accordingly. In 3D we would need to extend the velocity along 3 subsurface offsets, increasing further the dimensionality of the problem, and consequently its computational cost.

## OPTIMIZATION METHOD

The minimization of the objective function in 11 is a challenging optimization problem because of the non linearities in the modeling operator  $\tilde{\mathcal{L}}$  and the occasional contradictory search directions suggested by the gradients of the data fitting term and the focusing term. To overcome these challenges we devised and tested the nested optimization scheme with scale mixing described below.

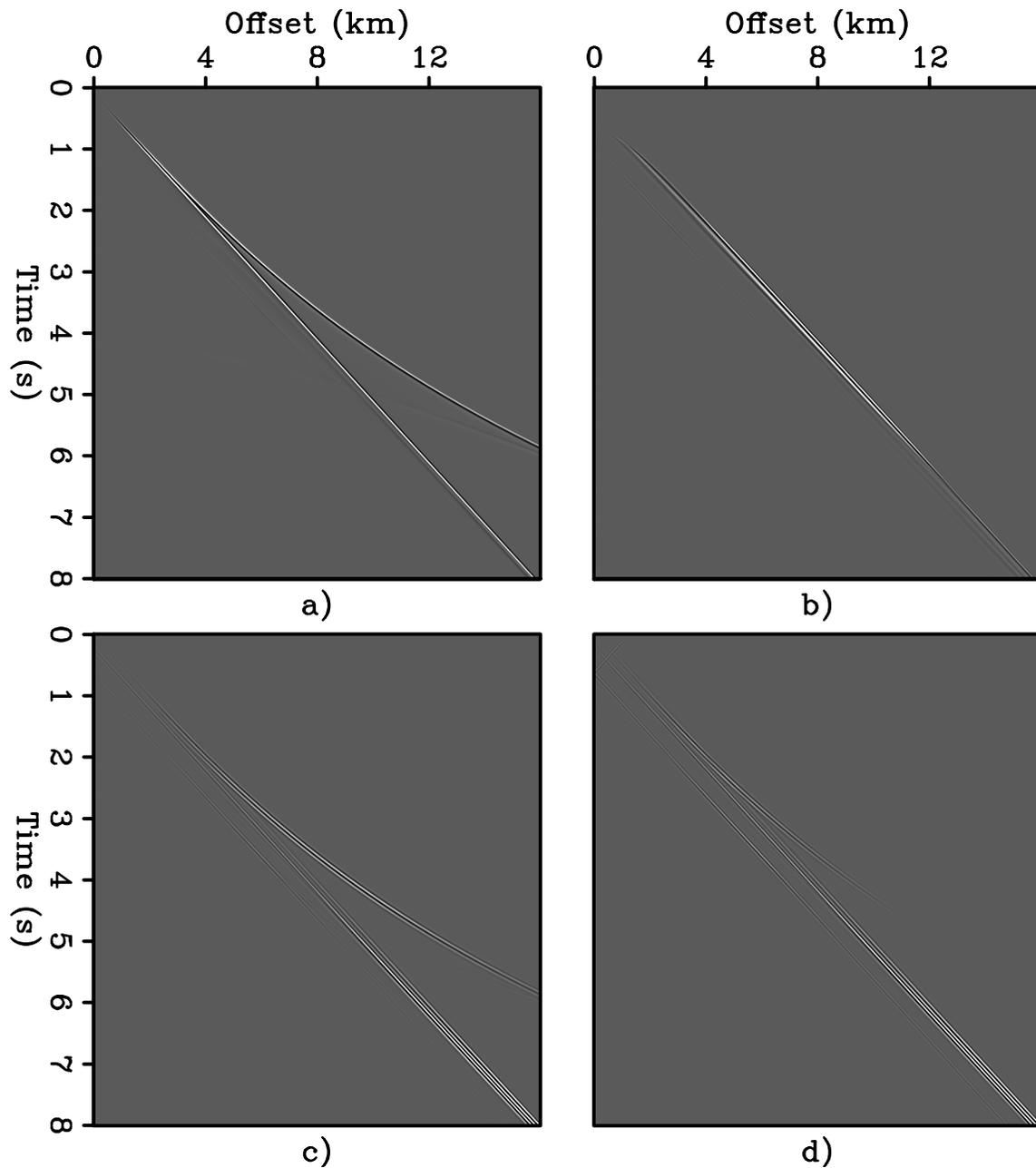


Figure 6: a) Difference between the background wavefield computed with the starting velocity and the wavefield propagated with the true velocity, b) data residual modeled from zero lag of the velocity perturbation cubes (front panel in cubes shown in both Figure 7 and Figure 8), c) data residual modeled from the velocity perturbations extended along the time-lag axis, (Figure 7), and d) data residual modeled from the velocity perturbations extended along the horizontal subsurface offset axis, (Figure 8). [CR] [biondo1/. diving-data-all](#)

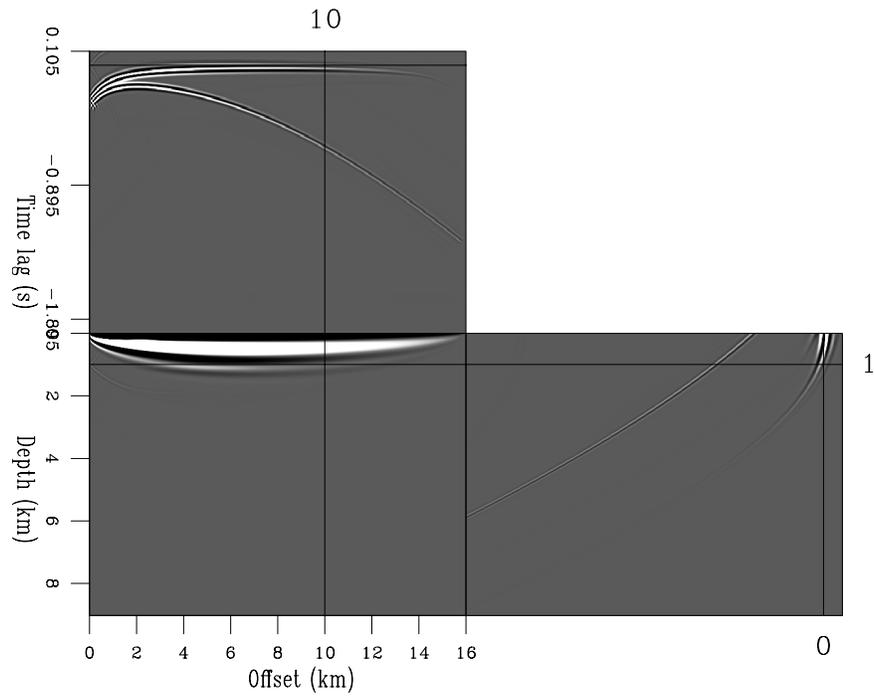


Figure 7: Velocity-perturbation cube extended along the time-lag axis and computed by backprojecting the data residuals shown Figure 6a. [CR] `biondo1/. diving-image-t-cube`

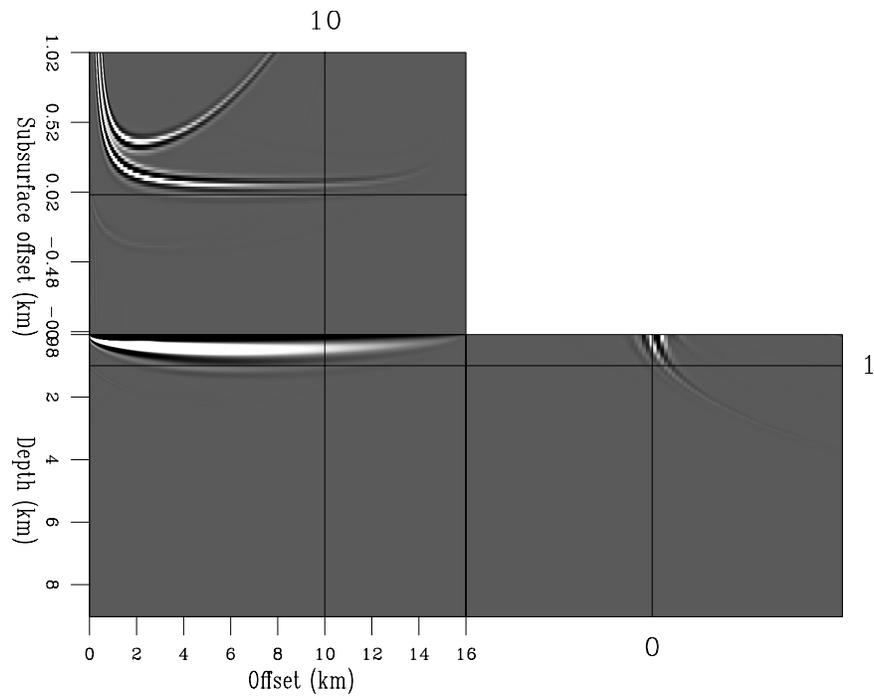


Figure 8: Velocity-perturbation cube extended along the horizontal subsurface-offset axis and computed by backprojecting the data residuals shown Figure 6a. [CR] `biondo1/. diving-image-x-cube`

## Nested Inversion

The proposed nested optimization scheme consists of an outer and an inner loop. In the outer loop we first compute the nonlinear data residual  $\Delta \mathbf{d} = \mathbf{d} - \tilde{\mathcal{L}}(\tilde{\mathbf{s}}_o)$ , where  $\tilde{\mathbf{s}}_o$  is the current slowness model evaluated at  $\tau = 0$ . The nonlinear residual is used as the “observed” data for the inner loop. The output of the inner loop is a search direction  $\Delta \tilde{\mathbf{s}}$ . We then perform a nonlinear line search that estimates the optimal step length  $\alpha$  that minimizes  $\left\| \tilde{\mathcal{L}}(\tilde{\mathbf{s}}_o + \alpha \Delta \tilde{\mathbf{s}}) - \mathbf{d} \right\|_2$ .

In the inner loop we formally separate the slowness model into a background model,  $\mathbf{b}$ , on which the operator  $\tilde{\mathbf{L}}$  depends non linearly, and a perturbation model  $\mathbf{p}^2(\tau)$ , on which the operator  $\tilde{\mathbf{L}}$  depends linearly. The objective function minimized in the inner loop is

$$J_{\text{ITFWI}}(\mathbf{b}, \mathbf{p}) = \frac{1}{2} \left\| \tilde{\mathbf{L}}(\mathbf{b}) \mathbf{p}^2 - \Delta \mathbf{d} \right\|_2^2 + \epsilon \left\| \tau \mathbf{p}^2 \right\|_2^2. \quad (12)$$

The starting models for the inner iterations are  $\mathbf{b}_o = \tilde{\mathbf{s}}_o(\tau = 0)$ , and  $\mathbf{p}_o = 0$ . The output of the inner loop after  $N$  iterations is  $\Delta \tilde{\mathbf{s}} = (\mathbf{b}_N - \mathbf{b}_o) + \mathbf{p}_N(\tau = 0)$ .

Unless the starting background slowness  $\mathbf{b}_o$  is very close to the true slowness, no choice of  $\mathbf{b}$  and  $\mathbf{p}$  will simultaneously zero the two terms in the objective function 12. This happens because of the particular choice of the data residual  $\Delta \mathbf{d}$ , that is the difference between the recorded data and the data modeled with  $\mathbf{b}_o$ . However, the models  $\mathbf{b}$  and  $\mathbf{p}$  that minimize this objective function provide an effective search direction  $\Delta \tilde{\mathbf{s}}$  for the outer loop of the nested optimization problem.

The modeling operator  $\tilde{\mathbf{L}}$  is linear with respect to perturbation  $\mathbf{p}$ , but nonlinear with respect to the background component  $\mathbf{b}$ . Therefore, another linearization around the “background” background is required to compute the gradient. The Born approximation is used (again) to linearize the  $\tilde{\mathbf{L}}$  operator with respect to the background resulting in a data-space tomographic operator,  $\mathbf{T}$ . Appendix A describes the derivation of this new data-space tomographic operator and how to numerically evaluate it and its adjoint. The expression of the two gradients at the inner iteration  $i$  is the following:

$$\nabla_{\mathbf{b}^2} = \mathbf{T}'(\mathbf{b}_i, \mathbf{p}_i) \left[ \tilde{\mathbf{L}}(\mathbf{b}_i) \mathbf{p}_i^2 - \Delta \mathbf{d} \right], \quad (13)$$

$$\nabla_{\mathbf{p}^2} = \tilde{\mathbf{L}}'(\mathbf{b}_i) \left[ \tilde{\mathbf{L}}(\mathbf{b}_i) \mathbf{p}_i^2 - \Delta \mathbf{d} \right] + \epsilon \tau^2 \mathbf{p}_i^2. \quad (14)$$

## Scale Mixing

In the inner loop, a straightforward use of the gradients is used to update their corresponding models directly. However, this would hinder the simultaneous inversion of different wavelengths of the model. This problem becomes apparent when we examine the result of the two operators in the inner loop. At the first iteration, the application of  $\tilde{\mathbf{L}}'$  to the data residual  $\Delta \mathbf{d}$  is equivalent to “migrating” the data and it could give a tomographic update which manifests as a low wavenumber update. In a conventional migration, this low wavenumber component is considered noise and filtered out. However, it is actually a tomographic component that should feed into the background model. The opposite argument is also true for the tomographic operator creating short-wavelengths perturbations.

Therefore, to improve our inversion results, we first mix the two gradients ( $\nabla_{\mathbf{b}^2}, \nabla_{\mathbf{p}^2}$ ) and then separate them in the Fourier domain to get the update of each model as follows:

$$\Delta \mathbf{b}^2(\vec{x}) = -\mathbf{C}_{\mathbf{b}^2} (\nabla_{\mathbf{b}^2}(\vec{x}) + \nabla_{\mathbf{p}^2}(\vec{x}, \tau = 0)), \quad (15)$$

where  $\Delta \mathbf{b}^2(\vec{x})$  is the search direction of the background model and  $\mathbf{C}_{\mathbf{b}}$  is a low-pass filter along the space coordinates  $\vec{x}$ . Similarly, we can compute the update of the perturbation model as:

$$\Delta \mathbf{p}^2(\vec{x}, \tau) = -\mathbf{C}_{\mathbf{p}^2} (\nabla_{\mathbf{b}^2}(\vec{x}) + \nabla_{\mathbf{p}^2}(\vec{x}, \tau)), \quad (16)$$

where  $\Delta \mathbf{p}^2(\vec{x}, \tau)$  is the search direction of the perturbation model and  $\mathbf{C}_{\mathbf{p}}$  is a high-pass filter along the space coordinates  $\vec{x}$ . In order to sum the two gradients properly, both of them need to have the same units as well as the same scale. This requires careful implementation of each operator at each linearization.

In the examples of this paper, we used a radial cut-off in the Fourier domain with a cosine squared taper. The wavelength cut-off is based on the dominant frequency in the data as well as the average slowness of the initial model. The two filters were designed such that they always sum to one at all wavelength to maintain the energy of the gradients. It is possible to design a more accurate filter that varies with frequency and slowness, but it is not necessary because both models will eventually be added to the slowness. This relative insensitivity is another benefit of applying the nested scheme we presented, in comparison with inverting the models separately and combining them only at the end. If the latter approach were implemented, the final results would be more sensitive to the choice of the scale-separation parameters.

This nested scheme has many benefits compared to a more conventional way of solving a tomographic inversion followed by an imaging inversion or even doing them simultaneously. The first benefit of this scheme is that the limitation of the linearized, first-order Born operator will not prevent the optimization from inverting higher order scatterings, such as multiple reflections and prismatic events. In this setup, the Born operator will attempt to match the first-order scattering from the background, which is not necessarily primary data since the background itself can generate many orders of scatterings if it contains sharp boundaries (and it will in later iterations). In other words, the linearized operator will account for one additional order of scattering at a time. Higher and higher scattering orders will be introduced with outer-loop iterations, until we eventually invert all the multiple-scattered events that are present in the data. This also means that higher-order scattered energy will initially be wrongly positioned in the model, but later iterations will correct for this mispositioning. The second benefit of this scheme is that it produces only one slowness model, because we keep pushing both background and perturbation into the slowness model, in contrast to the method presented in Almomin and Biondi (2012). Therefore, our goal is to drive the perturbation to a minimum and have the background explain the data.

## TFWI OF MARMOUSI DATA

A modified Marmousi model is used for the first synthetic examples where 500m of water layer is added to the top. Figure 9 shows the true velocity model. The thicker water layer reduces the amount of refracted energy being recorded in the data. Therefore, this dataset tests the capabilities of the inversion method to converge in presence of almost exclusively

reflected events. In contrast, the example presented in the next section tests the capabilities of the method to deal with mixed reflected/refracted arrivals.

There are 461 fixed receivers with a spacing of 20 m and 93 sources with a spacing of 100 m. We use a bandpassed wavelet with a frequency range between 5 Hz to 10 Hz and a small taper on both ends. The purpose of using this wavelet is to completely eliminate unrealistically low frequencies in the data. The initial model is shown in Figure 10 which is obtained by strongly smoothing the true model laterally after clipping the salt anomalies out. The inversion results after 900 outer-loop iterations, each of them with 10 inner iterations, are shown in Figure 11. The inversion shows remarkable results in reconstructing most features in the velocity model. The results are most accurate in the top and middle than at the bottom because they have the best illumination and coverage.

Figure 12 and 13 provide a clear illustration of the substantial improvements in the accuracy with which the final velocity model describes the data kinematics as compared with the initial one. Figure 12 shows the migrated image corresponding to the initial velocity; most of the reflectors are out of focus and mispositioned. In contrast, Figure 13 shows the migrated image corresponding to the final velocity. Reflectors are well focused and the structures are well imaged. In the middle of the section even the deepest reflectors are well focused and not grossly distorted.

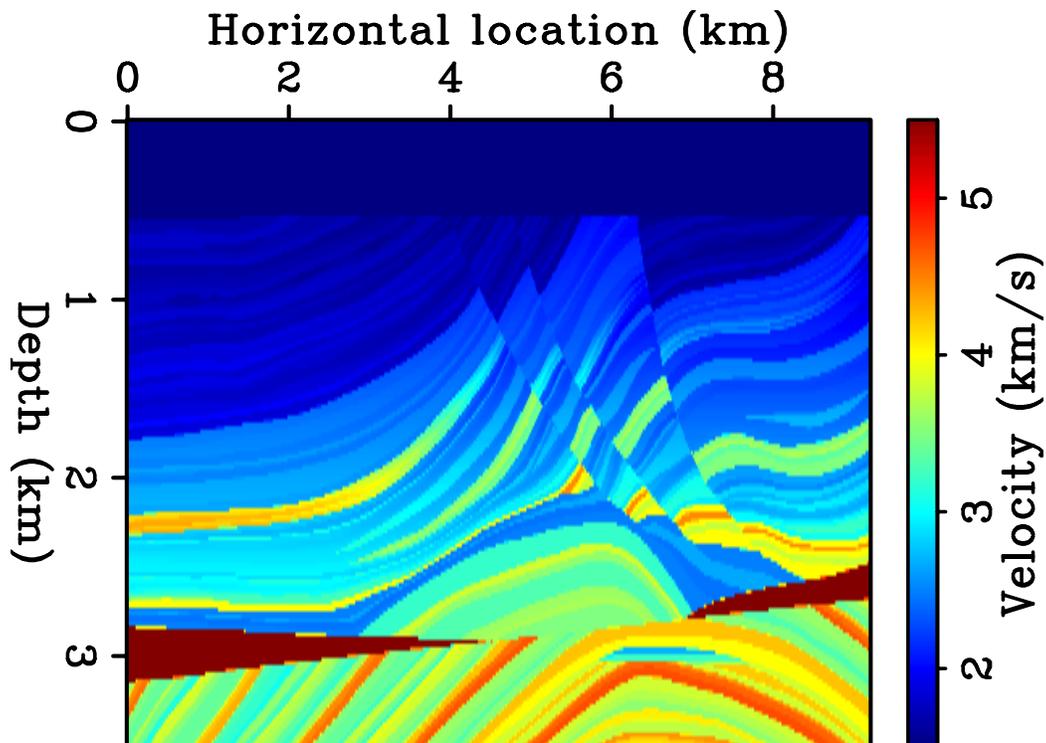


Figure 9: The true velocity of the Marmousi example. [ER]

`biondo1/. Geoph-marm-vel-true`

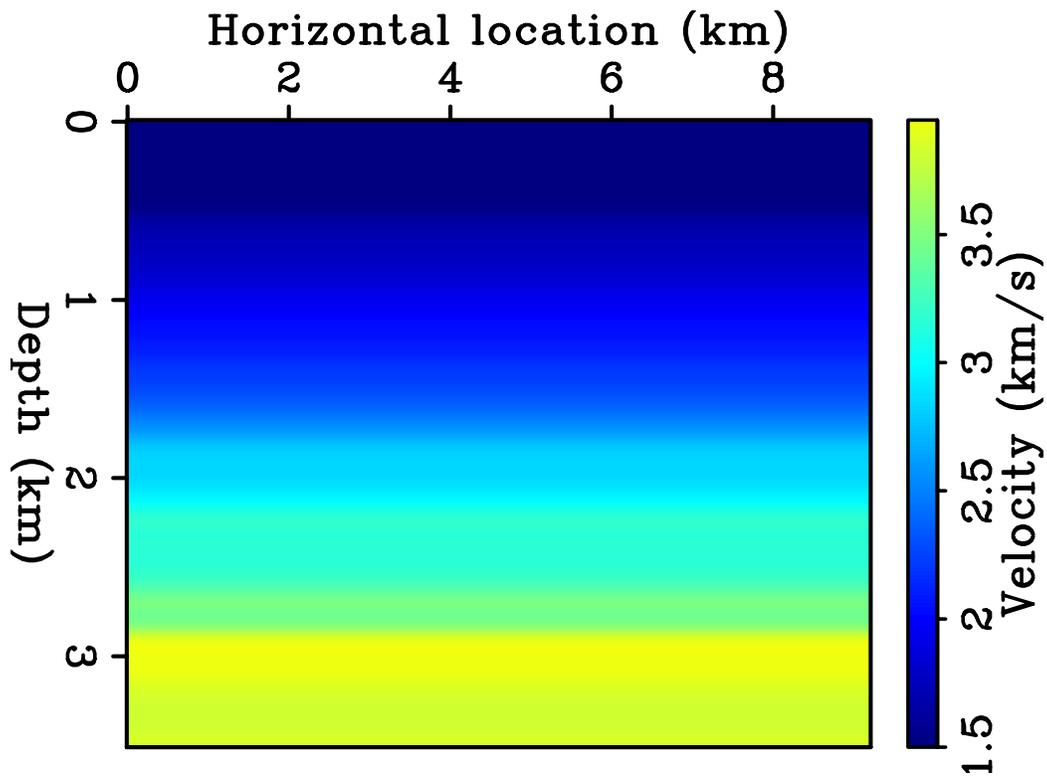


Figure 10: The initial velocity of the Marmousi example. [ER]

`biondo1/. Geoph-marm-vel-init`

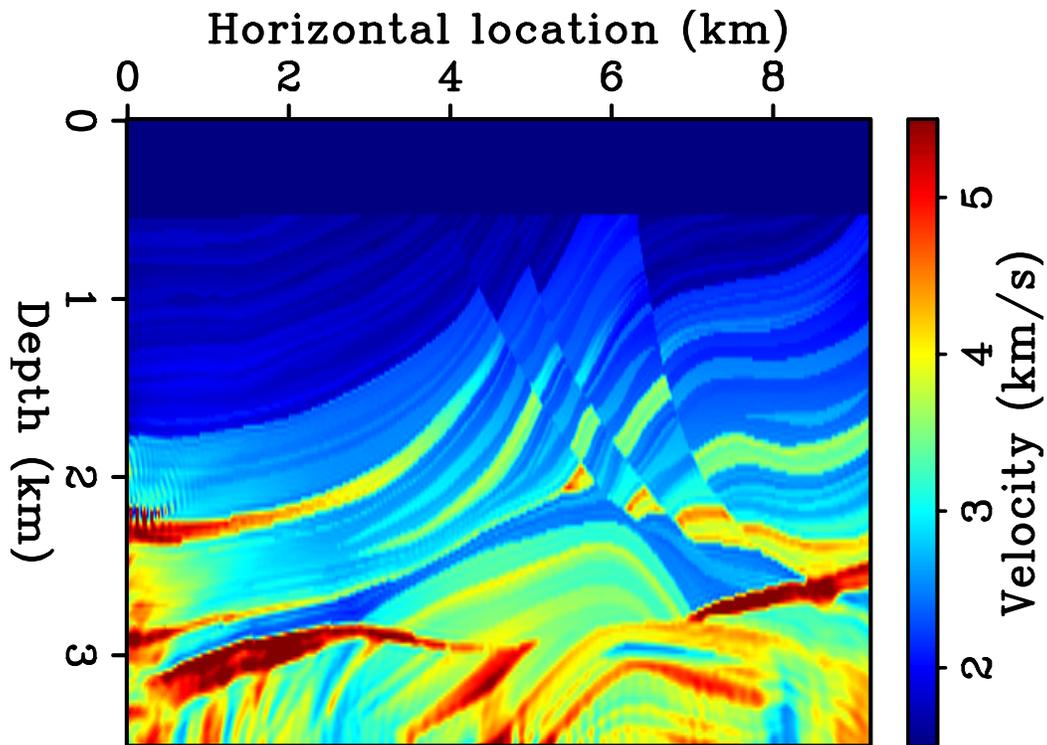


Figure 11: The inverted velocity of the Marmousi example. [CR]

`biondo1/. Geoph-marm-vel-final`

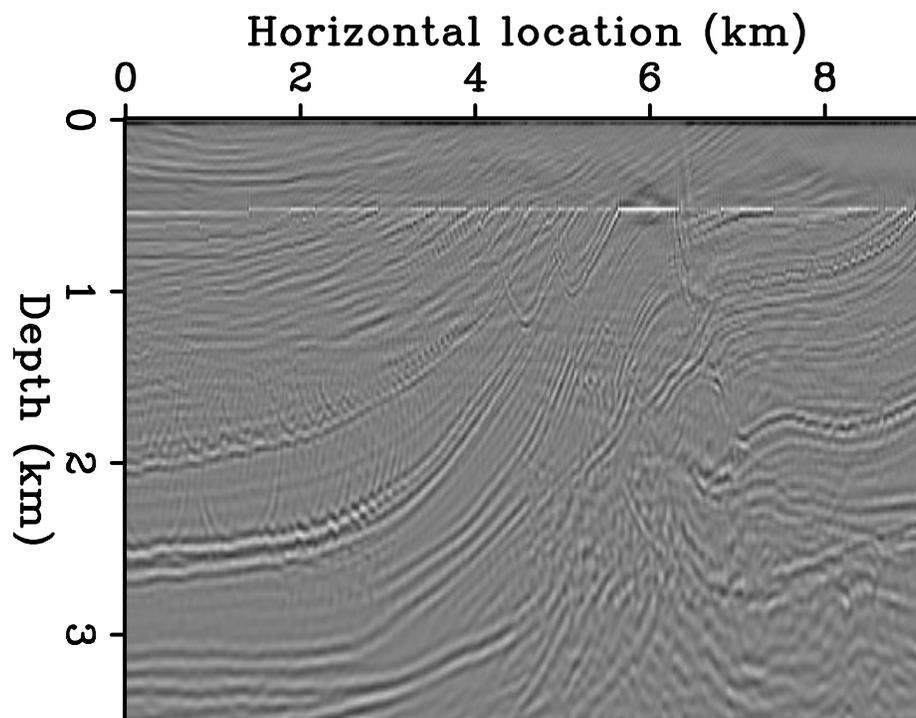


Figure 12: Migrated image using the initial velocity for the Marmousi example. [CR]

`biondo1/. Geoph-marm-init`

## TFWI OF LONG-OFFSET DATA

To verify the capabilities of the TFWI method based on a time-lag extension of the velocity model we tested the method on a synthetic data set recorded with long offsets. The data were generated over the “Caspian Sea” portion of the well-known BP velocity model, as shown in Figure 14. The receiver array was assumed to be fixed at the surface, and thus data with more than 20 kilometers long offsets were recorded. The source was a bandpassed wavelet between 5 and 10 Hz. No energy was present below 3 Hz.

Figure 15 shows the data recorded for the leftmost shot location. Strong and complex diving waves and refracted arrivals are visible in the data starting from approximately 8 kilometers offsets. These events carry useful information on the velocity, in particular in the shallow part of the section. In this data set, they are extremely useful to define the low-velocity anomalies present around the depth of two kilometers.

The starting model for the TFWI inversion was obtained by a strong horizontal smoothing of the true model, after the low and high velocity anomalies were removed. Figure 16 shows the shot gather located at the same location as the one shown in Figure 15, but modeled with the starting model, which is shown in Figure 17. As a direct comparison of Figure 15 with Figure 16 demonstrates, the differences between the true and starting models cause large time shifts in the diving-waves arrivals. The inaccuracies of the starting model, together with the lack of low frequencies in the data, prevent conventional FWI from converging to any useful model.

We solved the problem by a nested optimization algorithm as described earlier. The nested inversion converged towards the accurate model shown in Figure 18. The main features of the true model are accurately reconstructed. Some edge artifacts are present; they are caused by the finite span of the receiver array and the finite range of source locations. No sources were activated outside the displayed model, and the receiver array was fixed and covered the whole model.

Although the resolution of the model and the robustness of convergence are extremely attractive, the rate of convergence is slower than ideal; 500 iterations of the outer-loop iterations, each of them with 10 inner iterations, were required to estimate the model shown in Figure 18.

## DISCUSSIONS AND CONCLUSIONS

The integration of FWI and WEMVA into TFWI has the potential of yielding a waveform-inversion method that robustly converges to high-resolution models using the whole bandwidth of the seismic data simultaneously. We introduced a TFWI method based on the extension of the velocity along the  $\tau$  axis. This extension is based on a linear operator capable of correctly modeling transmitted events with large time shifts, as we demonstrate by two numerical examples based on simple 1D and layered models.

To minimize the TFWI objective function we propose a specialized nested inversion scheme. In the inner loop of this scheme the extended velocity model is separated into its background and perturbation components. The inversion scheme performs simultaneous inversion of different model scales by mixing the gradients of the two components and then separating them in the Fourier domain.

The results of the inversion of the Marmousi model demonstrate the strong convergence properties of the new method for typical reflection data. The results of the inversion of a long-offset data set recorded over the BP “Caspian Sea” demonstrate that the inversion method converges when both reflections and diving waves are recorded in the data, and inaccuracies of the starting velocity model create large errors in the kinematics of the diving waves.

The proposed optimization is effective to demonstrate TFWI convergence properties on realistic 2D data sets. However, the number of iterations required to produce the results shown in the paper are large: 900 for the Marmousi example and 500 for the “Caspian Sea” example. This suggests that further optimization of the inversion algorithm is needed for applications to 3D field data.

## ACKNOWLEDGMENTS

We would like to thank BP for making publicly available the velocity model of which we used a portion, and IFP for generating and distributing the Marmousi model. Almomin would like to thank Saudi Aramco for supporting his Ph.D. studies at Stanford.

## REFERENCES

- Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: SEG Technical Program Expanded Abstracts, **31**, 500–505.
- Bamberger, A., G. Chavent, C. Hemon, and P. Lailly, 1982, Inversion of normal incidence seismograms: *Geophysics*, **47**, 757–770.
- Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **31**, 547–552.
- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.
- Biondi, B. and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: *Geophysics*, **69**, 1283–1298.
- Mora, P., 1989, Inversion = migration + tomography: *Geophysics*, **54**, 1575–1586.
- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis—I: Theory: *Geophysical Prospecting*, **52**, 593–623.
- , 2004b, Wave-equation migration velocity analysis—II: Examples: *Geophysical Prospecting*, **52**, 607–623.
- Sava, P. and I. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: *Geophysics*, **73**, VE145–VE159.
- Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: *Geophysics*, **73**, VE49–VE59.
- Sun, D. and W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: SEG Technical Program Expanded Abstracts, **31**, 497–502.
- Symes, W. W., 2008, Migration velocity analysis and waveform inversion: *Geophysical Prospecting*, **56**, 765–790.

- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- Yang, T. and P. Sava, 2009, Wave-equation migration velocity analysis using extended images: *SEG Technical Program Expanded Abstracts*, **28**, 3715–3719.
- , 2010, Moveout analysis of wave-equation extended images: *Geophysics*, **75**, S151–S161.

## APPENDIX A

The minimization of the objective function 12 requires the computation of the gradients with respect to both  $\mathbf{b}$  and  $\mathbf{p}$ . These gradients can be computed by a perturbation analysis of the modeling operator  $\tilde{\mathbf{L}}$ . As discussed in the main text,  $\tilde{\mathbf{L}}$  is evaluated by solving equations 7 and 9. Rewriting these equations in terms of  $\mathbf{b}$  and  $\mathbf{p}$ , and the incident wavefield,  $\mathbf{P}_i$  and the scattered wavefield,  $\mathbf{P}_s$ , we obtain

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{A-1})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_s = \mathbf{p}_i^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \mathbf{P}_i. \quad (\text{A-2})$$

Introducing the perturbations,  $\delta \mathbf{b}^2$  and  $\delta \mathbf{p}^2$ , in the two model variables into equations C-2–C-3, and introducing the corresponding perturbations in the wavefields,  $\delta \mathbf{P}_i$  and  $\delta \mathbf{P}_s$  yields the following perturbed system:

$$[(\mathbf{b}_i^2 + \delta \mathbf{b}^2) \mathbf{D}_2 - \nabla^2] (\mathbf{P}_i + \delta \mathbf{P}_i) = \mathbf{f}, \quad (\text{A-3})$$

$$[(\mathbf{b}_i^2 + \delta \mathbf{b}^2) \mathbf{D}_2 - \nabla^2] (\mathbf{P}_s + \delta \mathbf{P}_s) = [\mathbf{p}_i^2(\tau) + \delta \mathbf{p}^2(\tau)] \overset{\tau}{*} \mathbf{D}_2 (\mathbf{P}_i + \delta \mathbf{P}_i). \quad (\text{A-4})$$

By setting  $\delta \mathbf{b}^2 = 0$  in equations C-4–C-5, we derive the following system of equations:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{A-5})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta \mathbf{P}_s = \delta \mathbf{p}^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \mathbf{P}_i, \quad (\text{A-6})$$

which defines again  $\tilde{\mathbf{L}}$  and can be used to evaluate the perturbations in the scattered wavefield  $\delta \mathbf{P}_s$ , and consequently in the recorded data  $\delta \mathbf{d} = \mathbf{S} \delta \mathbf{P}_s$ , caused by perturbations  $\delta \mathbf{p}^2$ .

We can derive the data-space tomographic operator,  $\mathbf{T}$ , which relates perturbations in the scattered wavefield,  $\delta \mathbf{P}_s$ , to perturbations in the background model,  $\delta \mathbf{b}^2$ , by setting  $\delta \mathbf{p}^2 = 0$  and neglecting the higher-order terms in  $\delta \mathbf{b}^2$  in equations C-4–C-5.

This tomographic operator is equal to the sum of two operators,  $\mathbf{T}_i$  and  $\mathbf{T}_s$ . The first operator,  $\mathbf{T}_i$ , models perturbations in the scattered wavefield caused by perturbations in the propagation of the incident wavefield:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{A-7})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta \mathbf{P}_i = \delta \mathbf{b}^2 \mathbf{D}_2 \mathbf{P}_i, \quad (\text{A-8})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta \mathbf{P}_s = \mathbf{p}_i^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \delta \mathbf{P}_i, \quad (\text{A-9})$$

The second operator,  $\mathbf{T}_s$ , models perturbations in the scattered wavefield caused by perturbations in the propagation of the scattered wavefield itself:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{A-10})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_s = \mathbf{p}_i^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \mathbf{P}_i, \quad (\text{A-11})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta \mathbf{P}_s = \delta \mathbf{b}^2 \mathbf{D}_2 \mathbf{P}_s, \quad (\text{A-12})$$

Both of these tomographic operators depend nonlinearly on  $\mathbf{b}_i$  and linearly on  $\mathbf{p}_i^2$ . They have zero output when  $\mathbf{p}_i^2$  is equal to zero; that is, at the first iteration of the inner loop because we set  $\mathbf{p}_o^2 = 0$  as a starting model of the inner iterations. However, as we update the linearization at each iteration, starting from the second iteration the output of  $\mathbf{T}$  becomes different from zero.

The data-space tomographic operator represented by equations C-8–C-10 and equations C-11–C-13 is analogous to the WEMVA operator (Biondi and Sava, 1999; Sava and Vlad, 2008) except that in the WEMVA operator, we keep the data fixed and vary the image; it is the other way around in this tomographic operator.

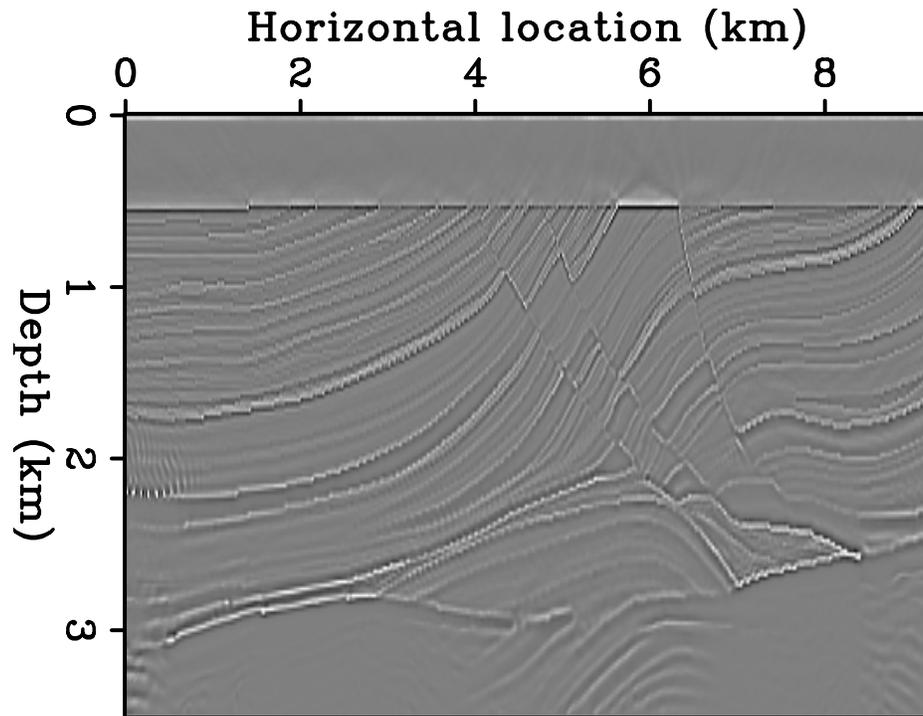


Figure 13: Migrated image using the final velocity for the Marmousi example. [CR] [biondo1/. Geoph-marm-final](#)

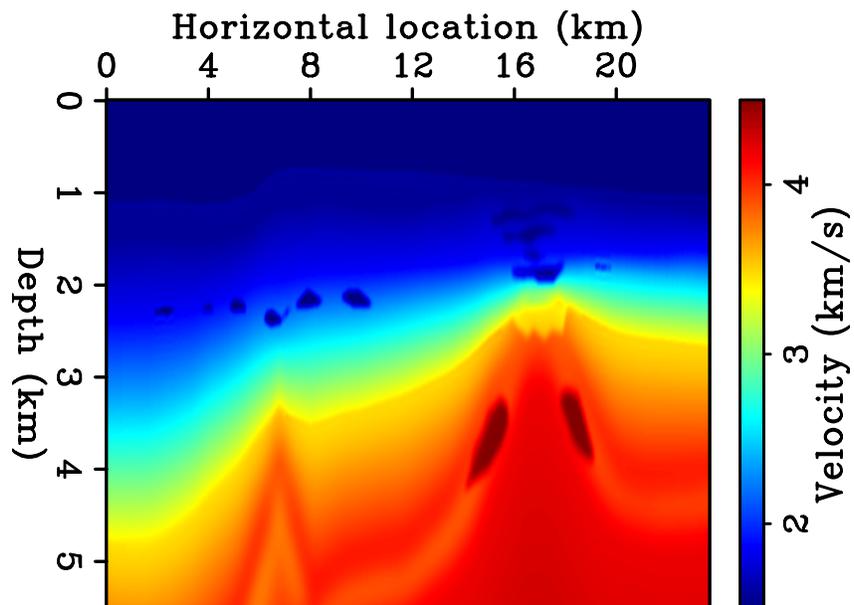


Figure 14: Portion of the BP velocity model used for the numerical test of the proposed TFWI method. The model contains both low-velocity anomalies (shallow gas) as well as high-velocity anomaly on the flanks of the mud volcano. [ER] [biondo1/. sepmeet-bp-true](#)

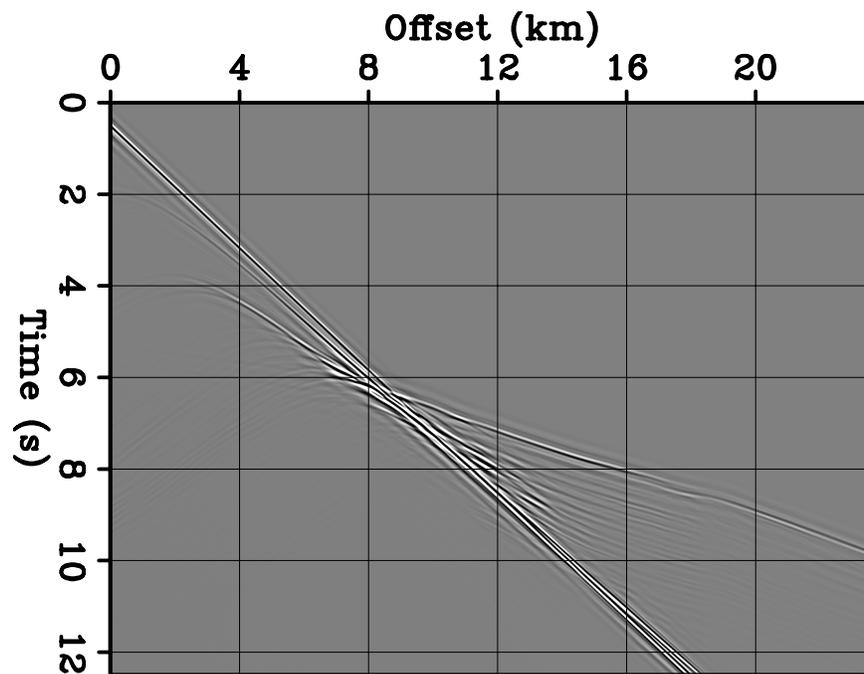


Figure 15: Leftmost shot profile recorded on the model shown in Figure 14. Notice several diving waves and refractions present in the data at offset larger than 8 kilometers. These events carry useful information for the estimation of the velocity anomalies present in the model. [CR] `biondo1/. data-bp-large`

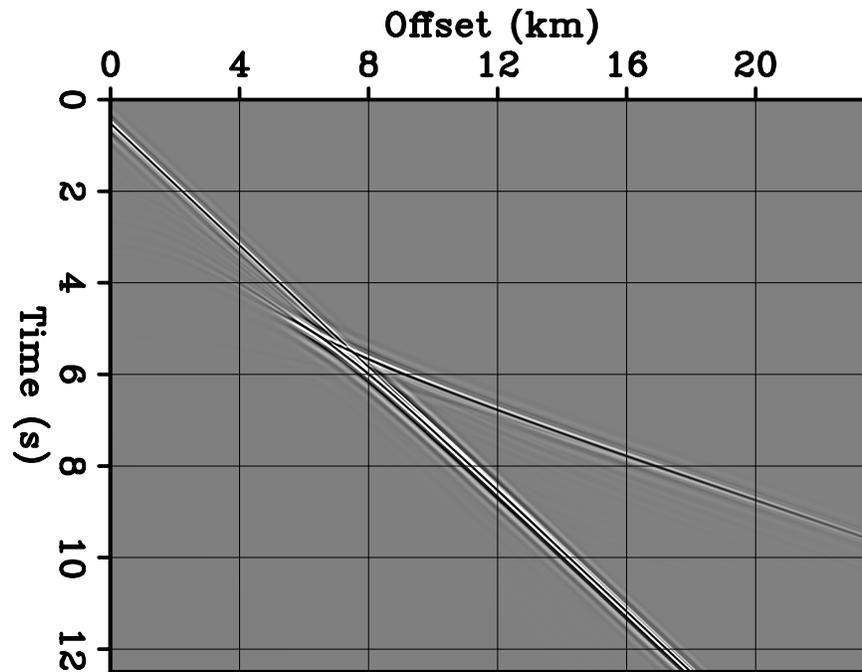


Figure 16: Shot gather modeled assuming the starting model shown in Figure 17 at the same shot location as the data shown in Figure 15. Notice the large time shifts between the diving-wave arrivals in this gather with the one shown in Figure 15. [CR] `biondo1/. data-bp-init-large`

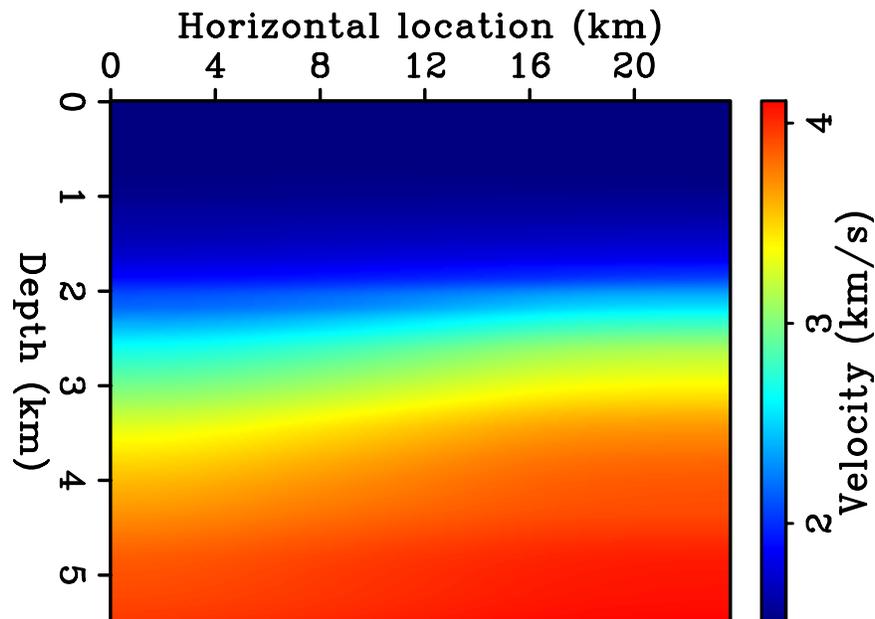


Figure 17: Starting model for the TFWI inversion. This model was obtained by strong horizontal smoothing of the model shown in Figure 14, after the low and high velocity anomalies were removed. The lack of low frequencies in the data makes this model inappropriate for starting a conventional FWI inversion. [ER] `biondo1/. sepmeet-bp-init`

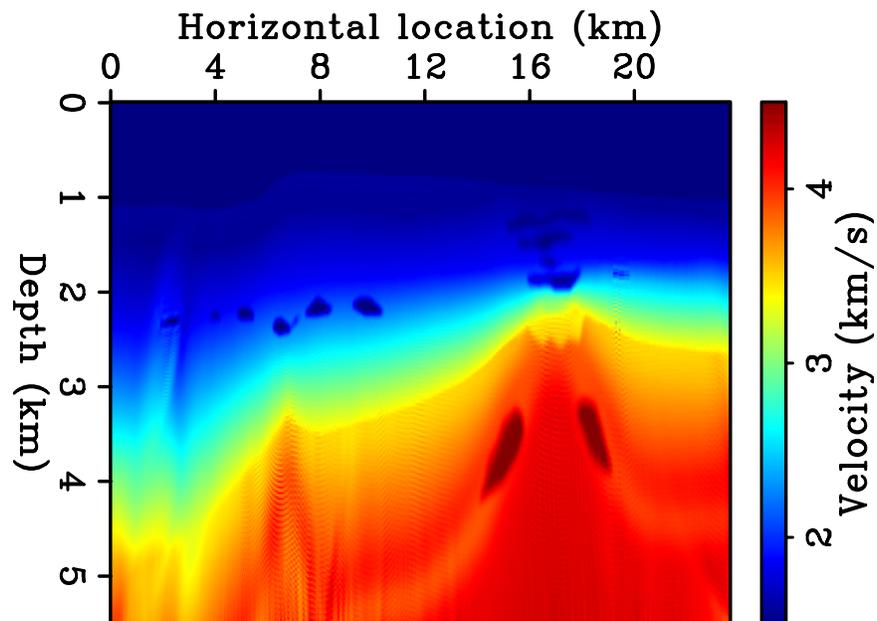


Figure 18: Estimated model after 500 iterations of the outer loop of the TFWI inversion based on time-lag extension of the velocity model. The main features of the true model are accurately reconstructed. Some edge artifacts are present; they are caused by the finite span of the receiver array and the finite range of source locations. No sources were activated outside the displayed model, and the receiver array was fixed and covered the whole model.

[CR] [biondo1/. sepmeet-bp-inv](#)

# Near-surface velocity estimation for a realistic 3D synthetic model

*Xukai Shen*

## ABSTRACT

I performed data-domain wave-equation tomography for a realistic synthetic near-surface model. From a starting model that misses some large scale velocity features as well as some small scale velocity features, both traveltimes tomography and waveform tomography were performed. First-break traveltimes tomography using wave equation not only results in correct updates of large scale velocity structure, but also gives hints of small scale velocity structures. The tomography result can be further refined by refraction waveform tomography. Refraction waveform tomography pin-point the location of small scale velocity features by using the waveform information in addition to the traveltimes information. However, direct refraction waveform tomography without traveltimes tomography can not resolve the missing large velocity features in the starting model, and easily converges to a local minima.

## INTRODUCTION

Conventionally, people use ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989) to derive near-surface velocity. Such smooth solutions may be adequate for areas with simple near-surface velocity structures, but in geologically complex areas, smooth velocities are not accurate enough for imaging deeper reflectors (Marsden, 1993; Bevc, 1995; Hindriks and Verschuur, 2001). In such cases, data-domain wave-equation tomography (Tarantola, 1984; Pratt et al., 1998; Mora, 1987; Luo and Schuster, 1991) tends to give more accurate results (Ravaut et al., 2004; Sheng et al., 2006; Sirgue et al., 2009) by simulating finite-frequency seismic wave propagation.

Data-domain wave-equation tomography methods update velocity using mismatches between observed data and modeled data. The mismatches usually include traveltimes (first-order) and waveform (second-order) mismatches. Both traveltimes and waveform mismatches are used by Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt et al., 1998; Mora, 1987), leading to high-resolution results but requiring an accurate starting model. On the other hand, Wave-equation Traveltime Inversion (WTI) (Luo and Schuster, 1991) estimates velocity model by minimizing only the traveltimes difference between observed data and modeled data using the wave equation. As a result, WTI differs from FWI in two ways: first, WTI is not affected by bad starting models with cycle skipping; second, WTI results tend to have low resolution. However, the low-resolution result from WTI can be a good starting model, and subsequent FWI can obtain a high-resolution result (Shen et al., 2012).

In this paper, I examine the resolution of first-break traveltimes wave-equation tomography and refraction waveform tomography, using a realistic near-surface model with complex velocity structures. The model contains velocity features of various spatial extend, making

it a good candidate for testing tomography resolution. First I will review data-domain wave-equation tomography methodology, then I will show why ray is unfavorable in such setting, finally I will compare and analyze tomography results.

## REVIEW OF DATA-DOMAIN EARLY ARRIVAL WAVE-EQUATION TOMOGRAPHY

Data-domain wave-equation tomography updates velocity or earth model in an iterative fashion by minimizing the difference between synthetically modeled data and recorded data. The tomography scheme can be summarized by the following pseudo code:

---

**Algorithm 1** Pseudo code of data-domain wave-equation tomography

---

**for** iteration =1,n **do**

    Calculate gradient of objective function regarding model parameters.

    Calculate update direction using the current gradient (and/or previous gradients).

    Calculate steplength of the update direction.

    Update model.

**end for**

---

Data-domain early arrival wave-equation tomography, as the name suggests, tries to model and match only early (temporal) arrivals in recorded data. Since early arrivals mainly travels through near-surface before returning to the surface, it is the ideal candidate when we are only inverting for near-surface velocity. Early arrivals include refractions, diving waves and shallow reflections. While shallow reflections can be used for near-surface velocity analysis(Shen, 2013), the tomography carried out in this paper used refraction and diving waves only. As will be shown in the example, to recover both large scale and small scale velocity features, both traveltimes and waveform of refraction/diving waves are needed.

Using waveform or traveltimes in the tomography is achieved by using different objective functions. More specifically, waveform tomography uses objective functions that subtract modeled data from recorded data, usually with some amplitude correction term to account for the amplitude difference caused by non-velocity factors(Shen, 2010). Traveltimes tomography, on the other hand, uses objective functions that measure the traveltimes difference between recorded and observed data, this can be done by one of the two options: 1. picking the time lag that maximize the correlation between recorded data and observed data; 2; picking the traveltimes of corresponding events ( such as first-breaks) in both recorded data and observed data, then subtract the two picks.

Using waveform and/or traveltimes of recorded data in the tomography naturally leads to the choice of time-domain(Mora, 1987) tomography scheme over frequency-domain scheme(Pratt et al., 1998) . Traditional time-domain schemes are very computationally expensive for two reasons: first, it involves iterative forward and backward two-way wavefield propagation; second, to compute velocity update direction, at least one of the wavefield needs to be saved. Because wavefields in large-scale 3D application can require terabytes of storage, I/O can be a non-trivial bottleneck. While computational power is growing very fast, it is less so for the memory size of computers. As shown in the pseudo code in Appendix A, using random boundaries can greatly reduce the memory requirement, hence eliminating the I/O bottleneck.

## COMPLEX VELOCITY STRUCTURE AND RAY BEHAVIOR

In the history of geophysical prospecting, people have identified and categorized many types of near-surface velocity features. However, to combine them into a realistic near-surface velocity model, it is better to generate those features via geological oriented process. Robert Clapp proposed one way of doing this in the last SEP report (Clapp, 2013), which was used to generate the near-surface velocity model in this paper.

The velocity model is generated by a series of geological processes. The most frequently used process is deposit, where a layer of constant thickness velocity is deposited onto the top of the current model, within the layer there are also fine layering and random lateral velocity variations. Both features are part of the true model (figure 1), and both are removed from the starting model. Depositing layers create a simple model that can be made more complex. More specifically, the simple model can be made more complex by certain “morphing” processes such as uplift (figure 2), compact and faulting. Those processes create macro velocity features, micro velocity features can be created by erosion processes. Current version of the software is capable of creating channel or “bowl” erosion, both types of which are used in creating the true model, but not the starting model for tomography. The velocity model created by the combination of all those processes are realistically complicated (figure 3).

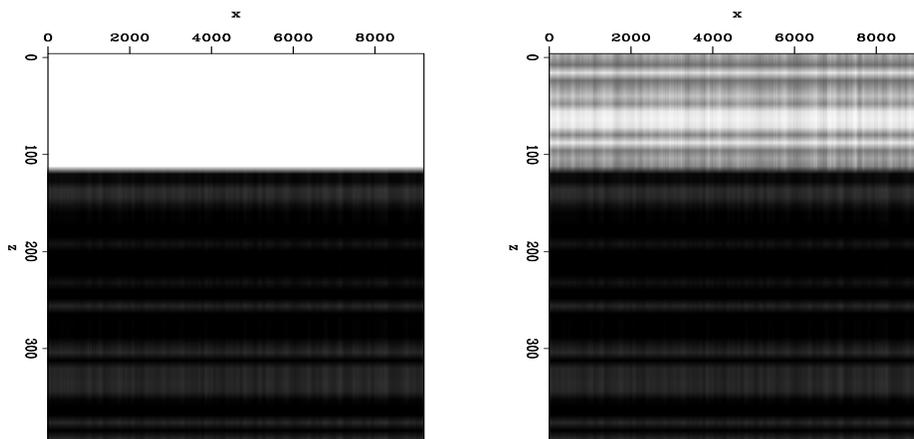


Figure 1: Velocity model building process, left: before deposition; right: after deposition.

[ER] xukai1/. deposit

Ray based methods have a very difficult time to deal with such complicated near-surface model. To illustrate this, I took a 2D slice of the 3D model, then performed three scenarios of ray tracing on the 2D slice (figure 4). All the scenarios use the same source to propagate four rays, each ray propagate ten seconds or hit the earth surface, whichever came first. The take-off angle of rays starts at forty-nine degrees from vertical down, with five and an half degrees increment. The only difference being the propagation time interval, where the first scenario use one milli-second interval, and the second and the third scenario use half and quarter Milli-second, respectively. Change of the time interval lead to slight lateral shift when rays hit the first interface. However, such small change in lateral arrival point translate into huge different in the later ray path. This is particularly obvious in the diving/refraction rays. Diving/refraction rays in such complicated velocity model is highly non-linear with ray parameters. Such non-linearity will be much worse in the full 3D model, making it

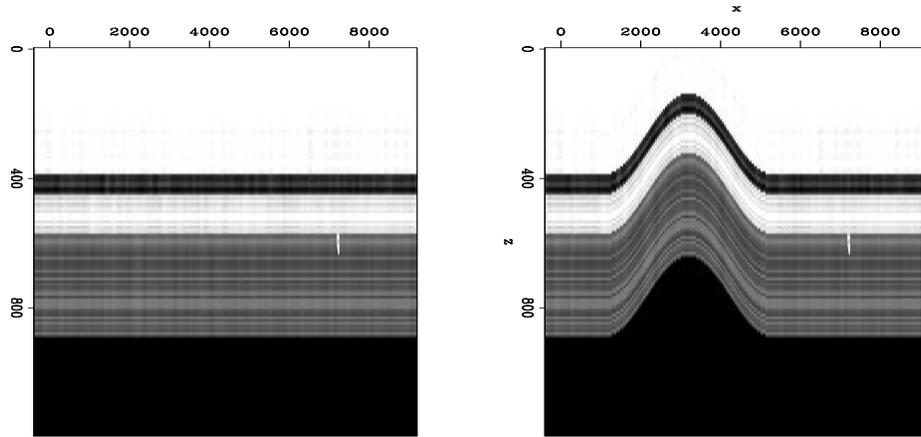


Figure 2: Velocity model building process, left: before uplift; right: after uplift. [ER]

xukai1/. uplift

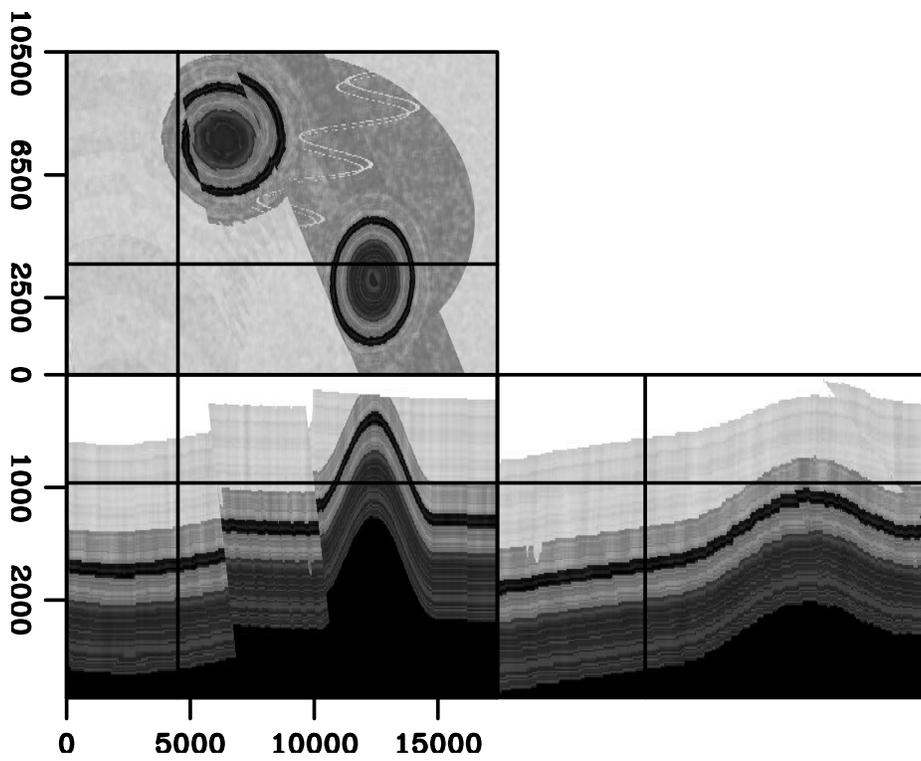


Figure 3: True velocity model for tomography. [ER] xukai1/. vtrue

completely unsuitable for high-resolution tomography. The starting model (figure 5) is

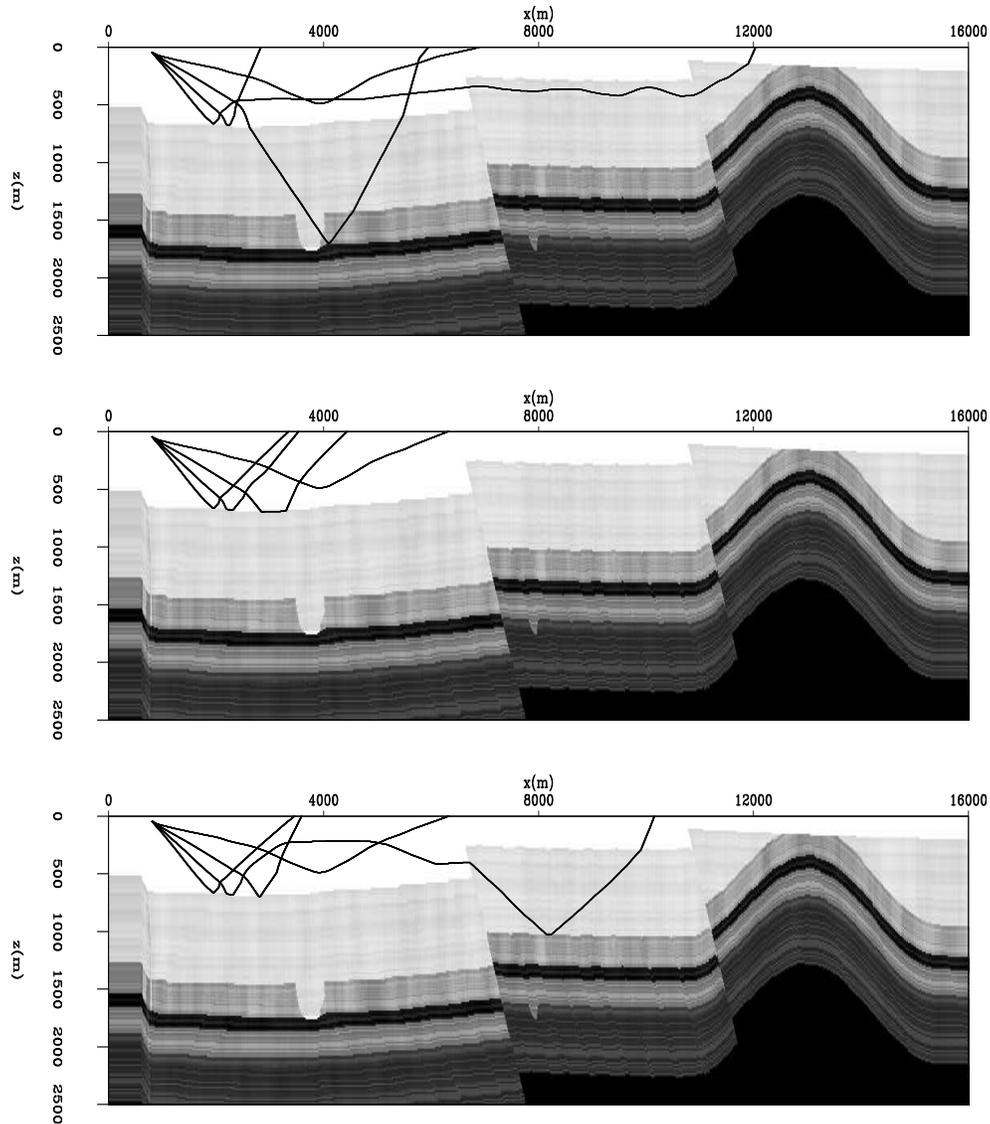


Figure 4: Ray tracing through the 2d slice of the true Velocity model with top: 1ms time interval; middle:0.5 ms time interval;bottom:0.25 ms time interval. [ER] `xukai1/. rays`

created by using most of the steps that created the true model. All the depositing steps are used, but fine layering, lateral variation of velocity within each layer and thin layer of high velocity are removed. All the “morphing” are applied exactly the same way. All the erosion steps are skipped. Also the model created by those steps are smoothed before using as input to tomography. Such a starting model simulates ray-based near-surface tomography result, where the model is usually smooth and lacks lateral resolution.

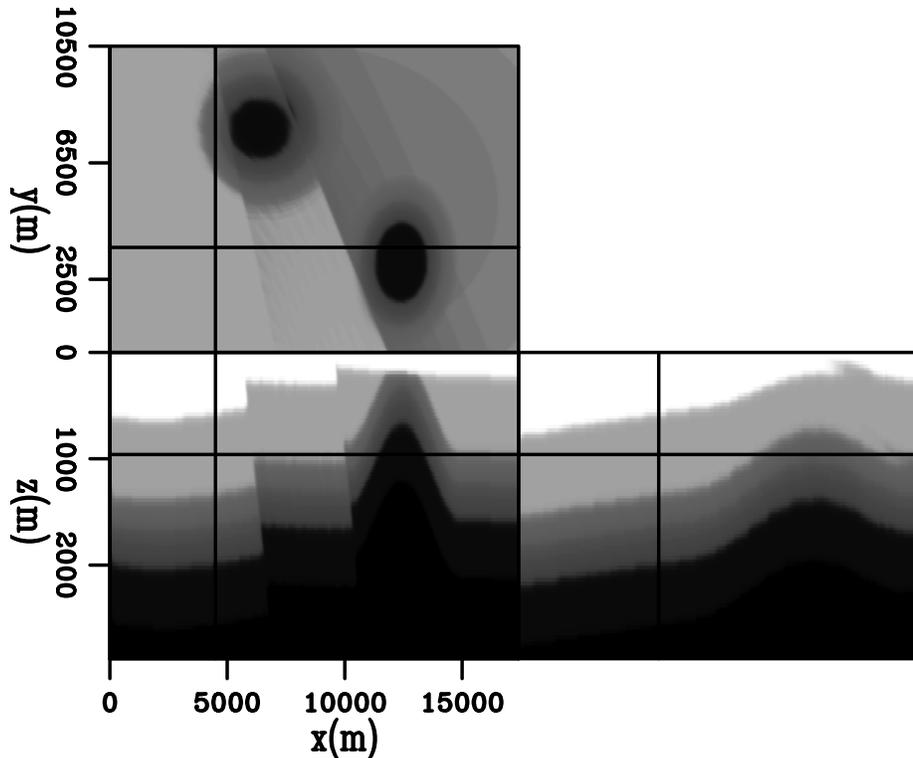


Figure 5: Starting velocity model for tomography. [ER] `xukai1/. vstart`

## TOMOGRAPHY AND RESULTS

Given the complicated true model, particularly the depth of various features, synthetic survey is designed such that most of the features are illuminated by diving waves/refractions. The survey simulates a typical cross-spread land survey setting. Source lines are along cross-line direction. Receiver lines are along inline direction, and are orthogonal to source lines. Source spacing is 120m along the cross-line direction, and 390m along the inline direction. Each shot point is centered in a rectangular patch of receivers. Receivers offset range from  $-5.1\text{km}$  to  $5.1\text{km}$  in the inline direction, with 30m spacing, and  $-2\text{km}$  to  $2\text{km}$  in the crossline direction, with 180m spacing. Since we are only interested in using diving wave/refractions, the recording time was only 3.5 seconds. A total number of 3240 shots are modeled with each shot contains up to 7840 traces. All the data are modeled on 15 m grid spacing in  $x$ ,  $y$  and  $z$ , with time sampling of 3 ms. The total amount of data are more than 100 GB.

To recover lateral velocity variation from starting model, it is important to choose data frequency that contains information of those velocity variations. I use the quarter-wavelength rule to select my temporal frequency. The rule states that seismic wave with spatial wavelength of  $x$  can resolve spatial velocity feature that is equal or larger than  $x/4$ . This empirical rule is very useful here since the true model contains not only small scale lateral velocity variations, but also two sinusoidal channels that are very close to each other in some places. To make sure we can invert for those small scale velocity variations, a Ricker wavelet of 12Hz was used for both modeling and tomography. The small scale velocity variation is on the order of ten numerical grids, hence for computational efficiency,

inversion were carried out on 30m spacing in x,y and z, with time sampling of 6 ms.

Three tomography tests are performed, the first one is the wave-equation traveltime tomography by minimizing the first break traveltime differences between observed data and modeled data. The second and third are both waveform tomography by minimizing refraction/diving waveform differences between observed data and modeled data. The two waveform tomography use different starting models, one uses traveltime tomography result, the other uses the starting model for traveltime tomography.

Traveltime tomography is less sensitive to cycle-skipping than waveform tomography. As a result, relatively bold steplength search is performed for each iteration. A total number of 10 iterations were run. The inverted model is in figure 6, it is still smooth compare with the true model. However, a close comparison to the starting model shows that most velocity features already start to appear. For example Given the velocity of about 2.5km/s where the two river channels are at, the quarter-wavelength rule suggest that the 12Hz peak frequency wavelet can detect velocity features down to about 50m. Since the channels have width and depth that are bigger than 100m, they can be resolved (figure 6. At the same time, other bigger velocity features such as lateral velocity variation, high velocity layer, and “bowl” type erosion are resolved as well. However, resolving in the content of traveltime tomography means the bulk of the velocity features start to appear, but not the exact boundaries of velocity features. Resolving the boundary requires us to at least correct the second order data misfit-the waveforms.

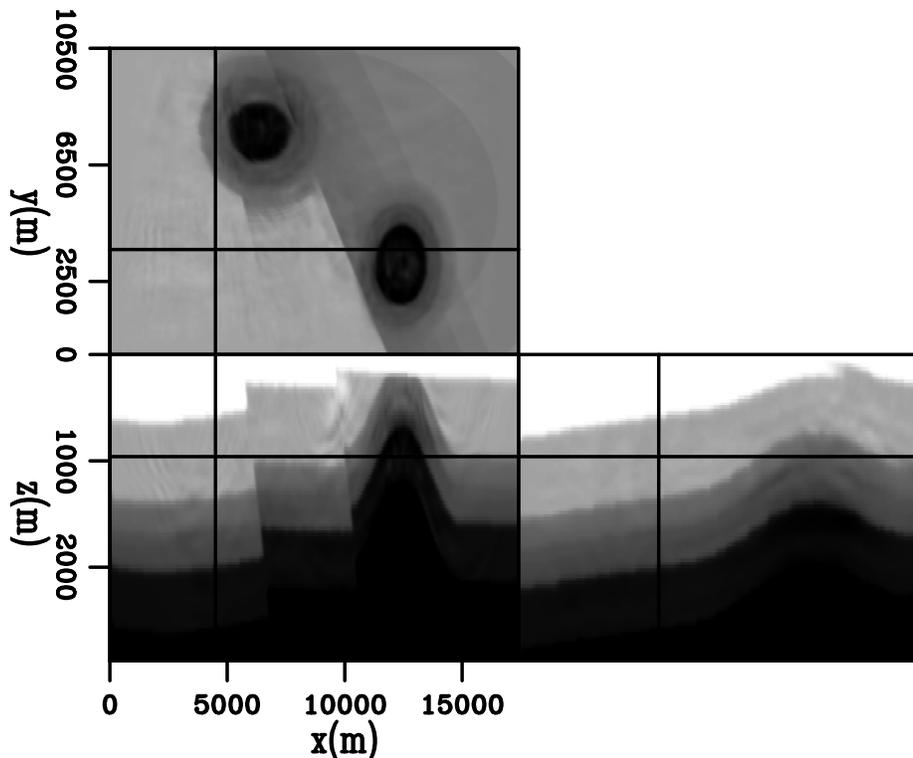


Figure 6: Traveltime tomography result. [CR] xukai1/. vwt

For the waveform tomography, since waveform dependency on velocity or slowness is less linear compare to that of traveltime, both waveform tomography tests use a cautious

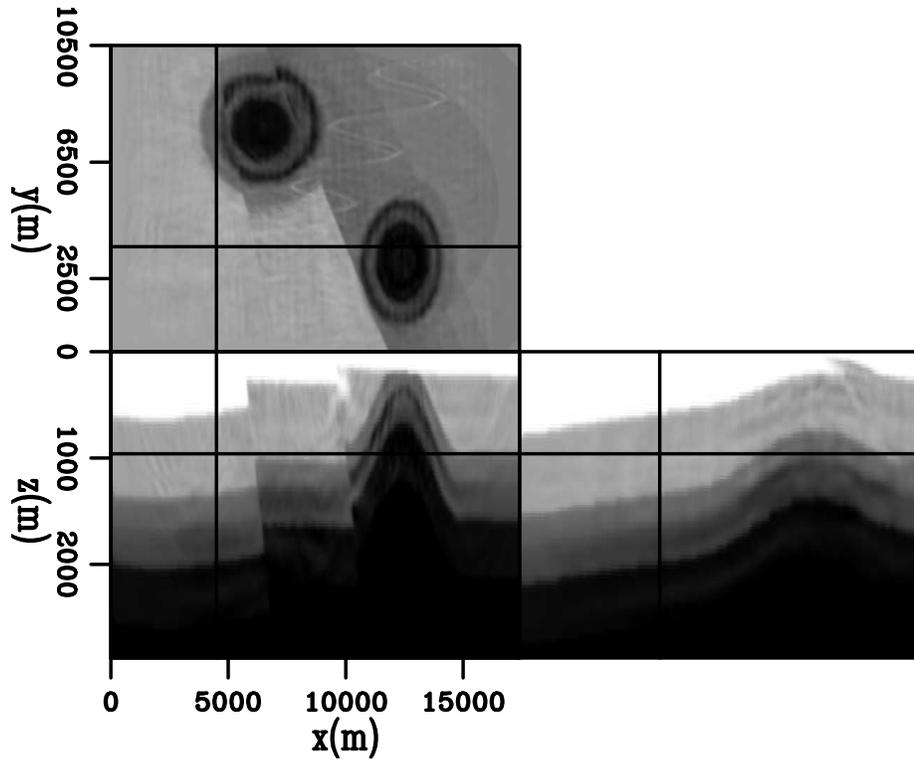


Figure 7: Waveform tomography result using traveltime result as starting model. [CR] xukai1/. vfwid

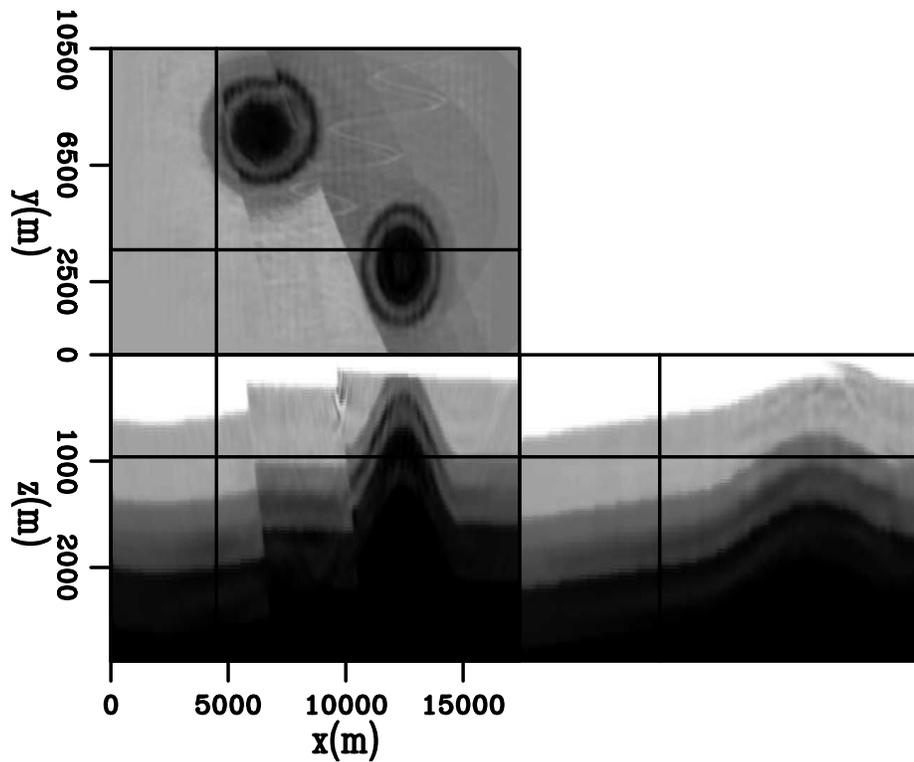


Figure 8: Waveform tomography result using the starting model. [CR] xukai1/. vfwid

steplength search. This result in smaller steplength per iteration and more iterations for tomography. Thirty iterations were run for waveform tomography from travelttime tomography result. Forty iterations were run for waveform tomography from starting model. This way, both waveform tomography results are derived using the same amount of computation, and comparison makes more sense from practical point of view. The other practicality comes into the choice of data for waveform tomography. Both waveform tomography used all the refraction/diving wave data rather than only using first break waveforms. Using the extra non-firstbreak refraction/diving wave barely increase computation while brings much more information into tomography, leading to better tomography results.

Waveform tomography result starting from travelttime tomography model is in figure 7. The resolution is much higher than that from travelttime tomography. Boundaries of all the velocity features become better defined. Even the two channels can be recognized where they are close to each other. Tomography from the starting model (figure 8) resolves the boundaries of velocity structures well, however, due to its inability of updating the bulk of the velocity structures, it did not converge to the correct velocity model. This is particularly obvious in the inline section, where direct waveform tomography solves part of the “bowl” type erosion boundary, yet is unable to solve the interior of the erosion.

## CONCLUSION

Wave-equation based methods perform better than ray based methods in building near-surface velocity model for geologically complex area. Physically, wave propagation is a better approximation than its high frequency approximations; numerically, rays easily become unstable with complex velocity structure. When both travelttime wave-equation tomography and waveform wave-equation tomography are performed, travelttime is better at resolving the bulk of velocity structures, while waveform is better at accurately define the boundaries of velocity structures. However, Waveform tomography can not converge to the true model if the bulk of the velocity structures are not resolved.

## REFERENCES

- Bevc, D., 1995, Imaging under rugged topography and complex velocity structure: SEP Ph.D Thesis.
- Clapp, R. G., 2013, Synthetic model building using a simplified basin modeling approach: SEP Report, **149**.
- Hampson, D. and B. Russell, 1984, First-break interpretation using generalized linear inversion: *Journal of CSEG*, **20**, 40–54.
- Hindriks, C. and D. Verschuur, 2001, Common focus point approach to complex near surface effects: SEG Expanded Abstracts.
- Luo, Y. and G. T. Schuster, 1991, Wave-equation travelttime inversion: *Geophysics*, **56**, 645–653.
- Marsden, D., 1993, Statics corrections-a review: *The Leading Edge*.
- Mora, P., 1987, Elastic wavefield inversion: SEP Ph.D Thesis.
- Olson, K. B., 1984, A stable and flexible procedure for the inverse modelling of seismic first arrivals: *Geophysical Prospecting*, **37**, 455–465.
- Pratt, R. G., C. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in

- frequency domain seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Ravaut, C., S. Operto, L. Imbrota, J. Virieux, A. Herrero, and P. Dell’Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: *Geophysical Journal International*, **159**, 1032–1056.
- Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: *SEG Expanded Abstracts*.
- , 2013, Anti-noise wave-equation travelttime inversion and application to salt estimation: *SEP Report*, **149**.
- Shen, X., T. Tonellot, Y. Luo, T. Keho, and R. Ley, 2012, A new waveform inversion workflow: Application to near-surface velocity estimation in saudi arabia: *SEG Abstracts*.
- Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Early arrival waveform tomography on near-surface refraction data: *Geophysics*, **71**, U47–U57.
- Sirgue, L., O.I.Barkved, J. V. Gestel, O. Askim, and R. Kommedal, 2009, 3d waveform inversion on valhall wide-azimuth obc: *EAGE 71th Conference*.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- White, D. J., 1989, Two-dimensional seismic refraction tomography: *Geophysical Journal International*, **97**, 223–245.

## APPENDIX B

### PSEUDO CODE FOR GRADIENT CALCULATION USING DIFFERENT BOUNDARY CONDITIONS

---

**Algorithm 2** Pseudo code of gradient calculation using absorbing boundary condition

---

```

for  $is = 1, ns$  do
  Forward wavefield propagation, generate modeled data and record source wavefield;
  Calculate data residual using recorded data and modeled data;
  Reverse-time propagation of residual data, correlate residual wavefield with source
  wavefield to generate gradient;
end for

```

---



---

**Algorithm 3** Pseudo code of gradient calculation using random boundary condition

---

```

for  $is = 1, ns$  do
  forward wavefield propagation using absorbing boundary, generate modeled data, do
  not record wavefield;
  forward wavefield propagation using random boundary, record last two time slices of
  source wavefield;
  calculate data residual using recorded data and modeled data;
  reverse time propagation of residual data and source wavefield using random boundary,
  generate gradient on the fly;
end for

```

---

# Hessian analysis of tomographic full waveform inversion operators

*Ali Almomin and Biondo Biondi*

## ABSTRACT

Tomographic full waveform inversion (TFWI) provides a framework to invert the seismic data that is immune to cycle-skipping problems. This is achieved by extending the wave equation and adding a spatial or temporal axis to the velocity model. For computational efficiency, the inversion is performed in a nested scheme. We examine the linearized component of the nested inversion scheme and present alternative fitting goals that have different properties compared to the original formulation. Then, we compute the Hessian matrix of both formulations as well as their individual operators to analyze the properties of each matrix. The analysis of the new formulation indicate an improved convergence behavior of inversion.

## INTRODUCTION

Tomographic Full Waveform Inversion (TFWI) (Symes, 2008; Biondi and Almomin, 2012) provides a way to overcome cycle-skipping problems by combining both FWI and wave-equation migration velocity analysis (WEMVA) techniques in a generalized framework. This generalized approach utilizes the components of all seismic data to invert for the medium parameters without cycle-skipping. This is achieved in two steps: first, extending the wave equation and adding a subsurface offset axis to the velocity model, and second, adding a regularization term that drives the solution towards the zero subsurface offset. Biondi and Almomin (2013b) presented an alternative extension using time shift instead of subsurface offset. In either setting, this velocity extension makes the propagation considerably more expensive because each multiplication by velocity becomes a convolution over the extended axis.

In a companion abstract (Biondi and Almomin, 2013a), we presented an approximation that significantly reduces the computational cost of TFWI by breaking the velocity model into a background component and a perturbation component. We achieve this in two steps. First, we set up a nested inversion scheme that utilizes the nonlinear modeling operator to update the residuals. Second, the two components of the gradient are first mixed and then separated based on a Fourier domain scale separation.

In this report, we examine the properties of the inner loop of the nested scheme. Then, we present alternative fitting goals that have different properties compared to the original formulation. Finally, we perform a Hessian analysis on both fitting goals and their operators in order to evaluate each formulation.

## FITTING GOALS

We start with the full waveform objective function,  $J_{\text{FWI}}$ , which we write as

$$J_{\text{FWI}}(\mathbf{s}^2) = \|\mathcal{L}(\mathbf{s}^2) - \mathbf{d}_{\text{obs}}\|_2^2, \quad (1)$$

where  $\mathbf{s}$  is the slowness model,  $\mathcal{L}(\mathbf{s}^2)$  is the wave-equation modeling operator, and  $\mathbf{d}_{\text{obs}}$  is the observed surface data. This first objective function represents the outer loop of the inversion. We now compute the nonlinear residuals  $\Delta\mathbf{d}$  as:

$$\Delta\mathbf{d}(\mathbf{s}^2) = \mathcal{L}(\mathbf{s}^2) - \mathbf{d}_{\text{obs}}. \quad (2)$$

The nonlinear residual will be used as the ‘‘observed’’ data for the inner loop. In the inner loop of the inversion, we need to separate the slowness model into a background and a perturbation as follows:

$$\mathbf{s}^2 = \mathbf{b}^2 + \mathbf{p}^2(\tau), \quad (3)$$

where  $\tau$  is the extension axis with time lags,  $\mathbf{b}^2$  is the background component, which is a smooth version of the slowness squared and  $\mathbf{p}^2(\tau)$  is the perturbation component. The perturbation component can extend across several subsurface offsets or time shifts so it is important to keep its extended axis. On the other hand, the background component is not expected to generate reflections that would be grossly time shifted with respect to the recorded data, and it thus safe to reduce its extension. This greatly reduces our cost since the convolution with slowness in propagation becomes a multiplication. The model separation allows us to use a linearized (Born) operator  $\tilde{\mathbf{L}}$  to model ‘‘linearized’’ data. The linearized objective function  $J_{\text{LTFWI}}$  can be written as:

$$J_{\text{LTFWI}}(\mathbf{b}^2, \mathbf{p}^2(\tau)) = \|\tilde{\mathbf{L}}(\mathbf{b}^2)\mathbf{p}^2(\tau) - \Delta\mathbf{d}\|_2^2 + \|\epsilon\mathbf{A}\mathbf{p}^2(\tau)\|_2^2, \quad (4)$$

where  $\tilde{\mathbf{L}}$  is the Born modeling operator,  $\mathbf{A}$  is the regularization operator that attempts to focus the extended model  $\mathbf{p}^2(\tau)$  and  $\epsilon$  is a scalar to balance the two terms of the objective function. The Born modeling operator is linear with respect to perturbation but nonlinear with respect to the background component. Hence, another linearization around the current value for both model parameters is required to compute the gradient. This is achieved by Taylor’s expansion of the operator around the model estimates at the current inner loop index  $i$ , i.e.  $\mathbf{b}_i^2$  and  $\mathbf{p}_i^2(\tau)$ , and dropping the higher-order terms as follows:

$$\begin{aligned} \tilde{\mathbf{L}}(\mathbf{b}_i^2 + \Delta\mathbf{b}^2) [\mathbf{p}_i^2(\tau) + \Delta\mathbf{p}^2(\tau)] &\approx \tilde{\mathbf{L}}(\mathbf{b}_i^2)\mathbf{p}_i^2(\tau) + \tilde{\mathbf{L}}(\mathbf{b}_i^2)\Delta\mathbf{p}^2(\tau) + \frac{\partial\tilde{\mathbf{L}}}{\partial\mathbf{b}^2}(\mathbf{b}_i^2) [\mathbf{p}_i^2(\tau)] \Delta\mathbf{b}^2 \\ &= \tilde{\mathbf{L}}(\mathbf{b}_i^2)\mathbf{p}_i^2(\tau) + \tilde{\mathbf{L}}(\mathbf{b}_i^2)\Delta\mathbf{p}^2(\tau) + \mathbf{T}(\mathbf{b}_i^2)[\mathbf{p}_i^2(\tau)]\Delta\mathbf{b}^2. \end{aligned} \quad (5)$$

The tomographic operator  $\mathbf{T}[\mathbf{p}^2(\tau)]$  correlates a background and a scattered wavefield from both the source and receiver sides. The scattered wavefields are computed by correlating a background wavefield with the perturbation model  $\mathbf{p}_i^2$  and then propagating again to all model locations. This operator is similar to the WEMVA operator except that in the WEMVA operator, we keep the data fixed and vary the image while here it is the other around for this tomographic operator.

We can now define the linearized form of  $J_{\text{LTFWI}}$  as follows:

$$\begin{aligned} J_{\text{LTFWI}}(\Delta\mathbf{b}^2, \Delta\mathbf{p}^2(\tau)) &= \|\tilde{\mathbf{L}}(\mathbf{b}_i^2)\Delta\mathbf{p}^2(\tau) + \mathbf{T}(\mathbf{b}_i^2)[\mathbf{p}_i^2(\tau)]\Delta\mathbf{b}^2 - (\Delta\mathbf{d} - \tilde{\mathbf{L}}(\mathbf{b}_i^2)\mathbf{p}_i^2(\tau))\|_2^2 \\ &\quad + \|\epsilon\mathbf{A}\Delta\mathbf{p}^2(\tau) - (-\epsilon\mathbf{A}\mathbf{p}_i^2(\tau))\|_2^2, \end{aligned} \quad (6)$$

where the fitting goals can be written in matrix notation as follows:

$$\begin{bmatrix} \Delta \mathbf{d} - \tilde{\mathbf{L}}(\mathbf{b}_i^2) \mathbf{p}_i^2(\tau) \\ -\epsilon \mathbf{A} \mathbf{p}_i^2(\tau) \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{L}}(\mathbf{b}_i^2) & \mathbf{T}(\mathbf{b}_i^2) [\mathbf{p}_i^2(\tau)] \\ \epsilon \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{p}^2(\tau) \\ \Delta \mathbf{b}^2 \end{bmatrix}. \quad (7)$$

To summaries, this scheme has three levels: an outer loop that tries to fit the data with the nonlinear operator (equation 1), an inner loop that tries to fit the nonlinear residual with the Born operator (equation 4), and finally a linearized loop that computes the updates for each model component (equation 6). When we start the inner loop, we set  $\mathbf{b}^2 = \mathbf{s}^2$  and  $\mathbf{p}^2 = 0$ . Once we iterate enough times within the inner loop, we update the velocity model with both the background and the perturbation at  $\tau = 0$ . Appendix A describes the derivation of the Born and tomographic operators and how to numerically evaluate them and their adjoints.

There is one potential issue in the previous formulation, particularly in equation 4. The Born operator will attempt to match the first-order scattering from the background while the focusing term will attempt to force the extended model into the zero lag. This means that the nonlinear residuals will be consistent with one term of the objective function  $J_{\text{LTFWI}}$  but not both, resulting in these parts of the objective function to compete with each other. This behavior could slow down the convergence of the inversion.

In order to avoid conflicting terms in the objective function, we propose to set alternative fitting goals such that the data fitting term is independent of the regularization term. One way to achieve this is to first compute a perturbation model  $\widehat{\mathbf{p}}^2(\tau)$  as follows:

$$\widehat{\mathbf{p}}^2(\tau) = \alpha \tilde{\mathbf{L}}'(\mathbf{b}_i^2) \Delta \mathbf{d}, \quad (8)$$

where  $'$  denotes the adjoint and  $\alpha$  minimizes the following objective function:

$$J_{\widehat{\mathbf{p}}^2}(\alpha) = \|\Delta \mathbf{d} - \alpha \tilde{\mathbf{L}}(\mathbf{b}_i^2) \tilde{\mathbf{L}}'(\mathbf{b}_i^2) \Delta \mathbf{d}\|_2^2. \quad (9)$$

Once the perturbation model is computed, the independent fitting goals can be set as follows:

$$\begin{bmatrix} \Delta \mathbf{d} - \tilde{\mathbf{L}}(\mathbf{b}_i^2) \widehat{\mathbf{p}}^2(\tau) \\ -\epsilon \tilde{\mathbf{L}}(\mathbf{b}_i^2) \mathbf{A}' \mathbf{A} \widehat{\mathbf{p}}^2(\tau) \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{L}}(\mathbf{b}_i^2) & 0 \\ 0 & \mathbf{T}(\mathbf{b}_i^2) [\widehat{\mathbf{p}}^2(\tau)] \end{bmatrix} \begin{bmatrix} \Delta \mathbf{p}^2(\tau) \\ \Delta \mathbf{b}^2 \end{bmatrix}. \quad (10)$$

The residual in the first row is only a data fitting residual caused by the difference between the adjoint and inverse of  $\tilde{\mathbf{L}}$  whereas the residual in the second row is only a model regularization residual caused by not focused  $\widehat{\mathbf{p}}^2(\tau)$ . The perturbation model  $\widehat{\mathbf{p}}^2(\tau)$  has a slightly different role than  $\mathbf{p}_i^2(\tau)$  in the first formulation. In the first formulation,  $\mathbf{p}_i^2(\tau)$  starts at zero but slowly change with the iterations of the inner loop whereas  $\widehat{\mathbf{p}}^2(\tau)$  is computed as a scaled perturbation model before optimizing the fitting goals in equation 10.

## HESSIAN ANALYSIS

A 1D model with a Gaussian anomaly (representing  $\Delta \mathbf{b}$ ) and a spike (representing  $\mathbf{p}$ ) is used for the synthetic examples. The true velocity is shown in Figure 1. The background velocity is 3 km/s and the Gaussian anomaly is located at 1 km. The perturbation spike is

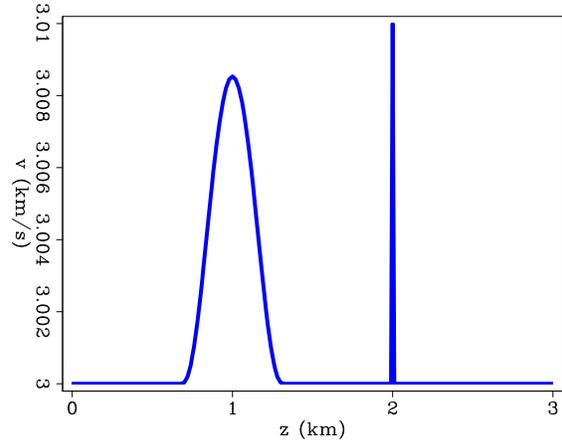


Figure 1: The true velocity model.

[ER] `ali1/. hessa-v`

located at 2 km. A Ricker wavelet with a fundamental frequency of 15 Hz is used to model the data. There is one source-receiver pair at the top of the model.

To test the  $\tilde{\mathbf{L}}$  operator, I start from a propagation velocity model that is the same as the true model except for the perturbation spike. The result of applying the adjoint of  $\tilde{\mathbf{L}}$  operator is shown in Figure 2. Next, I test the tomographic operator by computing the data residual due to removing the Gaussian anomaly. I use the results of applying the adjoint of  $\tilde{\mathbf{L}}$  as perturbation to estimate the update of the propagation velocity. The results of using the adjoint tomographic operator is shown in Figure 3. Both operators have the correct direction for the anomaly.

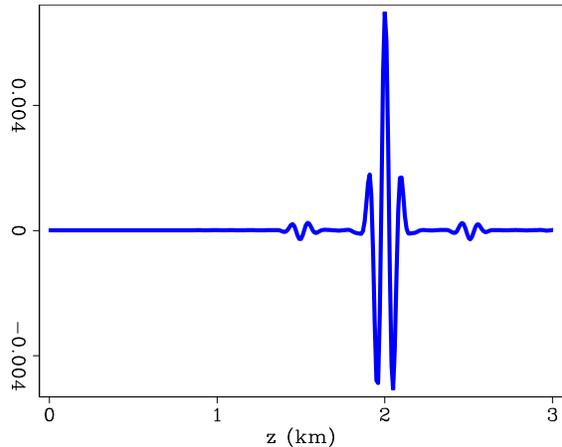


Figure 2: The Born operator adjoint on the data residual.

[ER] `ali1/. hessa-m`

To compare the properties of both total operators in equations 7 and 10, we first examine the properties of their individual operators. We first compute the Hessian matrix of the  $\tilde{\mathbf{L}}$  operator at  $\tau = 0$  as shown in Figure 4 and the singular values of that Hessian matrix as shown in Figure 5. The matrix is mostly diagonal with the imprint of the wavelet around the main diagonal and small artifacts due to the boundary conditions. The singular values drop significantly after approximately 100 values.

Next, we compute the Hessian matrix of the  $\tilde{\mathbf{L}}$  operator with  $\tau$  lags from -10 to 10 samples as shown in Figure 6 and the singular values of that Hessian matrix is shown in Figure 7. The Hessian matrix is now very non-diagonal due to the convolution by the

Figure 3: The tomographic adjoint on the data residual. **[ER]**  
 ali1/. hessa-gbm

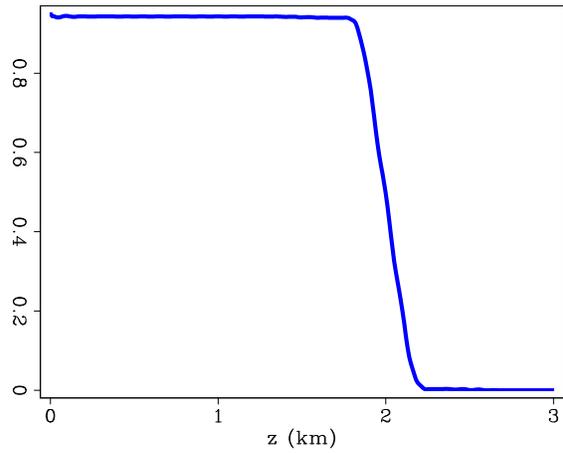


Figure 4: Hessian matrix of the  $\tilde{\mathbf{L}}$  operator at  $\tau = 0$ . **[ER]**  
 ali1/. hessa-hess10

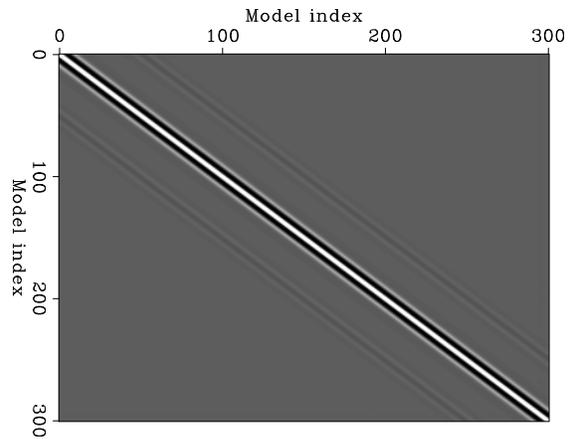
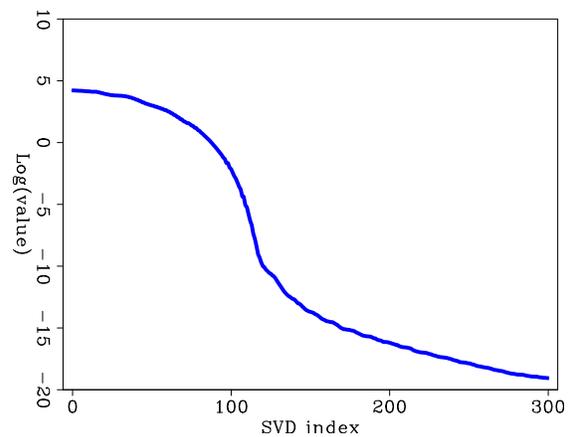


Figure 5: Singular values of Hessian matrix of the  $\tilde{\mathbf{L}}$  operator at  $\tau = 0$  in log scale. **[ER]**  
 ali1/. hessa-svdL0



extended model. The singular values are also higher than those shown in Figure 5.

Figure 6: Hessian matrix of the  $\tilde{\mathbf{L}}$  operator. [ER] ali1/. hessa-hess1

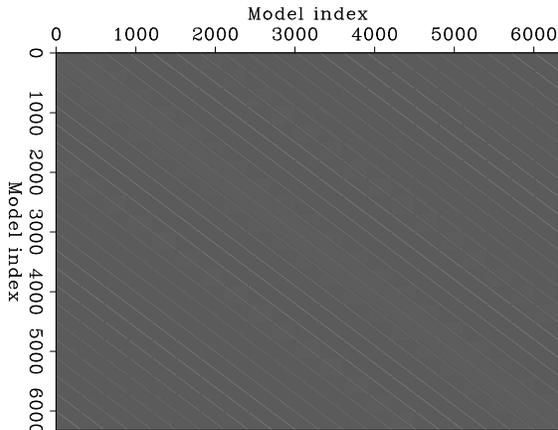
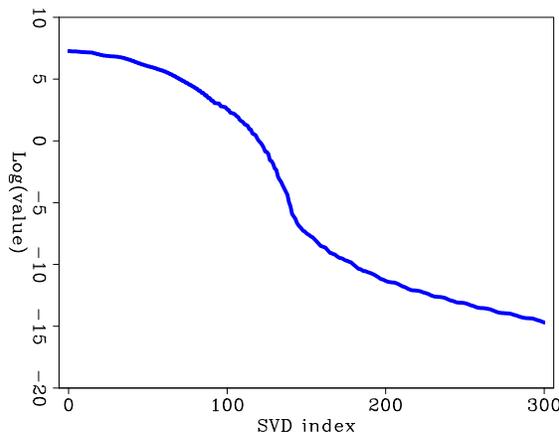


Figure 7: Singular values of Hessian matrix of the  $\tilde{\mathbf{L}}$  operator in log scale. [ER] ali1/. hessa-svdL



The last individual operator to examine is the tomographic operator  $\mathbf{T}[\mathbf{p}^2(\tau)]$ . The Hessian matrix of the  $\mathbf{T}[\mathbf{p}^2(\tau)]$  operator is shown in Figure 8 and the singular values of that Hessian matrix as shown in Figure 9. The hessian matrix is much less diagonal compared to the  $\tilde{\mathbf{L}}$  operator due to the integral nature of the tomographic operator. The plot of the singular values is very different than those of the  $\tilde{\mathbf{L}}$ . The singular values for the  $\mathbf{T}[\mathbf{p}^2(\tau)]$  operator have a relatively larger first singular value and the following values drop at a steeper rate.

Now, we compute the Hessian matrix of the total in Equation 7 and its singular values as shown in Figure 10. The singular values of this operator drop very slowly and become almost constant after 100 values. This is largely due to having  $\tilde{\mathbf{L}}$  at the off-diagonal of the total operator. Having that many large singular values indicate slow convergence rate of the inversion.

Finally, we compute the Hessian matrix of the total in Equation 10 and its singular values as shown in Figure 11. By comparing the these singular values to those in Figure 10, we see a large decrease in the values. Moreover, the singular values of the new operator drop at a much steeper rate. These changes indicate that inverting the new total operator will have better convergence rate and properties.

Figure 8: Hessian matrix of the  $\mathbf{T}[\mathbf{p}^2(\tau)]$  operator. [ER] `ali1/. hessa-hess2`

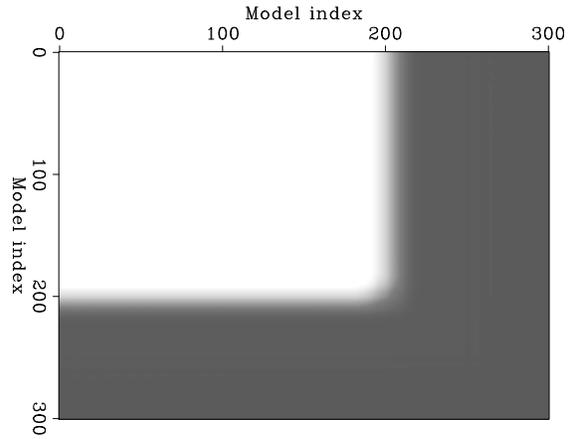


Figure 9: Singular values of Hessian matrix of the  $\mathbf{T}[\mathbf{p}^2(\tau)]$  operator in log scale. [ER] `ali1/. hessa-svdT`

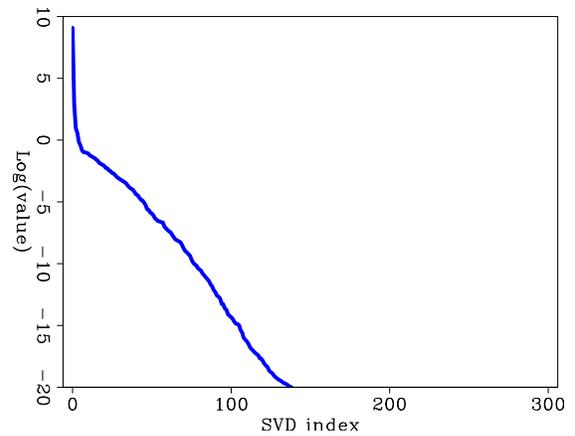


Figure 10: Singular values of Hessian matrix of the total operator in Equation 7 in log scale. [ER] `ali1/. hessa-svdO1`

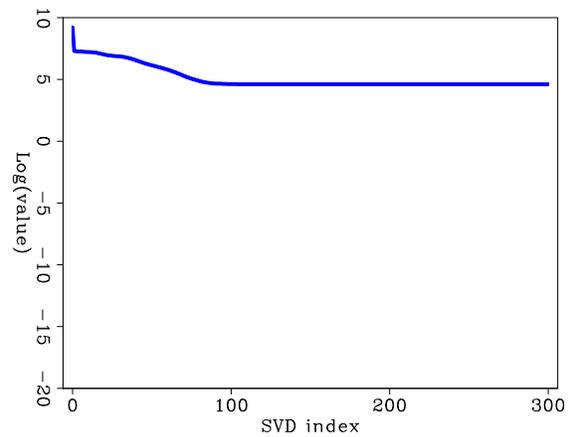


Figure 11: Singular values of Hessian matrix of the total operator in Equation 10 in log scale. [ER]

ali1/. hessa-svdO2



## CONCLUSIONS

Tomographic full waveform inversion (TFWI) requires separating the slowness model into background and perturbation in order to utilize an extended model efficiently. However, a direct minimization of the resulting objective function can potentially have bad convergence properties due to the conflicting terms in the objective function. This is also verified by computing the Hessian matrix of the total operator and its singular values. We proposed an alternative way to minimize the objective function of TFWI without conflicting terms in the objective function by separating the residuals into independent data-fitting residual and model regularization residual. The singular values of the new total operator indicate an improved convergence behavior of inversion. Further testing of both operator is needed to confirm these findings.

## APPENDIX A

The wave-equation modeling operator  $\mathcal{L}(\mathbf{s}^2)$  is evaluated by solving the equation

$$[\mathbf{s}^2 \mathbf{D}_2 - \nabla^2] \mathbf{P} = \mathbf{f}, \quad (\text{C-1})$$

where  $\mathbf{P}$  is the pressure field,  $\mathbf{D}_2$  is a finite-difference representation of the second derivative in time,  $\nabla^2$  is a finite-difference representation of the Laplacian, and  $\mathbf{f}$  is the source function. The minimization of the objective function 4 requires the computation of the gradients with respect to both  $\mathbf{b}$  and  $\mathbf{p}$ . These gradients can be computed by a perturbation analysis of the modeling operator  $\mathcal{L}(\mathbf{s}^2)$ . Rewriting the perturbed modeling equation in terms of  $\mathbf{b}$  and  $\mathbf{p}$ , and the incident wavefield,  $\mathbf{P}_i$  and the scattered wavefield,  $\mathbf{P}_s$ , we obtain the Born modeling operator  $\tilde{\mathbf{L}}$  by the coupled equations

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{C-2})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_s = \mathbf{p}_i^2(\tau) \overset{\tau}{*} \mathbf{D}_2 \mathbf{P}_i, \quad (\text{C-3})$$

where  $\overset{\tau}{*}$  denotes convolution in  $\tau$ . Introducing the perturbations,  $\delta \mathbf{b}^2$  and  $\delta \mathbf{p}^2$ , in the two model variables into equations C-2–C-3, and introducing the corresponding perturbations

in the wavefields,  $\delta\mathbf{P}_i$  and  $\delta\mathbf{P}_s$  yields the following perturbed system:

$$[(\mathbf{b}_i^2 + \delta\mathbf{b}^2) \mathbf{D}_2 - \nabla^2] (\mathbf{P}_i + \delta\mathbf{P}_i) = \mathbf{f}, \quad (\text{C-4})$$

$$[(\mathbf{b}_i^2 + \delta\mathbf{b}^2) \mathbf{D}_2 - \nabla^2] (\mathbf{P}_s + \delta\mathbf{P}_s) = [\mathbf{p}_i^2(\tau) + \delta\mathbf{p}^2(\tau)]^{\tau} \mathbf{D}_2 (\mathbf{P}_i + \delta\mathbf{P}_i). \quad (\text{C-5})$$

By setting  $\delta\mathbf{b}^2 = 0$  in equations C-4–C-5, we derive the following system of equations:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{C-6})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P}_s = \delta\mathbf{p}^2(\tau)^{\tau} \mathbf{D}_2 \mathbf{P}_i, \quad (\text{C-7})$$

which can be used to evaluate the perturbations in the scattered wavefield  $\delta\mathbf{P}_s$ , and consequently in the recorded data  $\delta\mathbf{d} = \mathbf{S}\delta\mathbf{P}_s$ , caused by perturbations  $\delta\mathbf{p}^2$ .

We can derive the data-space tomographic operator,  $\mathbf{T}$ , which relates perturbations in the scattered wavefield,  $\delta\mathbf{P}_s$ , to perturbations in the background model,  $\delta\mathbf{b}^2$ , by setting  $\delta\mathbf{p}^2 = 0$  and neglecting the higher-order terms in  $\delta\mathbf{b}^2$  in equations C-4–C-5.

This tomographic operator is equal to the sum of two operators,  $\mathbf{T}_i$  and  $\mathbf{T}_s$ . The first operator,  $\mathbf{T}_i$ , models perturbations in the scattered wavefield caused by perturbations in the propagation of the incident wavefield:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{C-8})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P}_i = \delta\mathbf{b}^2 \mathbf{D}_2 \mathbf{P}_i, \quad (\text{C-9})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P}_s = \mathbf{p}_i^2(\tau)^{\tau} \mathbf{D}_2 \delta\mathbf{P}_i, \quad (\text{C-10})$$

The second operator,  $\mathbf{T}_s$ , models perturbations in the scattered wavefield caused by perturbations in the propagation of the scattered wavefield itself:

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_i = \mathbf{f}, \quad (\text{C-11})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \mathbf{P}_s = \mathbf{p}_i^2(\tau)^{\tau} \mathbf{D}_2 \mathbf{P}_i, \quad (\text{C-12})$$

$$[\mathbf{b}_i^2 \mathbf{D}_2 - \nabla^2] \delta\mathbf{P}_s = \delta\mathbf{b}^2 \mathbf{D}_2 \mathbf{P}_s, \quad (\text{C-13})$$

Both of these tomographic operators depend nonlinearly on  $\mathbf{b}_i$  and linearly on  $\mathbf{p}_i^2$ . They have zero output when  $\mathbf{p}_i^2$  is equal to zero; that is, at the first iteration of the inner loop because we set  $\mathbf{p}_o^2 = 0$  as a starting model of the inner iterations. However, as we update the linearization at each iteration, starting from the second iteration the output of  $\mathbf{T}$  becomes different from zero.

## REFERENCES

- Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEG Expanded Abstracts, **31**, 275–279.
- , 2013a, Simultaneous inversion of full data bandwidth by tomographic full waveform inversion (TFWI): SEP-Report, **150**, 21–42.
- , 2013b, Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis: SEP-Report, **149**, 37–50.
- Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.



# Simultaneous time-lapse full waveform inversion

Musa Maharramov and Biondo Biondi

## ABSTRACT

We propose a technique for improving the robustness of time-lapse full waveform inversion by reducing numerical artifacts that contaminate inverted model differences. More specifically, we demonstrate that simultaneously inverting for baseline and monitor models in combination with a Tikhonov regularization applied to the model difference can reduce acquisition-related repeatability issues and spurious numerical artifacts arising in separate baseline and monitor inversions. We demonstrate our method using a synthetic model problem and describe a simplified “cross-updating” approach that can be applied to large-scale time-lapse industrial problems using the existing FWI inversion tools.

## INTRODUCTION

Time-lapse full waveform inversion (FWI) (Watanabe et al., 2004; Routh et al., 2012) is a promising technique for time-lapse seismic imaging where production-induced subsurface model changes are within the FWI resolution. However, like alternative time-lapse techniques, time-lapse FWI is sensitive to *repeatability issues* (Asnaashari et al., 2012). Non-repeatable acquisition geometries (e.g., slightly shifted source and receiver positions), acquisition gaps (e.g., due to previously absent obstacles), different source signatures and measurement noise – all contribute to differences in the data from different survey vintages. Differences in the input datasets due to repeatability issues may easily mask out valuable production-induced changes. However, even with the simple noise-free synthetic data in the absence of *acquisition* repeatability issues, numerical artifacts may easily contaminate the inverted difference of monitor and baseline when practical limitations are imposed on solver iteration count. We propose a computationally feasible and robust time-lapse FWI that minimizes model differences outside of areas of production-induced change by simultaneously inverting for multiple models, and imposing a regularization condition on model differences.

## THE METHOD

Full waveform inversion is defined as solving the following optimization problem (Virieux and Operto, 2009; Plessix et al., 2010)

$$\|\mathbf{M}\mathbf{u} - \mathbf{d}\| \rightarrow \min \quad (1)$$

where  $\mathbf{M}$ ,  $\mathbf{d}$  are the measurement operator and data,  $\mathbf{u}$  is the solution of a forward-modeling problem

$$\mathbf{D}(\mathbf{m})\mathbf{u} = \phi, \quad (2)$$

where  $\mathbf{D}$  is the forward-modeling operator that depends on a *model* vector  $\mathbf{m}$  as a parameter, and  $\phi$  is a source. The minimization problem 1 is solved with respect to both the model  $\mathbf{m}$  and source  $\phi$ , or just the model. In the frequency-domain formulation of the acoustic waveform inversion, the forward-modeling equation 2 becomes

$$-\omega^2 u - v^2(x^1, \dots, x^n) \Delta u = \phi(\omega, x^1, \dots, x^n) \quad (3)$$

where  $\omega$  is a temporal frequency,  $n$  is the problem dimension,  $v$  is the acoustic wave propagation velocity. Values of the seismic velocity at all the points of the modeling domain constitute the model parameter vector  $\mathbf{m}$ . Rock density can be incorporated in the problem but we omit it here for simplicity. Direct problem 2 can be solved in the frequency domain, or in the time domain followed by a Fourier transform in time (Virieux et al., 2009). Inverse problem 1 is typically solved using a multiscale approach, from low to high frequencies, supplying the output of each frequency inversion to the next step.

FWI applications in time-lapse problems seek to recover *production-induced* changes in the subsurface model (Barkved and Kristiansen, 2005) using multiple data sets from different acquisition vintages. For two surveys sufficiently separated in time, we call such datasets (and the associated models) *baseline* and *monitor*.

Time-lapse FWI can be carried out by separately inverting the baseline and monitor models (*parallel difference*) or inverting them sequentially with e.g. the baseline supplied as a starting model for the monitor inversion (*sequential difference*). Another alternative is to apply the *double difference* method, with a baseline model inversion followed by a monitor inversion that solves the following optimization problem

$$\|(\mathbf{M}_m^s \mathbf{u}_m - \mathbf{M}_b^s \mathbf{u}_b) - (\mathbf{M}_m \mathbf{d}_m - \mathbf{M}_b \mathbf{d}_b)\| \rightarrow \min \quad (4)$$

by changing the monitor model (Watanabe et al., 2004). The subscripts in equation 4 denote the baseline and monitor surveys,  $\mathbf{d}$  denotes the field data, and the  $\mathbf{M}$ s are measurement operators that project the synthetic and field data onto a common grid, the superscript  $s$  denotes the measurement operators applied to the synthetic data. Although the double difference method offers potential advantages for tackling acquisition-related repeatability issues, it was shown to be more sensitive to uncorrelated noise (Asnaashari et al., 2012) in addition to requiring data projection onto a common grid.

In all of the described techniques optimization is carried out with respect to a *single* model, albeit of different vintages at different stages of the inversion. In our method we propose to invert for the baseline and monitor models *simultaneously* by solving the following optimization problem:

$$\alpha \|\mathbf{M}_b \mathbf{u}_b - \mathbf{d}_b\|^2 + \beta \|\mathbf{M}_m \mathbf{u}_m - \mathbf{d}_m\|^2 + \quad (5)$$

$$\gamma \|(\mathbf{M}_m^s \mathbf{u}_m - \mathbf{M}_b^s \mathbf{u}_b) - (\mathbf{M}_m \mathbf{d}_m - \mathbf{M}_b \mathbf{d}_b)\|^2 + \quad (6)$$

$$\|\mathbf{W}_b \mathbf{R}_b(\mathbf{m}_b - \mathbf{m}_b^{\text{PRIOR}})\|^2 + \quad (7)$$

$$\|\mathbf{W}_m \mathbf{R}_m(\mathbf{m}_m - \mathbf{m}_m^{\text{PRIOR}})\|^2 + \quad (8)$$

$$\|\mathbf{W} \mathbf{R}(\mathbf{m}_m - \mathbf{m}_b - \Delta \mathbf{m}^{\text{PRIOR}})\|^2 \rightarrow \min, \quad (9)$$

with respect to both the baseline and monitor models  $\mathbf{m}_b$  and  $\mathbf{m}_m$ . The terms 5 correspond to separate baseline and monitor inversions, the term 6 is the optional double difference term, the terms 7,8 are optional separate baseline and monitor inversion Tikhonov regularization terms (Aster et al., 2012), and the term 9 represents Tikhonov regularization of the

model difference. The  $\mathbf{R}$  and  $\mathbf{W}$  denote regularization and weighting operators respectively, with the subscript denoting the survey vintage where applicable.

A joint inversion approach was proposed earlier in applications to the linearized waveform inversion (Ayeni and Biondi, 2012). In this work, we propose a simultaneous full waveform inversion with special emphasis on the regularization of model difference in equation 9. Constraining the model difference where production-induced changes are expected to be negligible while simultaneously solving for both baseline and monitor models can be expected to reduce both spurious numerical artifacts and non-repeatable acquisition related artifacts in the model difference. The traditional sequential and double difference methods, on the other hand, make artifact reduction less effective by allowing only one model to change.

An implementation of the proposed simultaneous inversion algorithm requires solving a nonlinear optimization problem of twice the data and model dimension of problems 1 and 4. To allow an immediate practical application of the proposed methodology using the existing single-model solvers, we propose a *cross-updating* technique that offers a crude but remarkably effective approximation to minimizing the objective function 5 and 9. This cross-updating technique consists in one standard run of the sequential difference algorithm, followed by a second run with the inverted monitor model supplied as the starting model for the second baseline inversion

$$\mathbf{m}_{\text{INIT}} \rightarrow \text{BASE INV} \rightarrow \text{MON INV} \rightarrow \text{BASE INV} \rightarrow \text{MON INV}, \quad (10)$$

and taking the difference of the latest inverted monitor and baseline models. Process 10 can be considered an approximation to minimizing 5 and 9 because non-repeatable footprints of both inversions are accumulated in both models, canceling out in the difference. In both the full simultaneous inversion and cross-updating methods, we expect the models to be brought closer together wherever either of them cannot be reliably resolved. Note that this process is not guaranteed to improve the results of the individual inversions but is only proposed for improving the model difference.

## RESULTS

We demonstrate our proposed simultaneous inversion method and its cross-updating approximation 10 on a one-dimensional model problem. We are trying to recover the difference between the true baseline and monitor models, as respectively shown in Fig 1(b), 1(a). Equation 3 is discretized on the equispaced grid of 32 points spanning the modeling interval  $[0, 1]$ . We use different baseline and monitor source wavelets (see Fig 2) and use a multiscale inversion from 0.2Hz to 1Hz with a .2Hz step, solving the optimization problems 1 and 5,9 using the hybrid Fletcher-Reeves-Polar-Ribière method (Nocedal and Wright, 2006). In the first experiment the data is noise-free and receivers are located at every node of the computational grid. In the second experiment we add 5% high-wavenumber noise to the data and allow receiver positions to shift randomly by up to half the grid step. No prior model or model difference was used in the objective function 5,9.

To demonstrate that the improvements delivered by the cross-updating procedure 10 are not simply due to the increased number of iterations, we use the following procedure

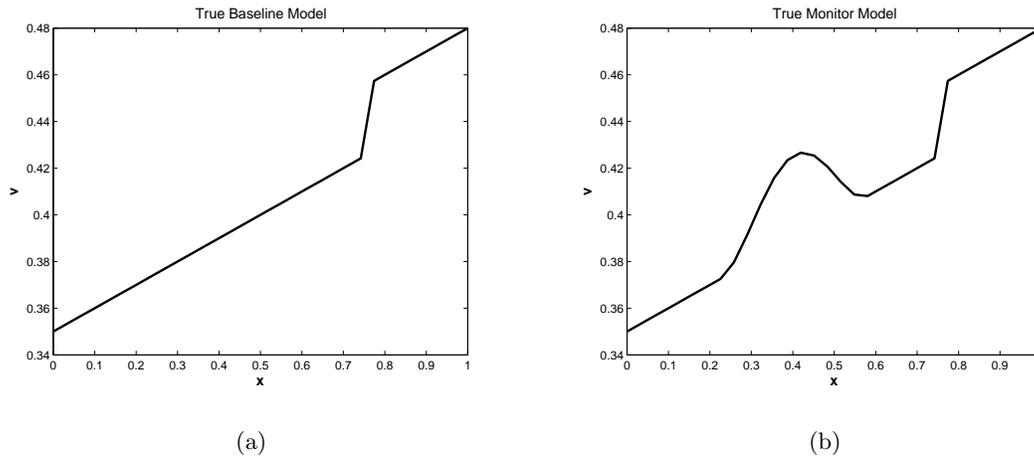
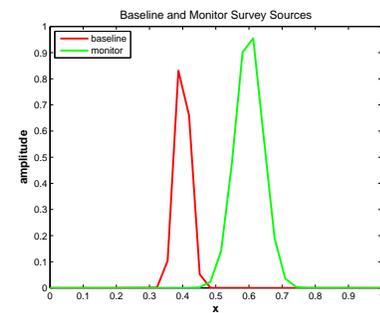


Figure 1: True baseline and monitor models for 1D simultaneous inversion test. [CR] `musal/. truebaseline,truemonitor`

Figure 2: Sources used for generating synthetic baseline and monitor surveys. [CR] `musal/. sources`



for “iterated” sequential inversion:

$$\begin{aligned} \mathbf{m}_{\text{INIT}} &\rightarrow \text{BASE INV} \rightarrow \text{MON INV} \rightarrow \text{MON INV} \\ \mathbf{m}_{\text{INIT}} &\rightarrow \text{BASE INV} \rightarrow \text{BASE INV}. \end{aligned} \quad (11)$$

Note that process 11 uses the same number of full waveform inversions as 10. We compare method 11 with the cross-updating 10 and the simultaneous inversion by optimizing the objective function 5,9. In our tests, we set the terms 6-8 to zero and use only a model difference regularization. We use Laplacian as the regularization operator  $\mathbf{R}$ , and the weighting operator  $\mathbf{W}$  is a simple mask zeroing out the area of expected model change.

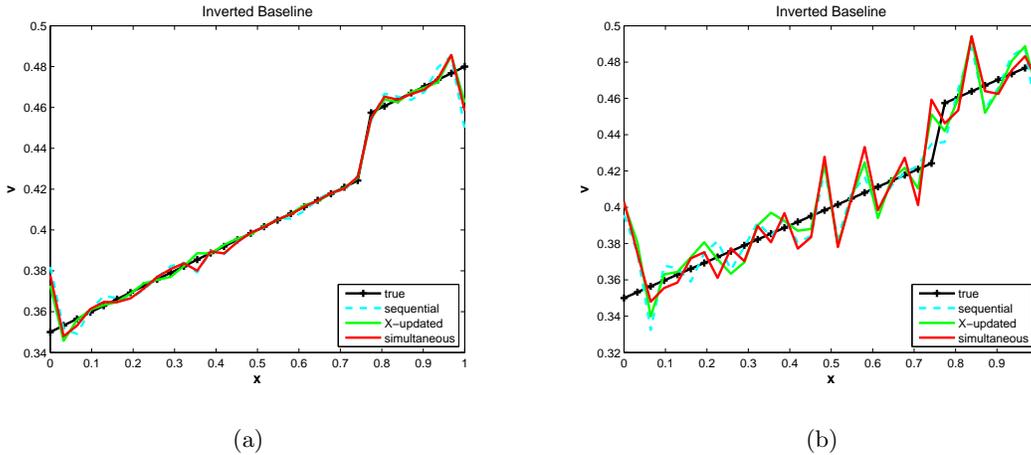


Figure 3: Inverted baseline: (a) no noise (b) 5% noise and  $.5\Delta x$  random receiver mispositioning. [CR] `musa1/. baseline,baselinenoise`

Fig 3(a), 4(a), 5(a) show the results of the three algorithms for the noise-free experiment. Both the simultaneous inversion result and the cross-updating approximation exhibit fewer artifacts outside of the area of production-induced changes than does the sequential approximation. Fig 3(b), 4(b), 5(b), 6(b) show the results of the second experiment with added noise and receiver mispositioning. The improvement achieved by the simultaneous inversion and cross-updating methods is quite dramatic, with the simple cross-updating method still delivering a good qualitative picture of the model difference. One key feature of the plotted models of Fig 3(b), 4(b) is that the inverted models seem to inherit each other’s peaks and deviations from the true model even though *different random noise* is added to the baseline and monitor data. This confirms our conjecture that the model-difference regularization will bring the models closer together where either model cannot be reliably resolved. Similar results have been achieved for multiple alternative noise realizations, with the simultaneous inversion always delivering the best *model difference* result and cross-updating providing a good qualitative approximation.

## CONCLUSIONS AND WAY FORWARD

We propose a new technique for time-lapse full waveform inversion that can provide a robust alternative to the existing methods. Applying the method to a one-dimensional

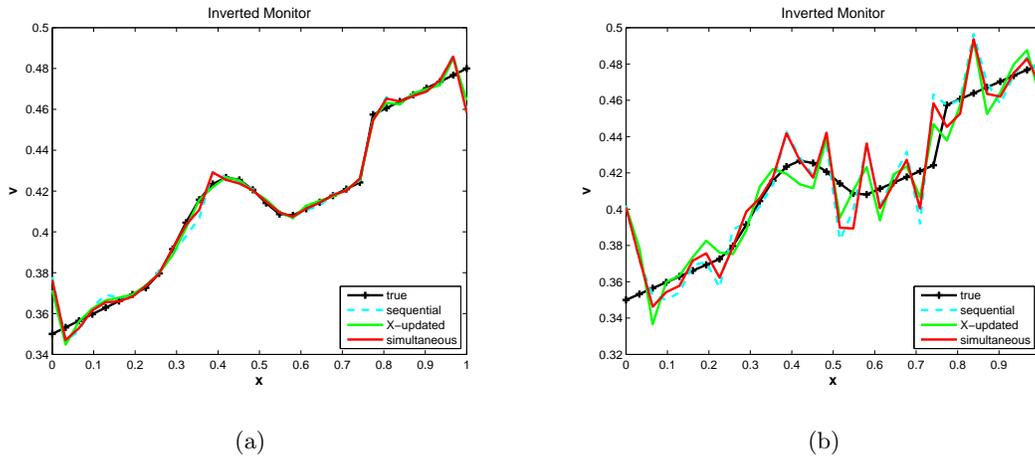


Figure 4: Inverted monitor: (a) no noise (b) 5% noise and  $.5\Delta x$  random receiver mispositioning. [CR] `musa1/. monitor,monitornoise`

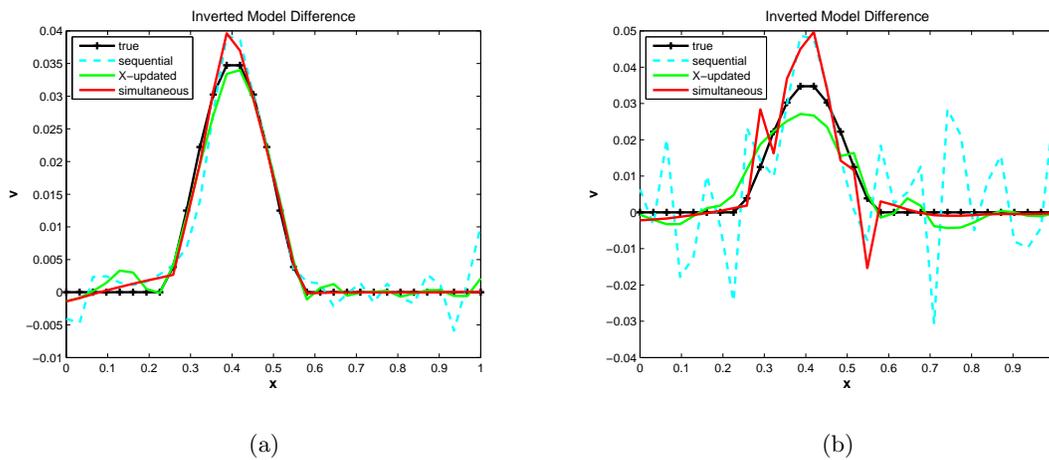


Figure 5: Inverted model difference: (a) no noise (b) 5% noise and  $.5\Delta x$  random receiver mispositioning. Note that the result of simultaneous inversion is closest to zero outside of the area of “production-induced” change. The cross-updating result is a good qualitative approximation to simultaneous inversion. [CR] `musa1/. difference,differencenoise`

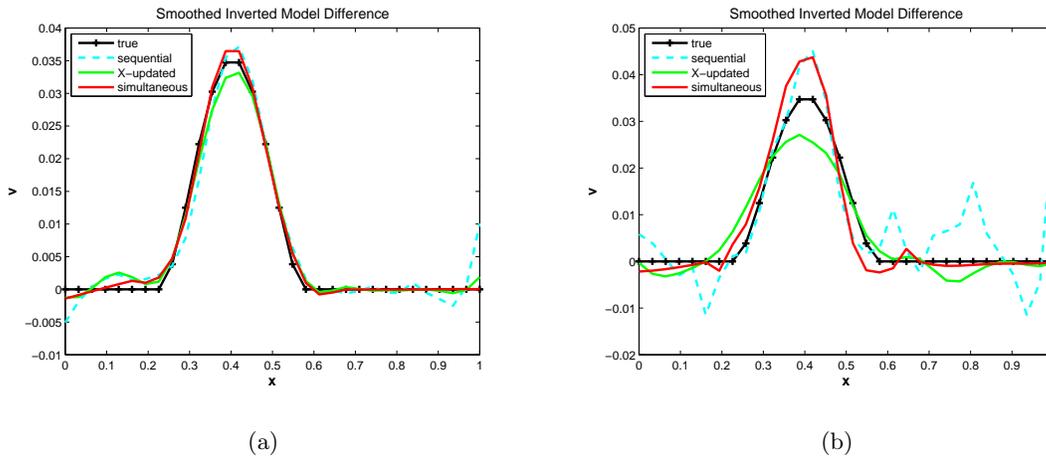


Figure 6: Smoothed inverted model difference: (a) no noise (b) 5% noise and  $.5\Delta x$  random receiver mispositioning. Even after smoothing the simultaneous inversion and its cross-updating approximation are superior to the result of iterated sequential inversion. [CR]

`musa1/. smoothedifference,smootheddifferencenoise`

test problem achieved a significant reduction of artifacts in the model difference. However, choice of the weighting operator  $\mathbf{W}$  is determined by prior knowledge of where production-induced velocity changes are likely to occur. In the absence of such prior knowledge, a frequency-dependent weighting of the model difference regularization term may have to be used to achieve a balance between the accuracy of data fitting and the desired properties of the model difference (Aster et al., 2012). In such a case, our proposed cross-updating method would offer an attractive alternative to the regularized simultaneous inversion as it delivers qualitatively accurate results without the need to specify regularization parameters.

The cross-updating technique does not require the development of any new tools and, in principle, can be applied to time-lapse FWI problems of any scale and complexity. The simultaneous inversion based on optimizing the objective function 5-9 can be applied to regularized time-lapse velocity inversion using a prior model difference derived from geomechanical studies (Maharramov, 2012).

We did not apply the regularized simultaneous inversion to the double difference method (i.e., with a nonzero term 6 in the objective function) and this will require further study. In our next work we intend to apply the simultaneous time-lapse FWI and cross-updating methods to field data, and study the feasibility of using geomechanical model-difference priors.

## REFERENCES

- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2012, Time-lapse imaging using regularized FWI: a robustness study: SEG Technical Program Expanded Abstracts, 1–5.
- Aster, R. C., B. Borchers, and C. H. Thurber, 2012, Parameter estimation and inverse problems: Elsevier.

- Ayeni, G. and B. Biondi, 2012, Time-lapse seismic imaging by linearized joint inversion a valhall field case study: SEG Technical Program Expanded Abstracts, 1–6.
- Barkved, O. and T. Kristiansen, 2005, Seismic time-lapse effects and stress changes: Examples from a compacting reservoir: *The Leading Edge*, **24**, 1244–1248.
- Maharramov, M., 2012, Identifying reservoir depletion patterns with applications to seismic imaging: SEP Report 147, 193–218.
- Nocedal, J. and S. Wright, 2006, Numerical optimization: Springer.
- Plessix, R.-E., S. Michelet, H. Rynja, H. Kuehl, C. Perkins, J. W. de Maag, and P. Hatchell, 2010, Some 3D applications of full waveform inversion in 3D Full Waveform Inversion Workshop – a Game Changing Technique?: 72nd EAGE Conference and Exhibition.
- Routh, P., G. Palacharla, I. Chikichev, and S. Lazaratos, 2012, Full wavefield inversion of time-lapse data for improved imaging and reservoir characterization: SEG Technical Program Expanded Abstracts, 1–6.
- Virieux, J. and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC127–WCC152.
- Virieux, J., S. Operto, B. H. Ali, R. Brossier, V. Etienne, F. Sourbier, L. Giraud, and A. Haidar, 2009, Seismic wave modeling for seismic imaging: *The Leading Edge - Special Section: Seismic Modeling*, 538–544.
- Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring: 74th SEG Conference and Exhibition, Expanded Abstracts, 2323–2326.

## Joint-LSRTM in practice with the Deimos ocean bottom field data set

*Mandy Wong, Biondo Biondi, and Shuki Ronen*

### ABSTRACT

We apply an adaptation of the least-squares reverse time migration (LSRTM) algorithm to the 3D Deimos ocean bottom field data set from the Gulf of Mexico. A simple data-fitting objective function may not be sufficient when applying LSRTM in practice. Some challenges arise because the recorded field data depart from the theory and assumption of the LSRTM operator. To optimize the inversion with the field data set, we include Laplacian preconditioning, salt-dimming data weighting, extended domain noise filtering, and regularization onto the LSRTM algorithm. Results from the 3D Deimos ocean bottom field dataset show an improvement when using joint LSRTM of primary and mirror signals over conventional imaging.

### INTRODUCTION

Depth migration images are often distorted by uneven subsurface illumination from acquisition geometry, complex overburden and band-limited wavefields. To correct for the effects of uneven illumination, the imaging problem can be posed as a linear inverse problem. Depth migration is equivalent to the adjoint of the Born-modelling operator. Instead of using the adjoint operator, we use the pseudo-inverse of the Born-modeling operator to optimally reconstruct the reflectivity. This inversion-based imaging method is also widely known as least-squares migration (LSM) (Lambare et al., 1992; Nemeth et al., 1999; Ronen and Liner, 2000) or linearized wavefield inversion (Clapp, 2005; Valenciano, 2008a).

In addition to least-squares migration, being able to co-image different types of seismic data can often add value to the image. In the case of ocean-bottom data, Wong et al. (2010) showed that using both primary (up-going) and mirror (down-going) signals can improve subsurface illumination. The mirror image has wider illumination. Depending on the offset range and the ocean-bottom depth, primary signal can often illuminate the subsurface at a wider reflection angle (Figure 1). This translates to higher image quality in the region illuminated by the primary. Instead of treating the primary image and the mirror image separately, we can combine the information from the two sets of data coherently by joint least-squares reverse-time migration (LSRTM)

LSM in general can boost up the signal for the true reflector and suppress noise. When applied to a field data set, the quality of the inversion can be compromised by various factors. As a result, adaptation and conditioning of the inversion is needed when applying to the field data set. We will point out a few of the challenges concerning the 3D Deimos ocean bottom data set and show some ways to address these issues. Some of the techniques include Laplacian preconditioning, salt-dimming data weighting (Wong et al.,

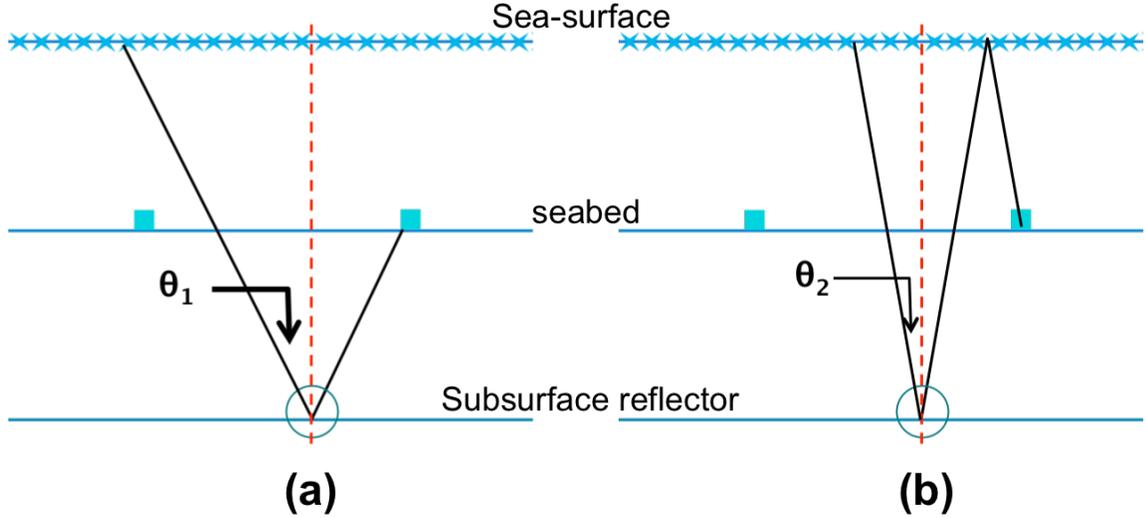


Figure 1: (a) shows the reflection angle ( $\theta_1$ ) by the primary reflection and (b) shows the reflection angle ( $\theta_2$ ) by the mirror reflection at the same sub-surface image location for ocean bottom node acquisition geometry. Given a survey with limited offset, the primary signal can often illuminate the subsurface at a wider reflection angle. [NR] mandy1/. primir

2010), extended domain noise filtering, and a regularization that impose continuity along the subsurface-angle axis.

In this paper, we will apply joint up-going (primary) and down-going (mirror) least-squares reverse-time migration (LSRTM) on the 3D Deimos ocean bottom node field data set. I will show some of the improvements as compared to conventional imaging.

## THEORY

### Least-squares reverse-time migration (LSRTM)

In reverse-time migration (RTM), the migration image is a linear operator applied to the recorded data,

$$\begin{aligned} m_{mig}(\mathbf{x}) &= \sum_{\mathbf{x}_r, \mathbf{x}_s, \omega} U_s^*(\mathbf{x}_s, \mathbf{x}, \omega) G^*(\mathbf{x}, \mathbf{x}_r) d(\mathbf{x}_r, \mathbf{x}_s, \omega), \\ &= \mathbf{L}' \mathbf{d}, \end{aligned} \quad (1)$$

where  $\omega$  is frequency and  $m(\mathbf{x})$  represents reflectivity at the image point  $\mathbf{x}$ . Additionally,  $U_s(\mathbf{x}_r, \mathbf{x}, \omega) = \omega^2 f_s(\omega) G(\mathbf{x}_r, \mathbf{x})$  is the source-side or incident wavefield.  $G(\mathbf{x}_1, \mathbf{x}_2)$  is the Green function that solves the two-way acoustic constant density equation. Often time, deconvolution is performed on the input data  $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$  and an appropriate waveform  $f_s(\omega)$  is used to generate the source-side wavefield. Note that the  $\omega$  dependence in  $G$  is suppressed. The RTM operator can be abbreviated into the operator form  $\mathbf{L}'$  acting on the data  $\mathbf{d}$ . In practice, the Green function is calculated using finite-difference time

domain technique and the multiplication in frequency domain is replaced by a zero-lag cross-correlation in the time domain.

To obtain a better reflectivity image, we can go beyond migration by formulating the imaging problem as a least-squares inversion problem. The solution  $m_{inv}(\mathbf{x})$  is obtained by minimizing the objective function  $S(\mathbf{m})$ , which is defined as the least-squares difference between the forward modeled data  $d_{mod}$  and the observed data  $d_{obs}$ .

$$S(\mathbf{m}) = \|\mathbf{d}_{mod} - \mathbf{d}_{obs}\|^2 = \|(\mathbf{L}\mathbf{m} + \mathbf{d}_o) - \mathbf{d}_{obs}\|^2 \quad (2)$$

The modeled data  $\mathbf{d}_{mod}$  can be broken down into two components: the linearly modeled data  $\mathbf{L}\mathbf{m}$  and the background data  $\mathbf{d}_o$ . The background data is the full wave-equation modeling of the data using the (background) migration velocity model.  $\mathbf{d}_o$  does not get updated, and remains the same over all iterations in the least-squares migration algorithm.

In LSRTM, the linear forward modeled data is defined to be the Born approximation of the linearized acoustic wave equation,

$$\begin{aligned} d_{mod}(\mathbf{x}_r, \mathbf{x}_s, \omega) &= \sum_{\mathbf{x}} U_s(\mathbf{x}_s, \mathbf{x}, \omega) m(\mathbf{x}) G(\mathbf{x}, \mathbf{x}_r), \\ \mathbf{d}_{mod} &= \mathbf{L}\mathbf{m}, \end{aligned} \quad (3)$$

It is important to point out that the forward modeling operator  $\mathbf{L}$  is the adjoint of the reverse-time migration operator  $\mathbf{L}'$ .

### Joint LSRTM of up/down-going P wave

Joint inversion of up- and down-going signals for ocean-bottom data can potentially be a better imaging technique than migrating either signal alone, because it combines information from both sets of signals. Ocean bottom data are first separated into acoustic up- and down-going components above the seafloor. The decomposed signals are then inverted to yield one optimally combined reflectivity image. The objective function for such an inversion is:

$$S(\mathbf{m}) = \|(\mathbf{L}_\uparrow \mathbf{m} + \mathbf{d}_o^\uparrow) - \mathbf{d}_{obs}^\uparrow\|^2 + \|(\mathbf{L}_\downarrow \mathbf{m} + \mathbf{d}_o^\downarrow) - \mathbf{d}_{obs}^\downarrow\|^2 \quad (4)$$

where  $\mathbf{L}_\uparrow$  and  $\mathbf{L}_\downarrow$  are modeling operators that linearly model the up- and down-going data. Two modified grids are used to forward model the lowest order of up- and down-going signals, namely the primary and the mirror reflection. The formulation of the modeling operator is summarized in Figure 2 and 3

Although LSM in general can boost up the signal for true reflectors and suppress noise, the quality of the inversion may be compromised when applied to a field data set. As a result, adaptation and conditioning of the inversion is needed when applying to the field data set. We will point out a few of the challenges and show some ways to address those issues.

### Laplacian Preconditioning

It has been observed that using the conventional cross-correlation imaging condition produces strong low-frequency migration artifacts in reverse-time migration. These artifacts result from the forward-scattered energy off sharp contrast in the velocity model. In LSRTM,

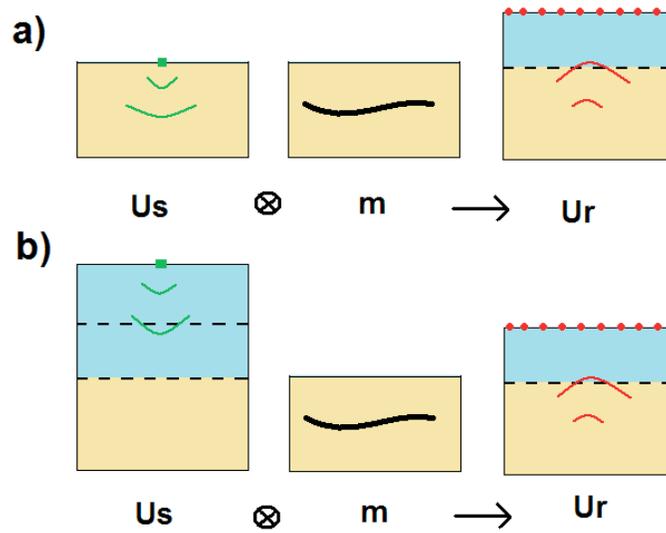


Figure 2: Forward modeling of (a) primary-only and (b) mirror-only data. The algorithm involves cross-correlating the source wavefield ( $U_s$ ) with the reflectivity model ( $m$ ) to generate the receiver wavefield ( $U_r$ ). Reciprocity is used here where the data, in common-receiver domain, is injected at the source location while the source wavelet is injected at the receiver location. Cross-correlation is done only with grid points below the seabed. [NR]

mandy1/. forward

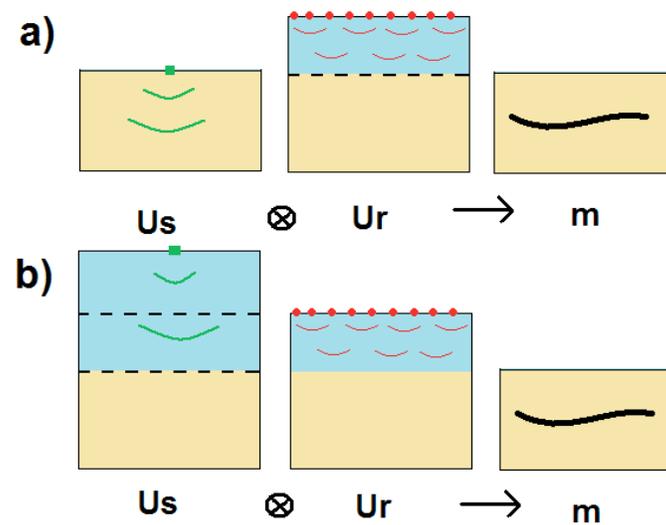


Figure 3: RTM of (a) primary-only and (b) mirror-only data. The algorithm involves cross-correlating the source wave field ( $U_s$ ) with the receiver wave field ( $U_r$ ) to generate the reflectivity model ( $m$ ). Cross-correlation is done only with grid points below the seabed. [NR]

mandy1/. reverse

a successive sequence of adjoint and modeling operations is required to iteratively change the initial solution into the inverted one. As a result, when the adjoint operation is used in LSRTM, the forward-scattering artifacts also exist in the gradient. Such artifacts hamper the convergence because unwanted energy is introduced into the gradient at every iteration. One way to suppress these artifacts is to apply a Laplacian filter to the stacked RTM image. Zhang and Sun (2009) have showed that applying the Laplacian filter (along with a  $1/\omega^2$  filter onto the data) is equivalent to applying a  $\cos^2\theta$  weight to the angle gather. This cosine factor down-weights contributions from large angles and effectively suppresses the forward-scattered artifacts when  $\theta = 90^\circ$ . A straightforward incorporation of this idea in LSRTM is to use the Laplacian filter as a preconditioner in the inversion. This can be done by doing a simple change of variable,

$$\mathbf{p} = \mathbf{A}\mathbf{m}, \quad (5)$$

where  $\mathbf{A}$  represents the Laplacian operator and  $\mathbf{p}$  is the preconditioned variable. Note that the Laplacian operation is self-adjoint. The objective function in terms of the preconditioned variable  $\mathbf{p}$  then becomes:

$$S(\mathbf{p}) = \|(\mathbf{L}_\uparrow \mathbf{A}\mathbf{p} + \mathbf{d}_\circ^\uparrow) - \mathbf{d}_{\text{obs}}^\uparrow\|^2 + \|(\mathbf{L}_\downarrow \mathbf{A}\mathbf{p} + \mathbf{d}_\circ^\downarrow) - \mathbf{d}_{\text{obs}}^\downarrow\|^2 \quad (6)$$

After inverting for  $\mathbf{p}_{\text{inv}}$ , the inverted model is then recovered by applying the inverse of the Laplacian operator,

$$\mathbf{m}_{\text{inv}} = \mathbf{A}^{-1}\mathbf{p}_{\text{inv}}. \quad (7)$$

The inverse of the Laplacian operator is singular. In practice, we applied a pseudo inverse in the wavenumber domain by multiplying a  $\frac{1}{k_z^2 + \epsilon}$  filter with a small  $\epsilon$  constant.

## Salt Dimming

The objective function in equation 2 requires the background data term  $\mathbf{d}_\circ$  to be subtracted from the observed data. Essentially,  $\mathbf{d}_\circ$  is the forward modeling using the migration velocity model. In most cases when the velocity is smoothly varying, this term is negligible and we can just ignore it in the inversion. When the velocity field has a sharp contrast, this term is non-trivial. However, subtracting the background data term is actually a difficult task. The theory that generates the background data is just an approximation to the complex Earth mechanism that generates the observed data. This makes following equation 2 impractical. Salt-dimming (Wong, 2013) is introduced as a way to work around this problem.

Salt-dimming aims to down-weight the salt reflection energy in the data space so that the inversion can minimize other regions in the model. This corresponds to the following objective function:

$$S(\mathbf{m}) = \|\mathbf{W}_s(\mathbf{L}\mathbf{m} - \mathbf{d}^{\text{obs}})\|^2,$$

where  $\mathbf{W}_s$  is the data weighting function that down-weights the salt reflection energy. This can be done by forward-modeling the salt reflection using the migration velocity. The next step is to calculate an envelope around the salt energy. The data weighting function can then be defined by assigning a small value to the salt reflection envelope. The objective function with preconditioning and salt-dimming then becomes:

$$S(\mathbf{p}) = \|\mathbf{W}_s^\uparrow(\mathbf{L}_\uparrow\mathbf{A}\mathbf{p} - \mathbf{d}_{\text{obs}}^\uparrow)\|^2 + \|\mathbf{W}_s^\downarrow(\mathbf{L}_\downarrow\mathbf{A}\mathbf{p} - \mathbf{d}_{\text{obs}}^\downarrow)\|^2. \quad (8)$$

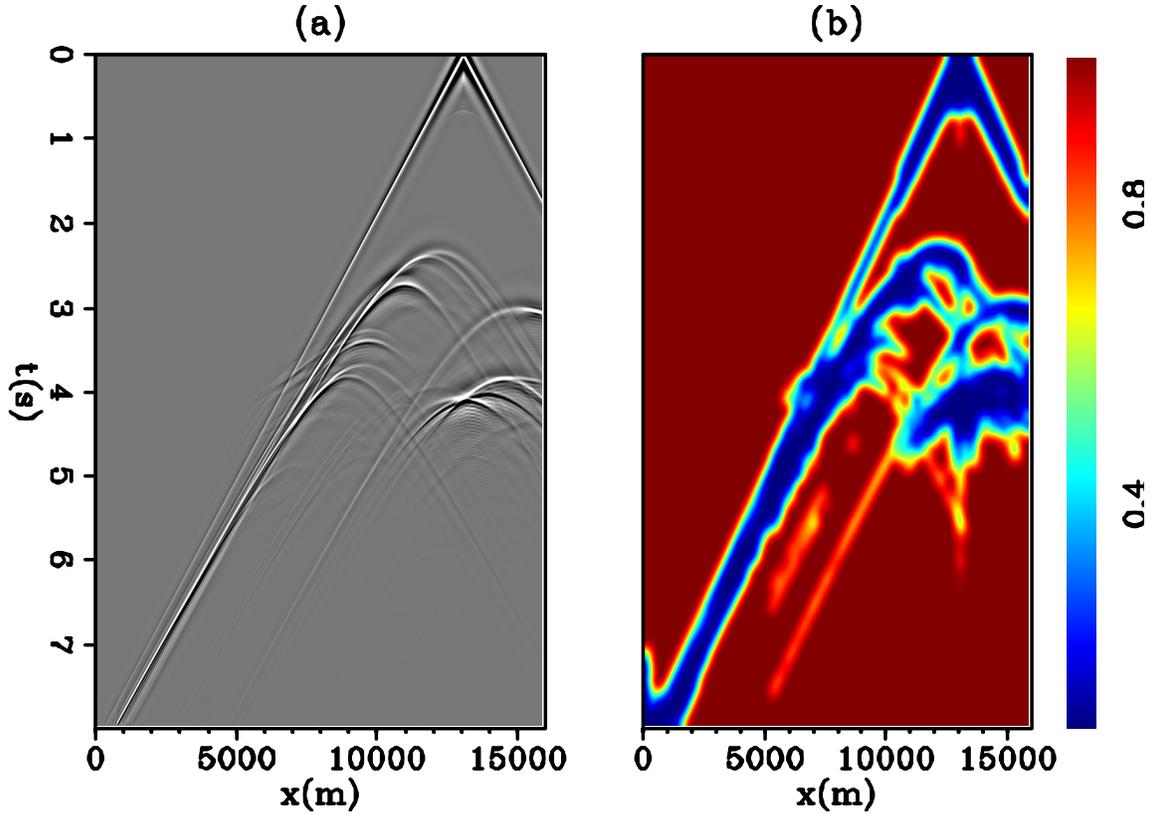


Figure 4: (a) Background data created by forward modeling with the migration velocity and (b) salt-dimming weight ( $W_s$ ) generated with the background data. [CR] mandy1/. Vfig5wgt

Figure 4a shows the forward modeling of one common-receiver gather in the synthetic Sigbees model. The salt reflection is then used to derive an envelope region to be down-weighted. The resulting weighting function is shown in Figure 4b. As shown in Figure 4b the down-weighted region (blue) corresponds mostly to the salt reflector.

### Noise removal in the extended domain

Unwanted noise sometimes appears in the migrated images. For example, when internal multiples are not properly removed, migrating that energy with an operator that only accounts for the kinematics of the primary would result in crosstalk noise in the image. In ocean bottom data sets, imperfect PZ summation or up-down decomposition can also result in artifacts in the migrated image. Another source of noise comes from our attempt to invert elastic data with acoustic waves theory. For example, converted wave are in the data but are not accounted for in the modeling and migration operator. Ideally, we want all unaccounted events to be removed in the field data before migration or least-squares migration. LSRTM can be hampered when noise energy remains in the data.

To make the algorithm more robust, we utilize the subsurface offset domain to filter some of the unwanted noise in the image space. Figure 5 shows an offset-domain common image gather (ODCIG) for one of the ocean bottom nodes. The front face corresponds to the image at zero subsurface offset and the side face corresponds to the depth-offset panel at a horizontal location. Notice how the energy in the depth-offset panel is tilted based on its relative position from the source. In Figure 5a, the horizontal position is less than  $x_{\text{shot}}$ . In this case, the prominent energy in the depth-offset panel is tilted with a negative slope. Similarly, in Figure 5c, the horizontal position is greater than  $x_{\text{shot}}$ . In this case, the prominent energy in the depth-offset panel is tilted with a positive slope. When we examine the depth-offset gather at a horizontal position that is close to the source, as shown in Figure 5b, the prominent energy has near zero slope.

This characteristic can be used to discriminate some of the noise in the image space. A filter was designed and applied for each ocean bottom node's ODCIG before summing together. For each ODCIG, we apply the Fourier transform of each depth-offset slice into the  $k_z - k_{\text{hx}}$  wavenumber space. In the wavenumber space, a dominant dip energy range can be selected as signal while the other energy can be filtered out as noise. The dip-range can be estimated based on the horizontal  $x$  position of the  $k_z - k_{\text{hx}}$  slice relative to the shot position  $x_{\text{shot}}$ . We used a large dip range (60 degrees) to avoid filtering out true reflector signal. Seismic waves are often severely distorted and bent by the salt, which could result in true signal forming at various dip in the  $z - h$  panel. We have used a different set of filtering criteria for the region beneath the salt so that the true signal will not be filtered out.

Physically, filtering the dips in the depth-offset domain is equivalent to filtering in angle domain. Recall the relationship that links the dips in the depth-offset domain to the aperture angle ( $\gamma$ ):

$$\frac{k_{\text{hx}}}{k_z} = -\tan \gamma. \quad (9)$$

We are making the assumption that, based on the horizontal distance between the image point with  $x_{\text{shot}}$ , signal can only be formed with a particular aperture angle range.

Although equation 9 is only true in 2D, an equivalent expression in 3D that include the reflector's tilt exist. For this particular dataset, the dipping along the crossline direction is minimal. An equivalent filtering procedure involving  $k_{\text{hy}}$  can be applied in 3D when the crossline dip is significant.

Figure 6 shows the result of extended domain filtering on a single prestack subsurface offset image gather. Most of the noise is removed above the salt reflection at  $z = 4000m$ . Figure 7 shows an enlarged section of Figure 6. The original prestack RTM image (Figure 7a) is decomposed into the signal part (Figure 7b) and the noise part (Figure 7c) by using extended-domain noise filtering. In our LSRTM algorithm, the extended-domain filtering is used to remove some of the noise in the gradient. A similar approach was used by Valenciano (2008b). Instead of filtering in the  $k_z - k_{\text{hx}}$  space, Valenciano (2008b) filter out salt-related internal multiples in the  $kx - k_{\text{hx}}$  wavenumber space.

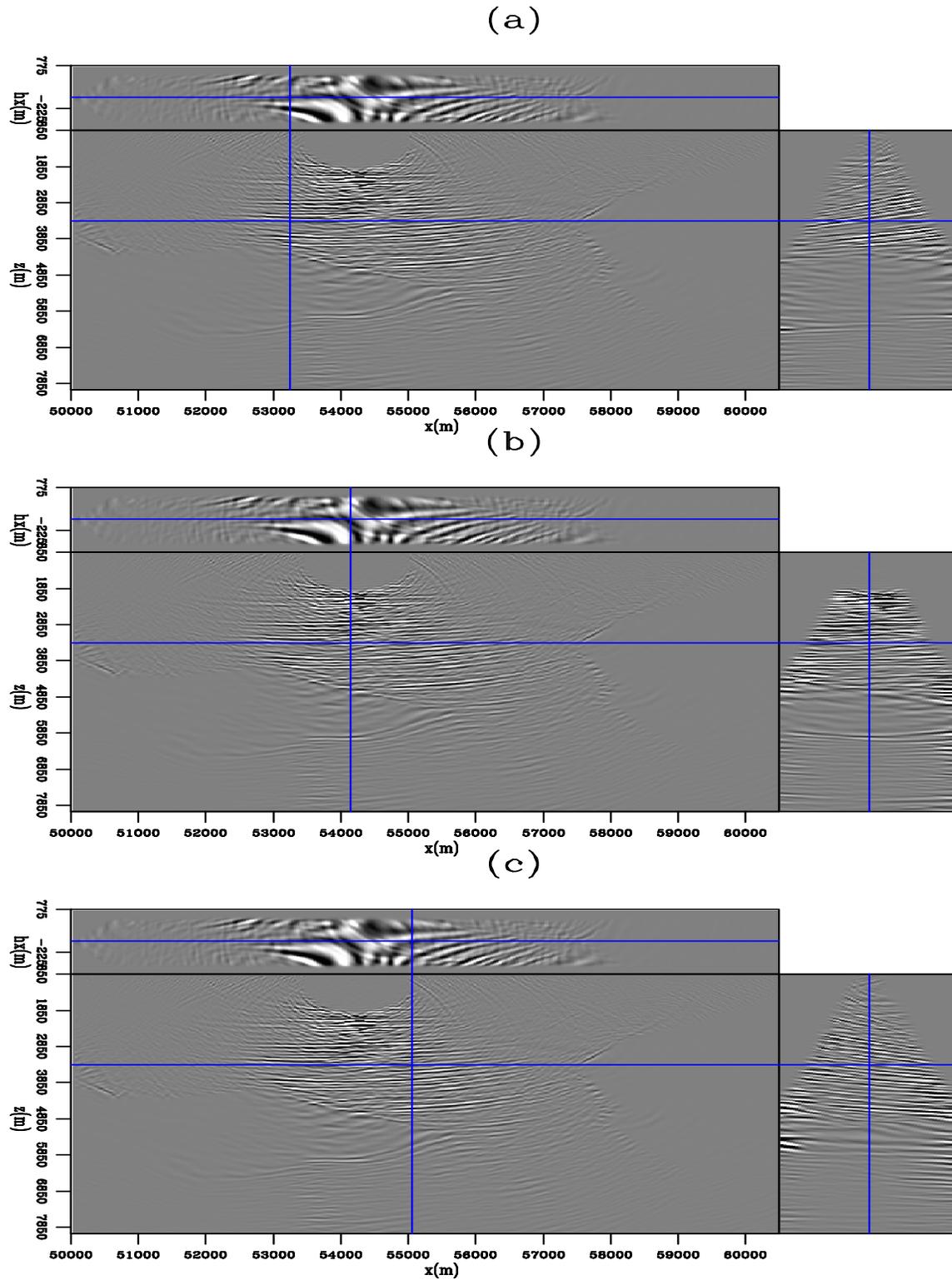


Figure 5: A prestack subsurface offset image gather with a single node located at  $x=54350$  m and  $y=34800$  m. The image cube is displayed at a constant crossline slice of  $y=54350$  m and with the horizontal slice displayed (a) to the left, (b) at the center, and (c) to the right of the horizontal shot location. Notice how the energy in the depth-offset panel is tilted based on its relative position from the source. [CR] `mandy1/.dmDei5`

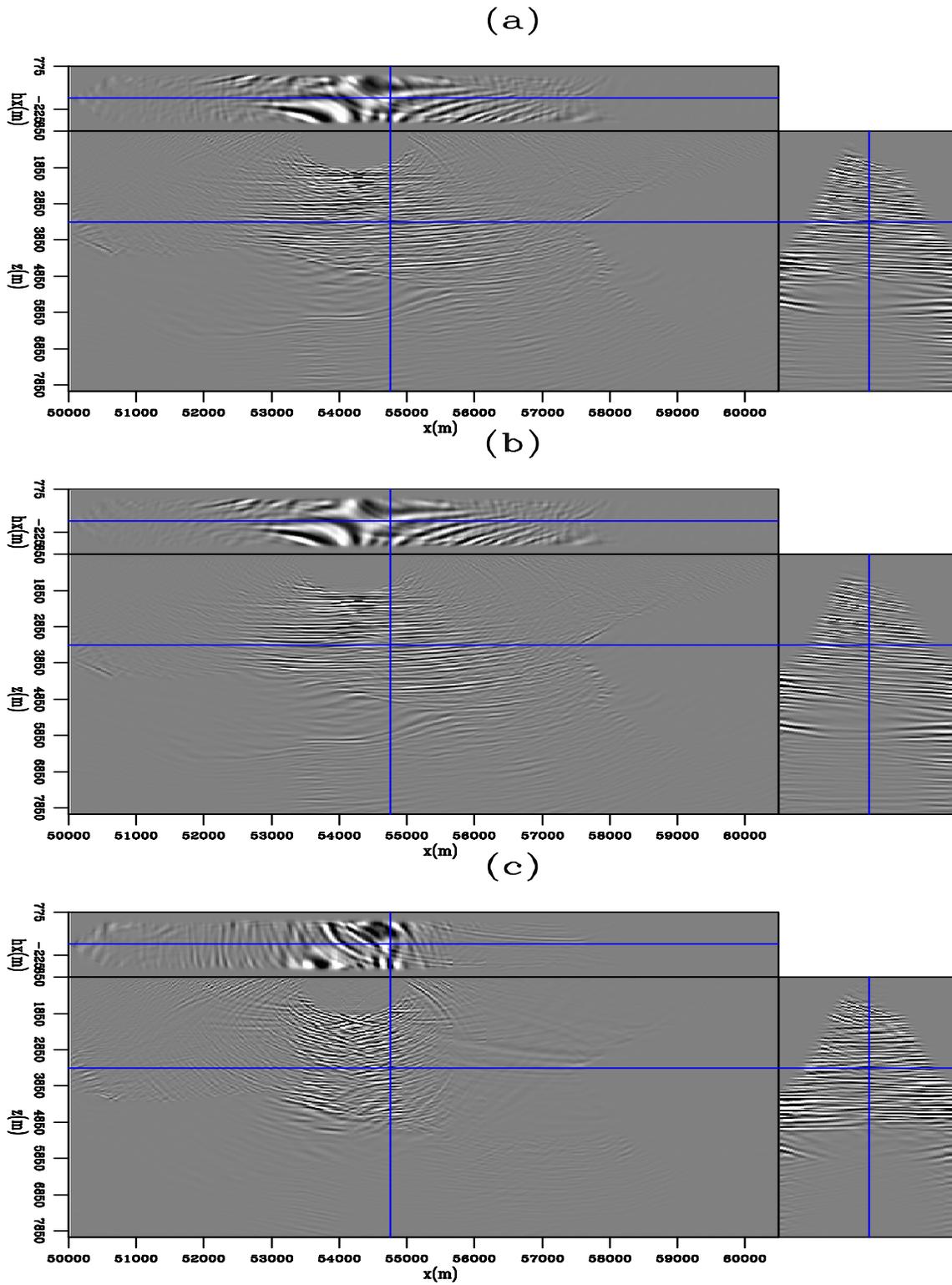


Figure 6: The same prestack subsurface offset image gather from Figure 5 with the horizontal slice displayed at  $x=54750$  m (a) before filtering (b) after filtering, and (c) filtered noise. The sum of (b) and (c) should be the same as (a). [CR] `mandy1/. filtdmDei5`

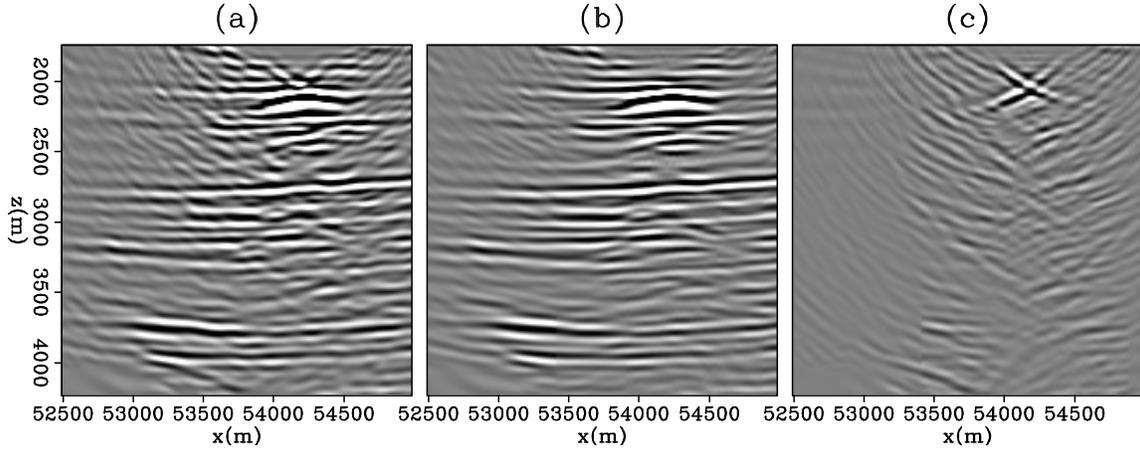


Figure 7: An enlarged section of Figure 6 (a) before filtering, (b) after filtering, and (c) filtered noise. The sum of (b) and (c) should be the same as (a). [CR] `mandy1/. zoomfiltmDei5`

### Regularization that penalizes discontinuity along the reflection angle

We used the regularization term in Clapp (2005) to penalize discontinuity along the subsurface reflection angle axis. This can be done by multiplying the factor  $h_x$  onto the ODCIG. We represent the regularization operator with  $\mathbf{D}$  as shown,

$$\epsilon \mathbf{D} \mathbf{p} = \epsilon h_x p(x, y, z, h_x). \quad (10)$$

In LSRTM, it can be added as a regularization term to help focus the subsurface-offset energy to zero-offset. The value of  $\epsilon$  is chosen such that the model fitting term is roughly about 5 percents of the data fitting term at the first iteration. As the data-fitting residual decreases over iterations, the relative contribution from the regularization increases. The final objective function looks like:

$$S(\mathbf{p}) = \|\mathbf{W}_s^\uparrow(\mathbf{L}_\uparrow \mathbf{A} \mathbf{p} - \mathbf{d}_{\text{obs}}^\uparrow)\|^2 + \|\mathbf{W}_s^\downarrow(\mathbf{L}_\downarrow \mathbf{A} \mathbf{p} - \mathbf{d}_{\text{obs}}^\downarrow)\|^2 + \|\epsilon \mathbf{D} \mathbf{p}\|^2, \quad (11)$$

where  $\mathbf{D}$  represents the DSO operator. Notice that the DSO is acting on the preconditioned variable  $\mathbf{p}$  instead of the model  $\mathbf{m}$ . By dampening on the preconditioned variable  $\mathbf{p}$ , the regularization term effectively penalizes large values in  $\|\epsilon \mathbf{D} \mathbf{A}^{-1} \mathbf{m}\|^2$ . This effectively penalizes discontinuity at smaller reflection angle more than at larger reflection angle.

## RESULTS

### Deimos 3D ocean bottom node dataset

The Deimos data set was recorded in an area approximately south-east of New Orleans in the Gulf of Mexico. The field was discovered in 2002, with first oil production in 2007 (Burch et al., 2010; Smit et al., 2008; Stopin et al., 2008). In 2007, Shell Exploration and

Production Company and their partner BP Americas commissioned Fairfield Industries to conduct a 3D ocean bottom node survey over the Deimos field in the Mississippi Canyon protraction area. The survey was acquired with a 400 x 400 m grid of 807 ocean bottom nodes with a 50 x 50 m shot grid. Before the 3D survey began, a pilot mini-survey of 16 ocean bottom nodes were deployed on a single 2D line at their normal 3D grid locations (Hays et al., 2008). A swath of seven dual-source sail lines nearest the node line were shot. This gives fourteen source lines on the nominal 50x50 m grid for about 3300 shots. Figure 8 shows the location of the sources and the receivers.

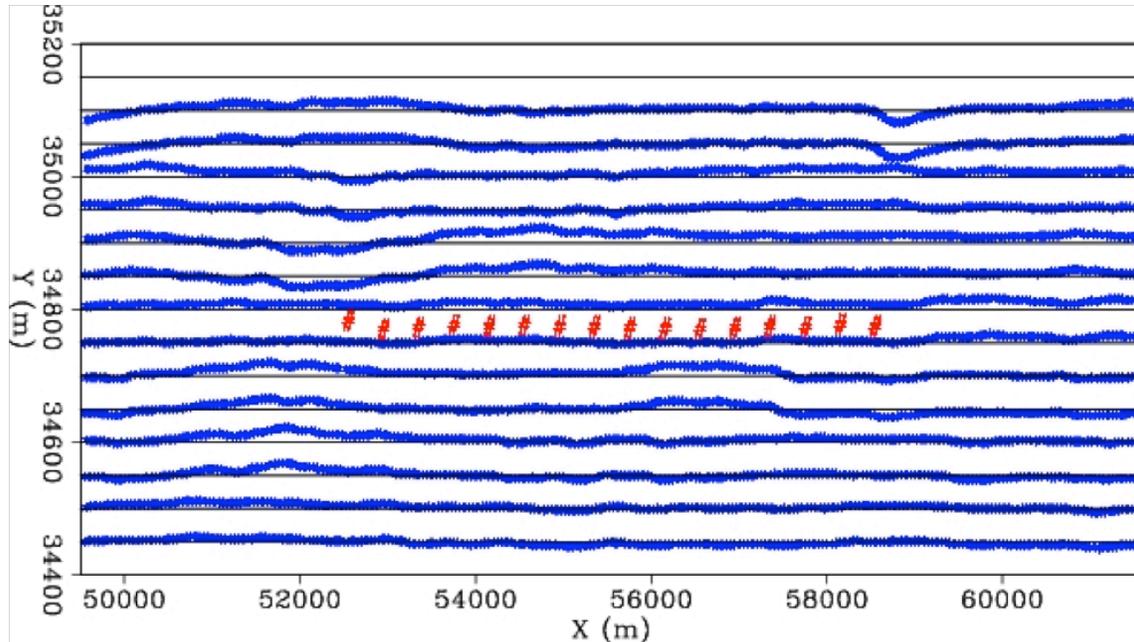


Figure 8: The acquisition geometry for the Deimos ocean bottom data set. The 14 source lines span a 50 x 50 m grid. The 16 ocean bottom node receivers are deployed on a 2D line with an approximate spacing of 400 m. [NR] `mandy1/. sourcedeimos`

This is effectively a narrow azimuth survey where the horizontal (in-line) extent is much greater than the vertical (crossline) extent. As a result, inline dipping reflectors should be much better resolved than crossline dipping reflectors. Prior knowledge of this field suggests that there are fewer structural variations along the crossline direction compared to the inline direction. Figure 9 shows the migration velocity model used in this study.

## Pre-processing

Shell Exploration and Production Company performed some pre-processing of this ocean bottom node dataset. PZ summation and multiple removal are performed to extract the up- and down-going signal from the dataset. Figure 10 shows the up-going and down-going data after pre-processing.

The conventional imaging scheme for ocean bottom node datasets migrate only the down-going mirror (Ronen et al., 2005; Dash et al., 2009) signal. Often time, more care is given to the down-going mirror signal than the up-going signal. We observe that the

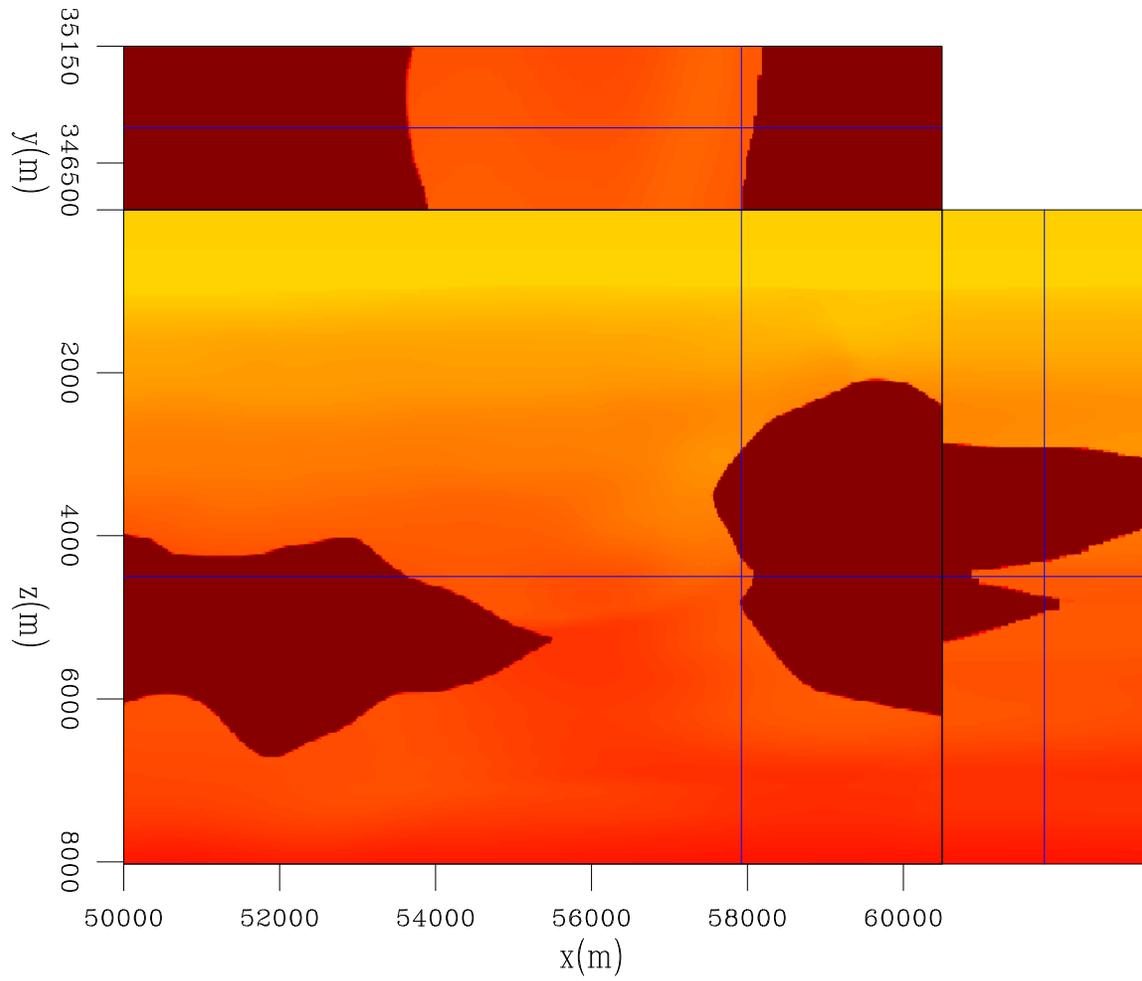


Figure 9: The migration velocity model used for the 16 nodes of the Deimos ocean bottom node survey. [CR] `mandy1/. veldeimos`

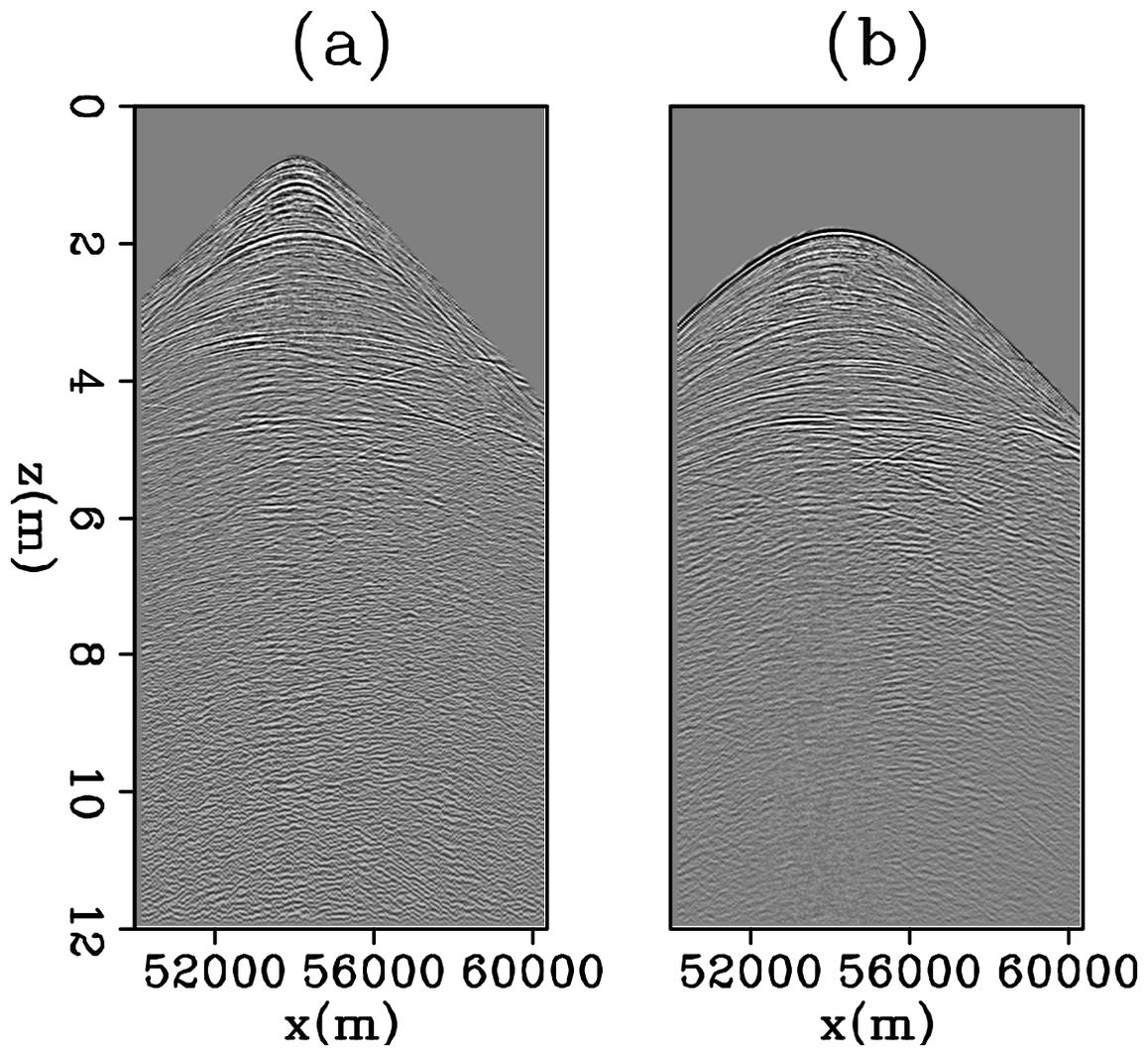


Figure 10: One common receiver gather of the (a) up-going primary and (b) down-going mirror data. [CR] `mandy1/. datadeimos`

down-going data is better separated than the up-going data. Figure 11a shows the imaging result when we apply the down-going migration operator onto the up-going data. We can represent this mathematically with:

$$\mathbf{m}_{walk} = \mathbf{L}_{\downarrow}^T \mathbf{d}_{\uparrow}. \quad (12)$$

Figure 11b shows the image from applying the down-going migration operator onto the down-going data. We can identify some of the events that are present in either of the two images. For example the reflectors annotated in Figure reffig:xtalkdeimos can be found in both images. This suggests that some residual down-going energy remains in the up-going data.

When unaccounted energy is present in the field dataset, it compromises both the RTM and the LSRTM results. In the LSRTM case, the inversion will adjust the solution to try to explain some of the crosstalk energy in the data. To alleviate this problem, I have applied the extended domain noise filtering scheme to remove some of the unwanted energy in the image space.

### Imaging with a single mode

In conventional imaging, the down-going mirror signal is used for migration because it provides a wider illumination area. Often time, the up-going signal is not used in imaging. This is because the illumination area of the primary reflection is often much narrower than that from the mirror reflection. Figure 12 shows the relative illumination area (highlighted in yellow) of the primary and mirror events. The mirror signal can clearly illuminate the sea-bottom region much better than the primary signal.

Figure 13 shows the RTM image using the primary signal and the mirror-only signal. Notice that the illumination area between Figure 13a and b is not as dramatically different as described by Figure 12. This is because the lateral extend of the source grid is only slightly larger than the lateral extend of the receiver grid along the inline direction. Although the illumination between the up-going and down-going signals are similar for this survey, joint-imaging could still be beneficial in terms of higher signal to noise ratio for the image. I will show you some of the joint imaging results in the next section.

### Joint imaging of up- and down-going ocean bottom signals with LSRTM

Figure 14a shows the joint-RTM image and Figure 14b shows the joint-LSRTM image. The joint-RTM image is essentially the first gradient of the joint-LSRTM algorithm. It can also be viewed as the sum of the up-going RTM image and the down-going RTM image. As compared to conventional mirror image (Figure 13b) and the joint-RTM image (Figure 14a), the joint-LSRTM image has better relative amplitude balance for the deeper reflectors.

One challenging area to image in the Deimos field is up against the Antares salt. Burch et al. (2010) proposed using 3D VSP survey to improve illumination in that region (Figure 15). We will look at how joint-LSRTM performs in a nearby region as depicted by Figure 15.

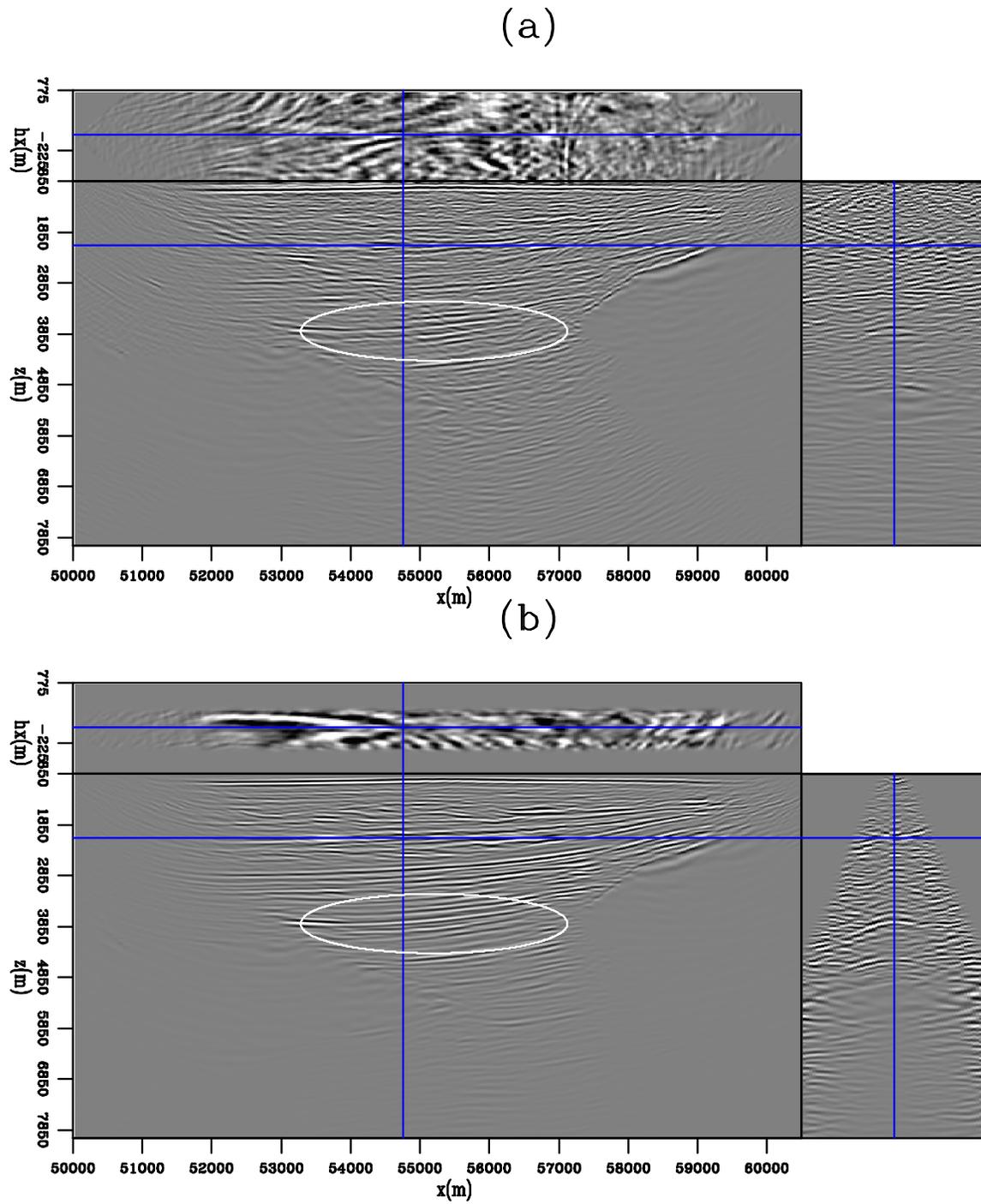


Figure 11: Image depicting crosstalk energy when applying the down-going migration operator onto the up-going data;  $\mathbf{m}_{xalk} = \mathbf{L}_{\downarrow}^T \mathbf{d}_{\uparrow}$ . [CR] `mandy1/. xtalkdeimos`

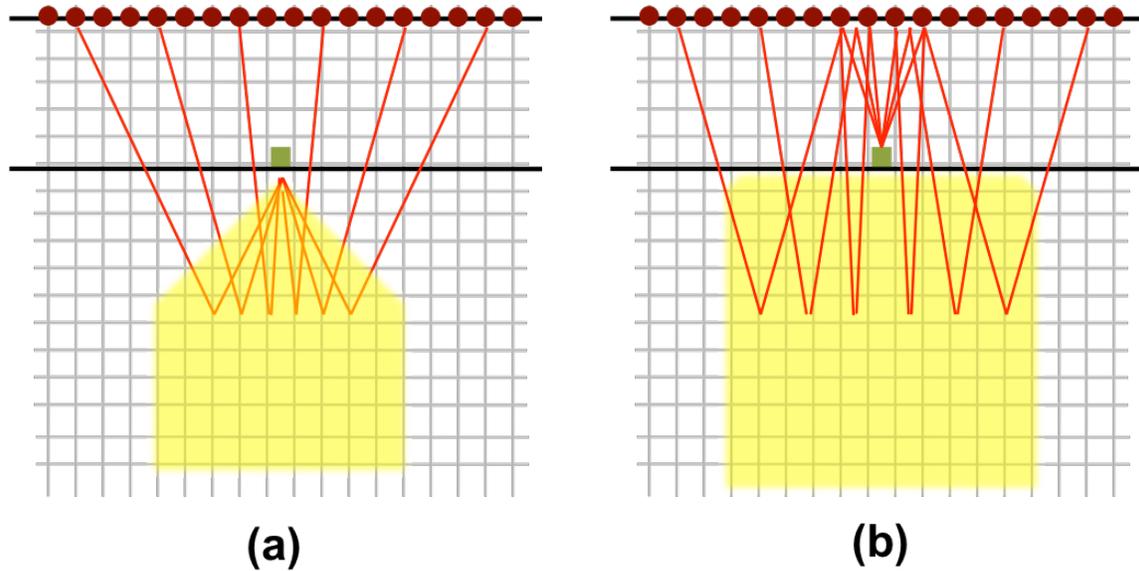


Figure 12: The relative illumination area (highlighted in yellow) of the (a) primary and (b) mirror events for ocean bottom node geometry. [NR] `mandy1/. illumupdown`

Figure 16 shows an enlarged section near the Antares salt for the down-going RTM and down-going LSRTM images. Figure 17 shows the corresponding enlarged sections of the joint-RTM and joint-LSRTM images. Comparing the two figures, we can see that the joint-LSRTM has improved illumination near the salt flank. The side-panel in Figure 17b shows that the energy is more focused at zero-subsurface offset for reflectors close to the salt flank.

## DISCUSSION

The result from Figure 17 is encouraging. However, more investigation is needed to study the contribution from the up-going and the down-going signal onto the joint-image. There are several future directions that we would like to pursue. One is to apply target oriented data space weighting to emphasize the subsalt region. Another direction is to use iterative reweighted least-squares for the inversion.

## CONCLUSION

Least-squares reverse time migration is an advanced imaging technique that can improve imaging with better relative amplitude information, fewer artifacts, and reduced noise. When applying to field data sets, the recorded data departs from the theory and assumptions of the LSRTM operator. It is found that a simple data-fitting objective function in LSRTM is not sufficient. We used Laplacian preconditioning, salt-dimming data weighting, noise filtering in the extended domain, and DSO regularization to condition the LSRTM algorithm when applied to the 3D Deimos ocean bottom field data set. By comparing the joint-LSRTM image of up- and down-going signal with the conventional down-going RTM or down-going

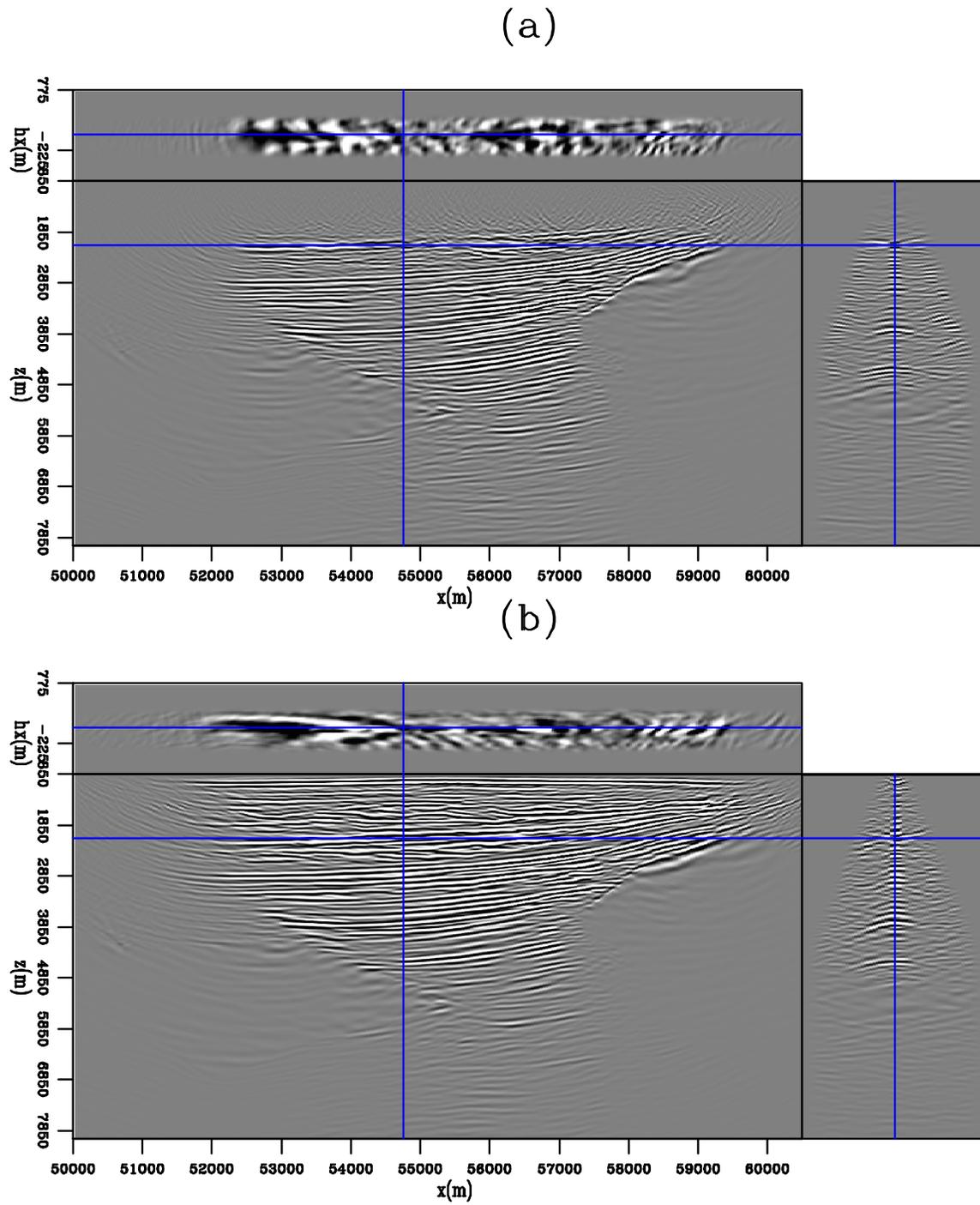


Figure 13: Subsurface offset RTM image gather using (a) the up-going primary signal and (b) the down-going mirror signal. The front face shows the image gather at zero-subsurface offset. [CR] [mandy1/. updownimagedeimos](#)

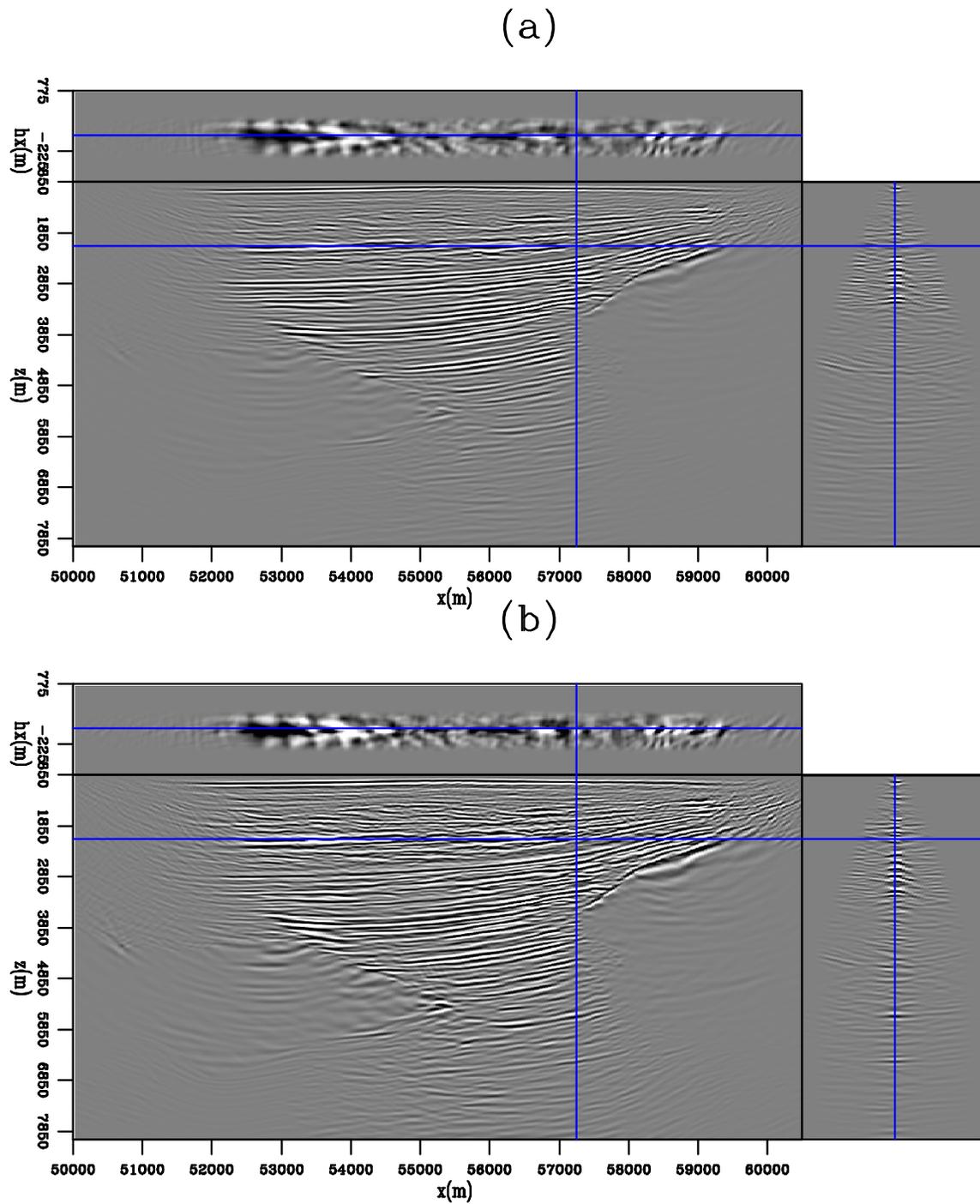


Figure 14: (a) joint-RTM image. (b) joint-LSRTM image at iteration 25. [CR]  
 mandy1/. jointimagedemos

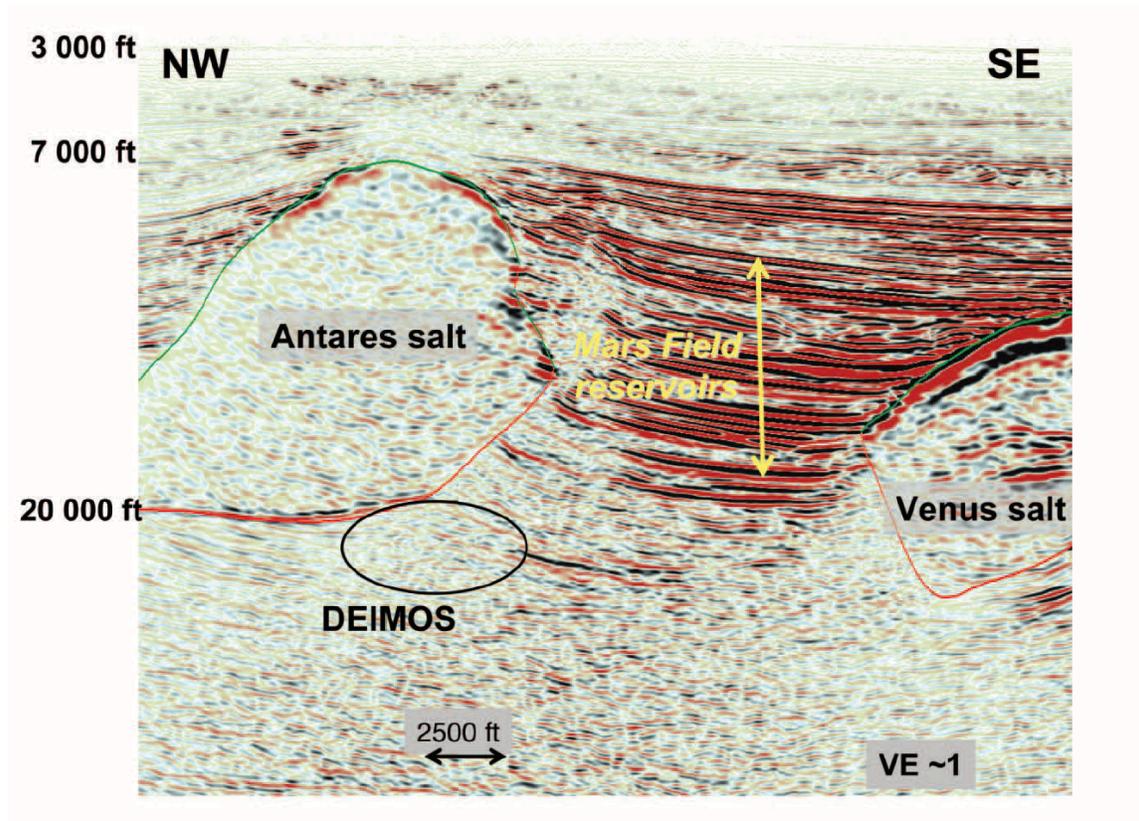


Figure 15: A migration image of the 2007 survey with 807 ocean bottom nodes. The Deimos field is indicated by the oval below the Antares salt. This figure is from Burch et al. (2010).  
[NR] [mandy1/. DeimosPub1image](#)

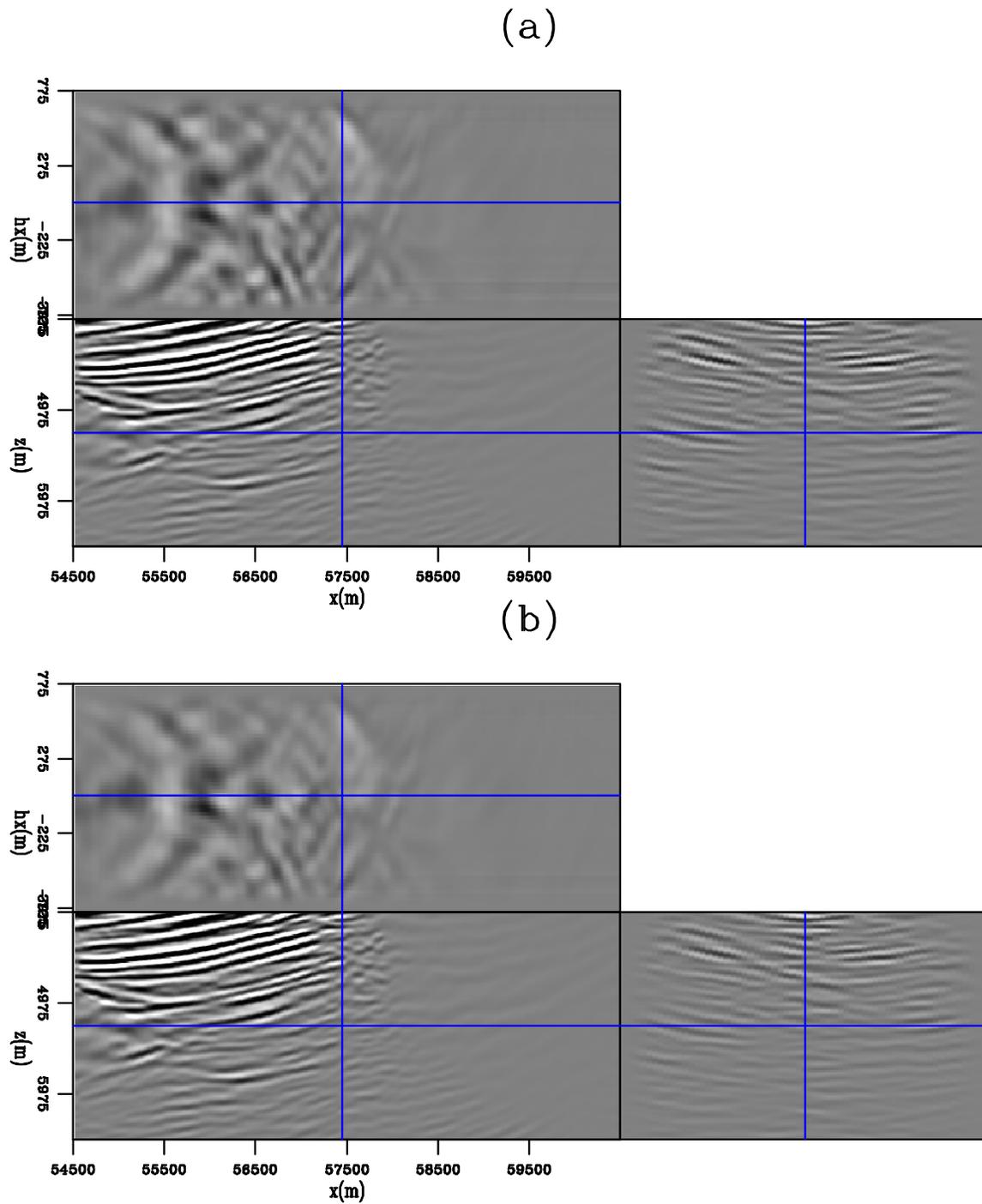


Figure 16: A enlarged section near the Antares salt for (a) the down-going RTM image and (b) the down-going LSRTM image at iteration 25. [CR] `mandy1/. zoomdowndeimos`

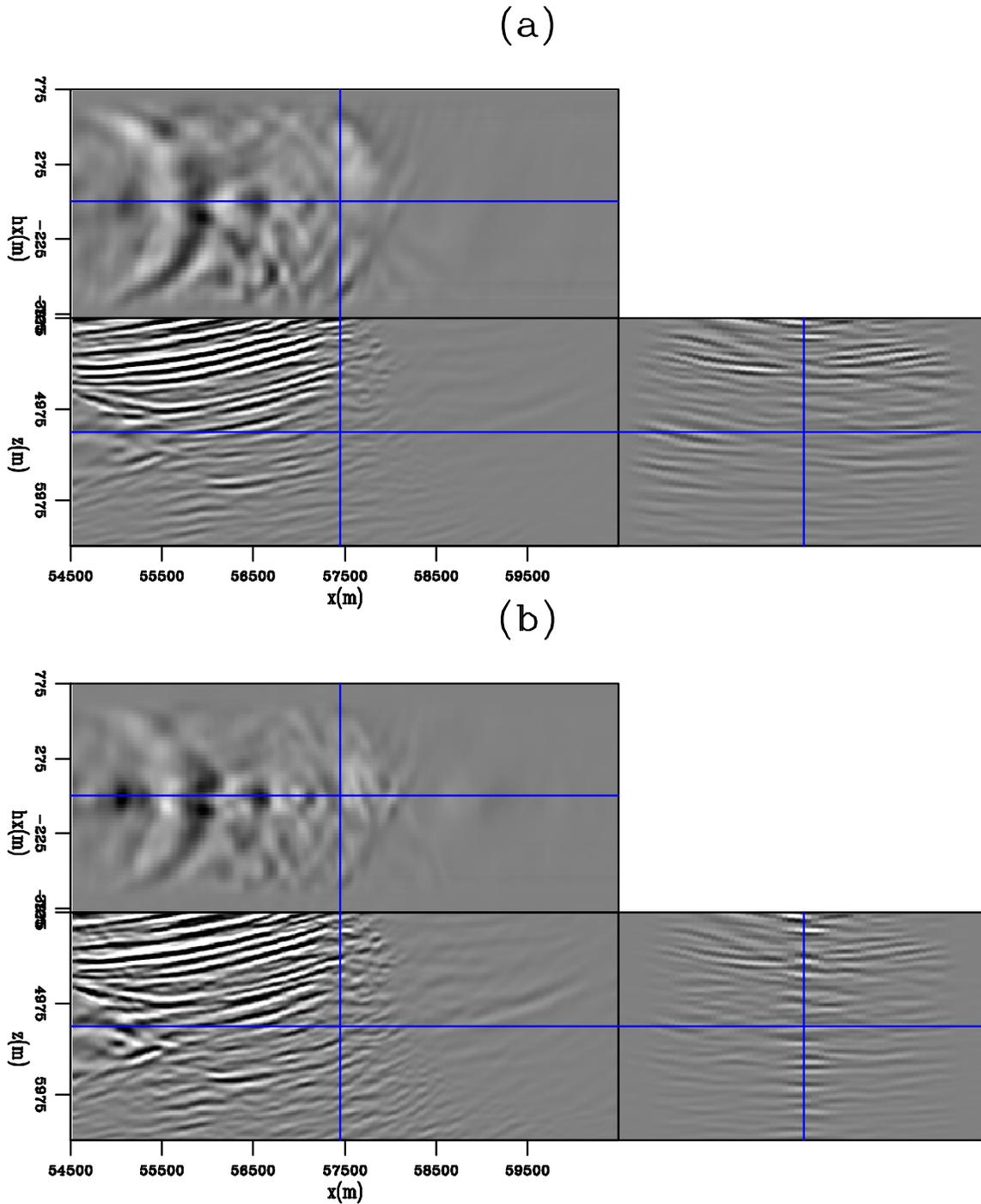


Figure 17: A enlarged section near the Antares salt for (a) the joint-RTM image and (b) the joint-LSRTM image at iteration 25. Notice that the reflectors against the salt flank near the Deimos field are better illuminated. The side-panel in (b) shows that the energy is more focused at zero-subsurface offset. [CR] `mandy1/. zoomjointdeimos`

LSRTM images, we see better relative amplitude balance for the reflectors and improved illumination near the Antares salt area.

## ACKNOWLEDGMENTS

The authors wish to thank Shell Exploration and Production Company, as well as BP Americas for permission to publish this work. Mandy thanks Michael Merritt, Richard Cook, Colin Perkins, Vanessa Goh, and Alexander Stopin for their data processing, helpful suggestions and discussions.

## REFERENCES

- Burch, T., B. Hornby, H. Sugianto, and B. Nolte, 2010, Subsalt 3d vsp imaging at deimos field in the deepwater gulf of mexico: The Leading Edge, **29**, 680–685.
- Clapp, M. L., 2005, Imaging under salt: illumination compensation by regularized inversion: PhD thesis, Stanford University.
- Dash, R., G. Spence, R. Hyndman, S. Grion, Y. Wang, and S. Ronen, 2009, Wide-area imaging from obs multiples: Geophysics, **74**, Q41–Q47.
- Hays, D., K. Craft, P. Docherty, and F. Smit, 2008, An ocean bottom seismic node repeatability study: SEG Technical Program Expanded Abstracts, 55–59.
- Lambare, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.
- Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.
- Ronen, S., L. Comeaux, and J. Miao, 2005, Imaging downgoing waves from ocean bottom stations: 75th SEG Annual Meeting Expanded Abstracts, 963–966.
- Ronen, S. and C. Liner, 2000, Least-squares dmo and migration: Geophysics, 1364–1371.
- Smit, F., C. Perkins, L. Lepre, K. Craft, and R. Woodard, 2008, Seismic data acquisition using ocean bottom seismic nodes at the deimos field, gulf of mexico: SEG Technical Program Expanded Abstracts, 998–1002.
- Stopin, A., M. M. Rae, L. Lepre, and B. Gaudin, 2008, Constructing an anisotropic velocity model for ocean bottom seismic node data: SEG Technical Program Expanded Abstracts, 993–997.
- Valenciano, A., 2008a, Imaging by Wave-Equation Inversion: PhD thesis, Stanford University.
- Valenciano, A. A., 2008b, Imaging by wave-equation inversion: PhD thesis, Stanford University.
- Wong, M., 2013, Handling salt reflection in least-squares rtm: SEG Expanded Abstracts, 3921–3925.
- Wong, M., B. Biondi, and S. Ronen, 2010, Joint least-squares inversion of up- and down-going signal for ocean bottom data sets: SEG Expanded Abstracts, 2752–2756.
- Zhang, Y. and J. Sun, 2009, Practical issues in reverse time migration: true amplitude gathers, noise removal and harmonic source encoding: First Break, **27**, 53–59.

# Designing an object-oriented library for large scale iterative inversion

*Chris Leader and Robert Clapp*

## ABSTRACT

A flexible library that allows the user to apply a variety of geophysical imaging/inversion techniques and leverage a selection of solvers can be a very powerful tool. However, constructing this to work in multiple dimensions and with a variety of options is a difficult task. The abstraction provided by object-oriented languages helps us to separate the geophysics from the solver, to use the same function calls for models of different dimensions and to create a single framework that has the potential to apply a range of imaging or inversion methods on heterogeneous computing systems.

## INTRODUCTION

Geophysical inversion acts to ‘un-do’ (or ‘invert’) certain physics of the Earth in order to estimate Earth properties. A data-set is provided that measures perturbations which relate to subsurface properties or interactions - these could be well-log data, GPS data, tilt data, seismic data, etc. Typically one can then forward model or simulate these physics over an estimated model and attempt to estimate our provided data set. The closer we get, the better our model and physics approximations. An inverse process endeavours to improve this model such that our recorded data are better represented. Given data sets can often be tera-bytes in size, and the appropriate model sampling can result in models that are giga-bytes or tera-bytes in size. Thus matrix-based processes become impractical and we must turn to iterative methods (Jupp and Vozoff, 1975).

Using object-oriented programming, we can separate the physics from the solver. This can be done using C++ (Nichols et al., 1993), Fortran 2003 (Clapp, 2010), java (Schwab, 1998) or Python (Clapp, 2005), amongst others. Furthermore, any contemporary iterative imaging and inversion methods have the same basic mathematical functions at their core - these are two-way wave propagation, correlation (imaging and extended imaging) and the vector operations (essentially a series of dot products) that comprise our solvers. A single library that contains these operations (and other necessary, less computationally intensive functions) can be constructed to permit minimal additional coding or effort needed to perform, say, Reverse Time Migration (RTM, Baysal et al. (1983)), linear inversion (occasionally called LSRTM, Nemeth et al. (1999)) or Full Waveform Inversion (FWI, Tarantola (1984)). This library can also be designed to accommodate modern distributed networks and heterogeneous computing, including GPUs (Foltinek et al., 2009). This is especially desirable for a group such as SEP, who possess several heterogeneous computing systems and perform many techniques that contain wave propagation and wavefield correlation at their core.

There has been an ongoing effort at SEP to combine our common interests into one

framework and the approach discussed herein has used C++ and an appropriate class hierarchy. This paper will discuss aspects of this class-based library, the benefits of this construction, and where it is headed in the future.

## OBJECT ORIENTED INVERSION

There are several methods to solve geophysical inverse problems. The majority fall under two categories - statistical methods (Duijndam, 1988) or gradient methods (Claerbout, 1999). Statistical/Bayesian methods can be useful when combining multiple sorts of data or for problems that typically contain very rough objective surfaces. A major advantage is that they allow us to include a priori information about our model. In gradient based methods we aim to find the model  $\mathbf{m}$  that minimizes the difference between our estimated data,  $\mathbf{d}_{est}$ , and our observed data,  $\mathbf{d}_{obs}$ . This amounts to trying to minimise our residual vector,  $\mathbf{r}$ , defined as

$$\mathbf{L}\mathbf{m} - \mathbf{d}_{obs} = \mathbf{r} \approx \mathbf{0} \quad (1)$$

where  $\mathbf{L}$  is a linear operator. In the simplest case we can use  $\mathbf{L}$  and its adjoint  $\mathbf{L}'$  in a solver; this will be discussed in more detail in the next section.

Our object oriented library seeks to solve a variety of problems of this nature.  $\mathbf{L}$  does not necessarily have to be linear in this inversion scheme. Provided we use a non-linear solver (and typically multiple forward modeling processes), this same framework can be used to solve non-linear geophysical inverse problems.

By defining our vectors ( $\mathbf{m}$  and  $\mathbf{d}$ ) and operators ( $\mathbf{L}$ , which could be a cascade of, say,  $\mathbf{A}$  and  $\mathbf{B}$ ) in an object-oriented fashion, we can easily separate the physics,  $\mathbf{L}$ , from the update scheme. Such a method also keeps the coding simple and flexible. Furthermore, we can use a basic set-up and framework for applying many different geophysical techniques (RTM, LSRTM, FWI, etc). Especially since RTM is simply the first gradient of LSRTM, and LSRTM is essentially the inner-loop in FWI.

## SOLVERS

The simplest solver is steepest descent. More complex solvers are almost all built on this basic approach, so we can use this as a case study for our library. Assuming our initial residual,  $\mathbf{r}$ , is simply  $-\mathbf{d}$  (zero starting model), we apply the adjoint of our operator  $\mathbf{L}$  to obtain our gradient vector  $\mathbf{g}$ . Next we map this gradient back into the data-space by applying  $\mathbf{L}$ , obtaining  $\mathbf{r}\mathbf{r}$ . We then perform a series of dot-products to estimate a scaling factor,  $\alpha$ , that makes  $\mathbf{r} + \mathbf{r}\mathbf{r}$  as small as possible. This process is repeated until  $\mathbf{r}$  is sufficiently small or we run out of time. This can be summarised as

where ‘stepper’ is our chosen solver.

As the pseudo-code suggests, the solver acts to update the model and the residual by using  $\mathbf{g}$  and  $\mathbf{r}\mathbf{r}$ . These are each scaled by  $\alpha$  and then summed to  $\mathbf{m}$  and  $\mathbf{r}$ , respectively. This solver does not need to know  $\mathbf{L}$ , or any detailed properties about  $\mathbf{m}$  or  $\mathbf{d}$ . In the case

---

**Algorithm 1** General linearised inversion

---

```

r = Fm0 - dobs
while iter < niter; iter++ do
  g = F'r
  rr = Fgg
  (m, r) = stepper(m, r, g, rr)
end while
Output m

```

---



---

**Algorithm 2** Steepest descent (a possible 'stepper')

---

```

α = -dotproduct(g, rr)/dotproduct(g, g)
m = m + αg
r = r + αrr

```

---

of updating the residual it only needs to calculate a dot-product, to scale **rr**, and to add this vector to **r**.

More complicated solvers are generally comprised of these basic operations; they often just remember details of previous scaling factors and gradients. For example, the code for conjugate directions is only a few lines longer and uses the same three functions.

## IMPLEMENTATION IN C++

Using C++ we can construct a hierarchy of classes. Each class has a series of functions and variables that can be private or shared, and by making some functions abstract we can name a series of methods that others classes can use and define. This paper will not go into detail about all the classes and the hierarchy of the library, but it will mention the vector and operator classes since these comprise our solver.

We know that for our linear solver we need three operations - dot-product, scale and add. These are all functions that can be placed within our vector class, and by keeping a level of abstraction we can use these same operations and function calls to act on arrays of different dimensions. Other operations we can define in 'vector' are 'clone,' 'zero,' and 'set.' These six operations allow us to do most of the necessary operations on the vectors within our inversion. It is wise to also include some vector 'tests', such as checking whether two vectors being operated on belong to the same space.

The operator class is constructed in a similar manner, although it is a little simpler. In this case we define a forward and an adjoint, and then allow two functions that set the 'domain' of the operator and the 'range'. These are essentially the dimensions of the model and the data vectors, respectively. Calling an operator has three inputs and a fourth optional input. Inputs are 'add' (whether we want to zero the output array or not), then the vector 'model' and the vector 'data.' The optional input is a iteration number argument, as some techniques require the current iteration number.

Setting up our classes as such results in simple, readable code for the inversion. The exact code for the solver looks as follows:

```

my_vector *r=data->clone_vec();
r->scale(-1.);
for(iter=0; iter < niter; iter++){
    oper->adjoint(false,g,r,iter);
    oper->forward(false,g,rr,iter);
    bool valid=st->steepest_descent(iter,m,r,g,rr,&val);
}

alpha=-r->dot(rr)/rr->dot(rr);
s->scale(g,alpha);
ss->scale(rr,alpha);
m->add(s);
r->scale(ss);

```

this keeps things simple and readable, and the actual code looks almost the same as the pseudo-code discussed in the previous section. Note that the vectors here could be 2D, 3D or higher (provided they have the same number of dimensions as each other) and the operator could be any forward-adjoint pair. For example, this code snippet could be performing a basic radon inversion, or FWI.

With the operators coded and the various dimensions and data initialized, calling the solver and producing a result is as simple as:

```

step *st=new cgstep();
solver *solv=new solver(st,&data,&operator);
solv->solve(no_iter);
oc_float *result=(oc_float*) solv->return_model();
image.add(result);

```

## HETEROGENEOUS COMPUTING

We must also design these codes to work on a selection of computers, in particular a single core node, a multi-core node, a multi-node network, single GPUs, multiple GPUs on a single node, multiple GPUs across nodes, and out-of-core. It is possible to keep a lot of details about the computing out of the majority of the operator and solver code, however the distributed network code involves a lot more work with MPI.

For single node solutions (single or multi-core, single or multiple GPU) it is only the propagation and the correlation that must be distributed across units. The solver code does not have to change. Furthermore, the details of the decomposition can be kept separate within these aforementioned routines. Provided the code is told which method is desired, the majority of the operator does not have to change either.

Currently the library works with single and multiple GPUs on a given node. This provides an incredible amount of computation power with few of the classic GPU memory restrictions. Whether the user wants to use a single or multiple devices is simply a command line argument. For multiple GPUs domain decomposition is performed (Micikevicius, 2012); however all details of this are kept out of the operator and solver codes. Only one code,

`gpu_funcs_3d.cu` has any knowledge of how this is done, making the majority of the library very clean and readable.

Efficient CPU based multi-core and multi-node implementations are currently being improved upon as part of a SEP initiative. Introducing them to the system is very simple; they can either be included in `gpu_funcs_3d.cu` or a separate file, then whichever method is used can be determined by an input argument. Keeping these numerically intensive aspects somewhat separate from the operators further helps to make the code flexible.

## PHASE ENCODING

Extending the code to use phase encoded inversion (Gao et al., 2010; Morton and Ober, 1998) is less straightforward. In phase encoded inversion we combine groups of shots into supershots, and then aim to recover a common model between them. In the most extreme case we can combine all shots into one supershot, so that we only perform two forward model operations and one adjoint operation per iteration. By changing the encoding matrix that we use between iterations, very fast model convergence (as a function of cost) can be seen (Krebs et al., 2009; Romero et al., 2000).

The inversion set-up needs knowledge of two data-sets: the original/separated data-set and our encoded dataset. However our current operator class only gives us the option of passing one data vector. This means we have to introduce two data vectors within our phase encoded operator routine in the forward routine. This is not a difficult task, but it limits the flexibility of using the same code as LSRTM for phase encoded LSRTM without introducing a series of if statements.

## CONCLUSIONS

An object oriented framework creates a flexible, clean and readable inversion library. The class system means our solver can neatly deal with vectors of multiple dimensions while the internal workings of our operator (the geophysics) is kept separate. The details of how the propagation and correlation are performed on the network are kept separate again from the main operator construction. This creates further flexibility for general operator coding and means only a few command line arguments can be used to greatly vary how the computationally intensive aspects of the inversion are dealt with.

## FUTURE WORK

There are many ways of extending this library in the future. Current effort is being put into multi-core and multi-node solutions for propagation and correlation, into ways of accelerating extended imaging, and into non-linear solvers.

## REFERENCES

Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: *Geophysics*, **45**, 1514–1524.

- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project.
- Clapp, R., 2010, Hybrid-norm and fortran 2003: Separating the physics from the solver: SEP, **142**.
- Clapp, R. G., 2005, Inversion and fault tolerant parallelization using Python: SEP-Report, **120**, 41–62.
- Duijndam, A., 1988, Bayesian estimation in seismic inversion. part i: Principles1: Geophysical Prospecting, **36**, 878–898.
- Foltinek, D., D. Eaton, J. Mahovsky, P. Moghaddam, and R. McGarry, 2009, Industry scale reverse time migration on GPU hardware: SEG Technical Program Expanded Abstracts, **28**, 2789–2793.
- Gao, F., A. Atle, and P. Williamson, 2010, Full waveform inversion using deterministic source encoding: SEG Technical Program Expanded Abstracts, **29**, 1013–1017.
- Jupp, D. L. B. and K. Vozoff, 1975, Stable iterative methods for the inversion of geophysical data: Geophysical Journal of the Royal Astronomical Society, **42**, 957–976.
- Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, WCC177–WCC188.
- Micikevicius, P., 2012, Multi-GPU programming: Presentation at GPU Technology Conference 2012.
- Morton, S. A. and C. C. Ober, 1998, Fastershot-record depth migrations using phase encoding: SEG Technical Program Expanded Abstracts, **17**, 1131–1134.
- Nemeth, T., C. Wu, and G. Schuster, 1999, Least squares migration of incomplete reflection data: Geophysics, **64**, 208–221.
- Nichols, D., H. Urdaneta, H. I. Oh, J. Claerbout, L. Laane, M. Karrenbach, and M. Schwab, 1993, Programming geophysics in C++: SEP-Report, **79**, 313–471.
- Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.
- Schwab, M., 1998, Enhancement of discontinuities in seismic 3-D images using a Java estimation library: **99**.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

## Velocity dispersion at Long Beach

*Jason P. Chang*

### ABSTRACT

Previous studies from the dense Long Beach, California, seismic array have shown that both surface waves and P-waves can be recovered by seismic interferometry. In this report, I focus on constraining the apparent velocities of the wave types by performing  $\tau$ -p transforms and generating dispersion images. From  $\tau$ -p transforms, I find that the velocity of P-waves is approximately 2500 m/s. My dispersion analysis reveals that I am recovering the fundamental Rayleigh wave mode, as well as the first-order mode. Both observed Rayleigh-wave modes are dispersive, with the waves travelling faster at lower frequencies than at higher frequencies. I also find that the first-order mode travels at a greater velocity than the fundamental mode over the range of frequencies that I investigated.

### INTRODUCTION

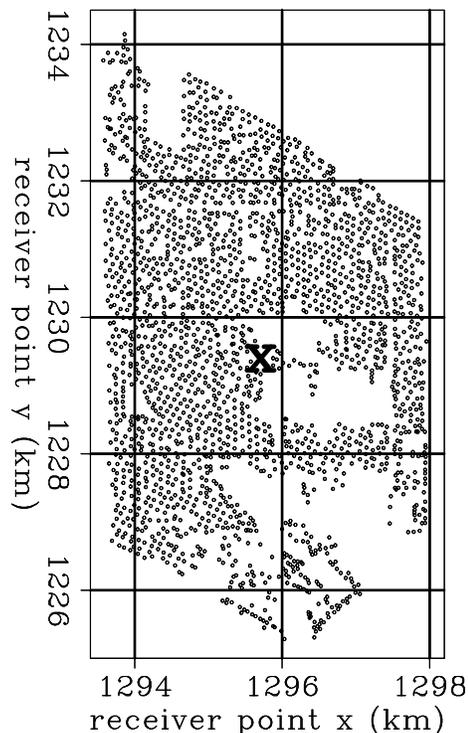
Characterizing near-surface regions ( $< 100$  m depth) is of primary importance in a number of fields, including engineering seismology, where shear wave velocities are useful for earthquake hazard analysis, and hydrogeology, where the depth of the water table can be determined from seismic data. Traditional methods for collecting near-surface seismic information in urban environments are disruptive and difficult to perform. Seismic interferometry provides an alternative method, and the Long Beach, California, seismic array is particularly well-suited for urban seismic interferometry. Deployed by Nodal Seismic, Inc. in January 2012, the array is both dense, consisting of 2400 vertical-component geophones with an average spacing of 100 m (Figure 1), and of long duration, recording continuously (24 hrs/day) for over three months.

From this Long Beach dataset, Chang et al. (2013) showed that Rayleigh waves originating from the Pacific Ocean could be recovered between 0.5 and 2.0 Hz from a week of ambient seismic noise recordings. Furthermore, they showed that by stacking over multiple virtual sources, P-waves could be recovered between 2.0 and 4.0 Hz from nearly three weeks of ambient noise recordings. From this virtual super-source gather, they estimated the P-wave to be propagating at a velocity of approximately 2000 m/s.

The goal of this report is to better constrain the apparent velocity estimates from Chang et al. (2013) for both Rayleigh and P-waves. First, I generate virtual source gathers using seismic interferometry. I then perform  $\tau$ -p transforms and generate dispersion images from those virtual source gathers to better constrain the apparent velocities of the wave types.

Figure 1: Map of the Long Beach array. 'X' marks the location of the virtual source. Coordinates are NAD27, CA State Plane, Zone 7, kilometers. [ER]

`jason1/.lb-midmap`



## VIRTUAL SOURCE GATHERS FROM LONG BEACH

Seismic interferometry states that the Green's function between two receivers can be obtained by cross-correlating records of ambient seismic noise from two simultaneously recording receivers over a long period of time (Wapenaar et al., 2010). In other words, I can generate virtual seismic sources using seismic interferometry. To generate my estimated Green's functions (EGFs), I follow the processing method adapted from Bensen et al. (2007). I first break up the time series into tapered, non-overlapping two-hour time windows. I then whiten the input traces and bandpass filter prior to cross-correlating. In the frequency domain, the cross-correlation of pre-whitened traces can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left( \frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left( \frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle, \quad (1)$$

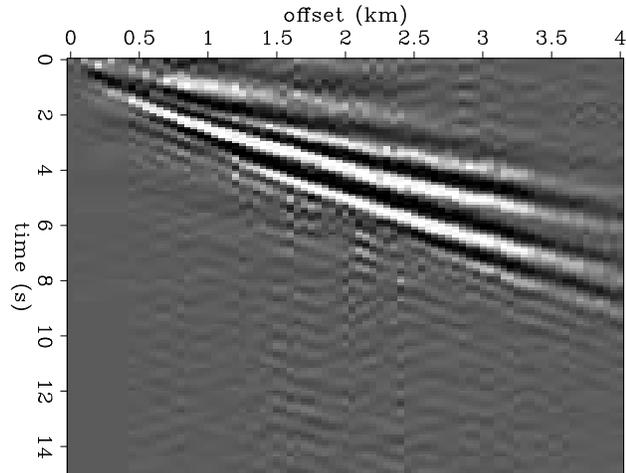
where  $G$  is the Green's function between two receiver locations  $(x_A, x_B)$ ,  $U(x, \omega)$  is the spectrum of the wavefield at a given receiver location  $x$ ,  $*$  represents the complex conjugate,  $\langle \cdot \rangle$  represents the time-averaged ensemble,  $|\cdot|$  represents the real absolute value of the spectrum, and  $\{\cdot\}$  represents a 0.003 Hz running window average used for whitening the signal. By dividing the standard cross-correlation procedure by smoothed amplitude spectrums at the two receivers, I am deconvolving an estimate of the noise source signal from the correlations in the time domain to obtain a better EGF.

Examples of virtual source gathers obtained from seismic interferometry can be seen in Figures 2 and 3. In Figure 2, I show traces that are bandpassed for 0.5-2.0 Hz. The virtual source location is shown in Figure 1. I average the correlations from 100 two-hour time windows, resulting in an average over 8 days. I then sort the traces by radial offset into 50 m bins and stack. Due to the inhomogeneous noise source distribution at Long Beach

(Chang et al., 2013), I choose to improve the quality of the correlations by averaging the positive and negative time lag signals. In other words, I take the symmetric part of the correlations.

Figure 2: A virtual common source gather for frequencies between 0.5 Hz and 2.0 Hz. The virtual source location is labeled by 'X' in Figure 1. Traces have been sorted into 50 m radial offset bins. [CR]

jason1/. sh-ph2-down-Jan-mid-1-si-offaz



In Figure 3, I display a virtual super-source gather for frequencies between 2 and 4 Hz. This is made by creating virtual source gathers from 100 randomly distributed source locations for each of the 240 two-hour time windows. This amounts to nearly 3 weeks of data and over 50 million correlations. I then sort all correlations by radial offset into 50 and 200 m bins, stack, and then symmetrize across zero time lag. Chang et al. (2013) used 200 m bins (Figure 3(a)) because it clearly revealed a P-wave that they estimated to be travelling at 2000 m/s based on the slope of the stepout. Here, I show that the P-wave is still visible with 50 m binning (Figure 3(b)), and that the surface waves are better resolved with finer binning. Hence, I will use the super-source gather with 50 m binning for further analysis.

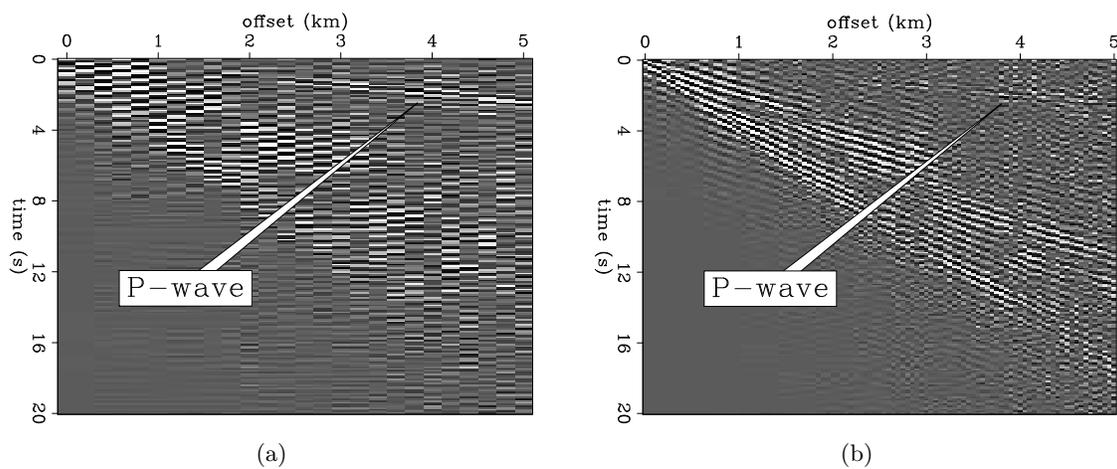


Figure 3: Virtual super-source gathers for frequencies between 2 and 4 Hz. Left: 200 m offset binning. Right: 50 m offset binning. [CR]

jason1/. superdupergather1-coarse-si-offaz-a,superdupergather1-fine-si-offaz-a

## $\tau$ -P TRANSFORMS

A  $\tau$ -p transform, or slant stack, maps the gathers from the radial offset-time domain to the horizontal slowness-intercept time domain, or the  $\tau$ -p domain. The transform is represented as

$$u'(p, \tau) = \int u(x, \tau + px) dx, \quad (2)$$

where  $u$  is the wavefield,  $u'$  is the transformed wavefield,  $x$  is radial offset,  $p$  is slowness or ray parameter ( $dt/dx$ ), and  $\tau$  is the time intercept or zero-offset reflection time (Claerbout, 1985; Stoffa, 1989). The transform extracts slowness information, as any coherent energy travelling at a certain slowness will stack constructively and appear as a strong event in the  $\tau$ -p domain.

Figure 4(a) shows the gather in the  $\tau$ -p domain for the lower frequency range (0.5-2.0 Hz). The anomaly appears quite spread out, meaning the Rayleigh waves do not travel at just one phase velocity. The slownesses are primarily bounded between 0.5 and 2.0 ms/m, or between 500 and 2000 m/s. Figure 4(b) shows the super-source gather in the  $\tau$ -p domain for the higher frequency range (2-4 Hz). Again, there is a large anomaly that is representative of the Rayleigh waves. Here, the slownesses are primarily bounded between 1.4 and 2.7 ms/m, or between 370 and 710 m/s. It makes sense that Rayleigh-wave phase velocities are greater for lower frequencies than for higher frequencies, as lower frequencies typically sense deeper (and hence sense greater velocities) than higher frequencies.

There is also a smaller anomaly in Figure 4(b) near zero intercept time with a slowness of 0.4 ms/m, or 2500 m/s. This event is representative of the P-wave arrival. Whereas Chang et al. (2013) estimated the velocity of this arrival to be approximately 2000 m/s based on measuring the slope of the moveout with 200 m offset binning, the  $\tau$ -p transform reveals that the apparent velocity is faster. This transform also shows that although the P-wave is best revealed when radial offset is sorted into 200 m bins, the velocity is better measured with increasingly smaller bins.

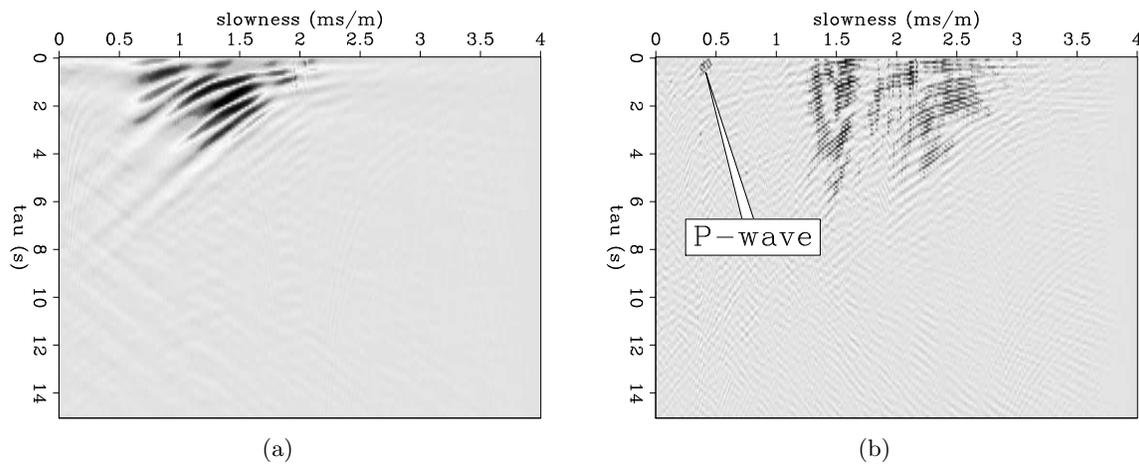


Figure 4:  $\tau$ -p transforms of the virtual source gathers with 50 m binning. Left: using the virtual source gather in Figure 2 for frequencies between 0.5 and 2.0 Hz. Right: using the virtual super-source gather in Figure 3(b) for frequencies between 2 and 4 Hz. [CR]

jason1/. sh-ph2-down-Jan-mid-1-si-tp,superdupergather1-fine-si-tp-a

With a P-wave velocity of 2500 m/s, it is unlikely that this event is a refraction off the water table. However, more correlations will need to be stacked in to improve the quality of the gather at short offsets so that I can determine whether this P-wave is a critically refracted arrival or a direct arrival.

## DISPERSION ANALYSIS

From the  $\tau$ -p transformed gathers, it is clear that Rayleigh waves covered a range of velocities that varied with frequency content. To better see this behavior, I create dispersion images in the frequency-slowness domain. This is done by taking a one-dimensional Fourier transform along the intercept time axis of the  $\tau$ -p transform. The resulting spectral peaks in the dispersion images indicate the phase slowness(es) associated with each frequency.

Figure 5(a) shows the dispersion image for a virtual source gather for the lower frequency range. There are two dispersive wave modes: the fundamental and the first-order Rayleigh wave modes. The stronger fundamental mode travels slower than the weaker first-order mode for a given frequency. For instance, at 0.8 Hz, there is a spectral peak centered near 1.5 ms/m, which corresponds to the fundamental mode travelling at 660 m/s, and a minor spectral peak near 0.9 ms/m, which corresponds to the first-order mode travelling at 1100 m/s.

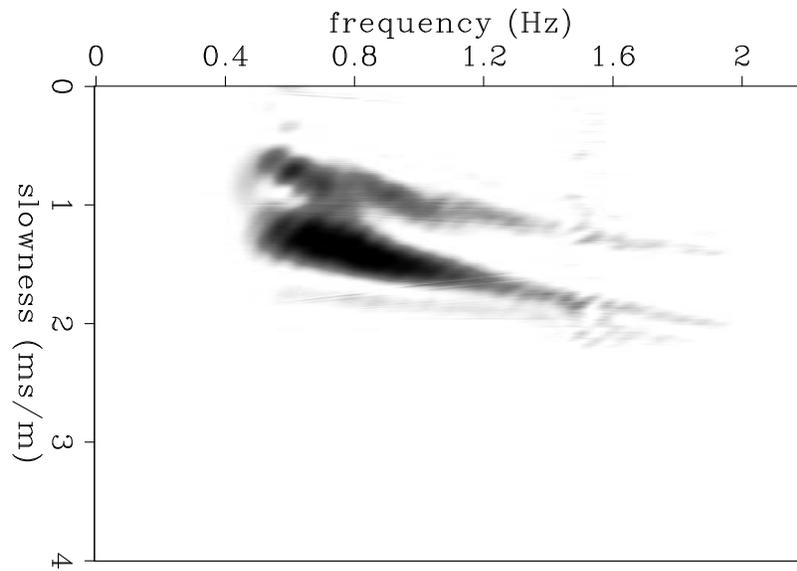
Figure 5(b) shows the dispersion image for the virtual super-source gather for the higher frequency range. I examine the super-source gather rather than a normal source gather for this frequency range due to the lack of convergence for the latter scenario. Much like the lower frequency range, there are the fundamental and first-order Rayleigh wave modes. At 2.2 Hz, there is a spectral peak near 2.2 ms/m, which corresponds to the fundamental mode travelling at 450 m/s, and a minor spectral peak near 1.5 ms/m, which corresponds to the first-order mode travelling at 660 m/s.

These observations fall in line with expectations for dispersive Rayleigh wave modes. Again, for a given wave mode, lower frequencies sense deeper and thus travel at higher velocities than at higher frequencies. Furthermore, the fundamental mode travels slower than the higher-order mode for the same frequency because higher-order modes typically sense deeper (Xia et al., 2003). Higher-order Rayleigh wave modes can potentially be used to help constrain shear wave velocities obtained from inverting dispersion curves.

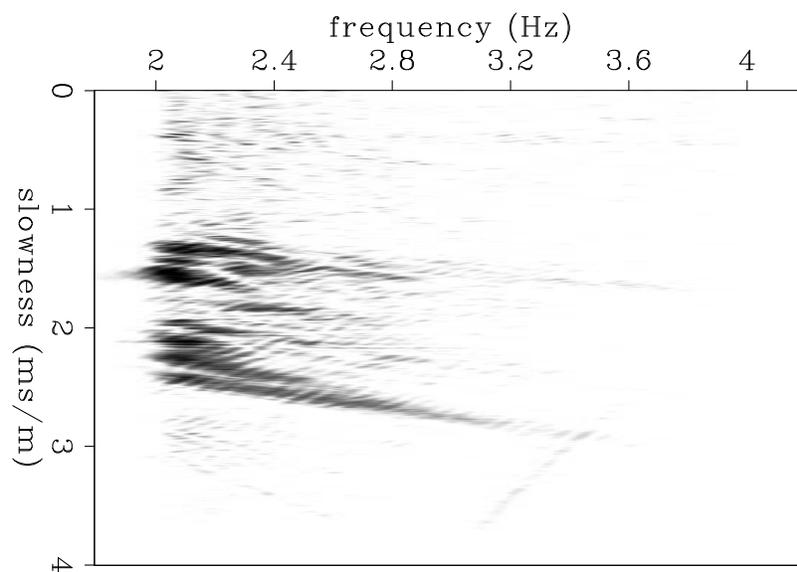
## SUMMARY AND FUTURE WORK

The goal of this report is to constrain the apparent velocity estimates for the Rayleigh and P-waves obtained from seismic interferometry. Chang et al. (2013) estimated the P-wave apparent velocity to be 2000 m/s, based on the slope of the stepout in the virtual super-source gather with 200 m offset binning. Here, I use a  $\tau$ -p transform on the same super-source gather with 50 m offset binning to show that the velocity of the P-wave is closer to 2500 m/s. My dispersion analysis reveals two dispersive wave modes: the fundamental and the first-order Rayleigh wave modes. For a given frequency range, I find that the fundamental mode travels slower than the first-order mode.

The next step is to create phase velocity maps for different frequencies. From these maps, I can generate local dispersion curves that can be inverted for 1D shear velocity profiles. The



(a)



(b)

Figure 5: Dispersion images generated from the  $\tau$ - $p$  transformed gathers in Figure 4. Left: for frequencies between 0.5 and 2 Hz. Right: for frequencies between 2 and 4 Hz. [CR] `jason1/. sh-ph2-down-Jan-mid-1-si-wp,superdupergather1-fine-si-wp`

presence of the first-order Rayleigh wave mode might help constrain those local dispersion curves and thus improve the shear velocity estimates. While I have estimated the velocity of the P-wave, I still need to determine whether this is a direct or a critically refracted P-wave. I therefore intend to stack more correlations into the super-source gather to improve the signal-to-noise ratio at shorter offsets.

## ACKNOWLEDGMENTS

I gratefully acknowledge Signal Hill Petroleum, Inc. for access to this dataset and permission to publish. I also would like to thank Sjoerd de Ridder and Biondo Biondi for their guidance and support throughout this research.

## REFERENCES

- Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: *Geophysical Journal International*, **169**, 1239–1260.
- Chang, J. P., S. de Ridder, and B. Biondi, 2013, Noise characterization and ambient noise cross-correlations at Long Beach: SEP-Report, **149**.
- Claerbout, J., 1985, *Imaging the Earth's Interior*: Blackwell Science, Inc.
- Stoffa, P., 1989, *Tau-p: A Plane Wave Approach to the Analysis of Seismic Data*: Kluwer Academic Publishers.
- Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010, Tutorial on seismic interferometry: Part 1—basic principles and applications: *Geophysics*, **75**, 75A195–75A209.
- Xia, J., R. D. Miller, C. B. Park, and G. Tian, 2003, Inversion of high frequency surface waves with fundamental and higher modes: *Journal of Applied Geophysics*, **52**, 45–57.



## Scholte-wave azimuthal-anisotropic phase-velocity images in the near surface at Ekofisk from seismic noise correlations

*Sjoerd de Ridder, Dave Nichols and Biondo Biondi*

### ABSTRACT

In this report we summarize work done on an ambient seismic noise recording made at the Ekofisk LoFS array. We first isolate the double frequency microseism noise and synthesize virtual seismic sources by cross-correlation. A dispersion analysis shows that these sources contain fundamental mode Scholte-waves. Using Eikonal tomography on the phase-delay times extracted by using the unwrapped instantaneous phase, we construct maps of Scholte wave phase-velocities and elliptical anisotropy. A high velocity anomaly is found in the center of the array, surrounded by a lower velocity region. Under the the southern end of the array we find higher velocities again. We retrieve azimuthal anisotropy that relates to the subsidence pattern.

### AMBIENT SEISMIC FIELD RECORDED BY LOFS AT EKOFISK

Ekofisk has a Life of Field Seismic (LoFS) array installed over a large portion of the North Sea field. Figure 1 contains a map with station locations. The inline and cross-line station spacings are approximately 50 m and 300 m respectively. We received a dataset spanning 41 hours, 28 minutes and 40 seconds starting October 24, 2011, UT 00:17:00. There is about 1 hour and 24 minutes of recording missing before noon on October 24. For 119 stations located in a radius of 750 m of UTM (514,6261) km (blue stations in Figure 1) we plot the spectral amplitudes over time in Figure 2. This spectrogram is computed by Fourier transformation of 2.5-minute recording windows with 50% overlap. Below 0.15 Hz the hydrophone is dominated by ocean swells, which are wind generated gravity waves in the sea (Munk, 1950). These cause pressure variations on the sea floor, and when multiple swells constructively interfere they excite interface waves traveling along the sea-floor at double the frequency (Longuet-Higgins, 1950). This noise is therefore coined double frequency microseism noise. We observe the microseism noise between 0.35 – 1.35 Hz. At frequencies above 1.5 Hz, a variety of predominantly field-operational noise sources dominates the recordings. This study focuses on the microseism noise because it travels along the sea-floor, and potentially provides subsurface information on the top hundreds of meters in the near surface immediately below the sea floor.

To investigate from which direction the ambient noise is incident at Ekofisk, we performed beamforming experiments for the same 119 stations used in the spectrogram calculation. First the recordings are bandpassed for energy between 0.55 – 0.65 Hz and a slant stacked transforming the data to  $\tau - p$  domain. We then average the result for 2.5 minutes and smooth with a triangle over 10 minutes. Figure 3 contains 13 snap shots, each 3 hours apart, from the beamforming experiment. High amplitudes (red) indicate the slowness and azimuth of incoming energy. Generally we observe a circle with absolute slowness

Figure 1: Map of station locations in the Ekofisk LoFS array. Each black dot denotes a station. The stations used to create the spectrogram in Figure 2 and for the beamforming experiments in Figure 3 are blue. [CR] sjoerd1/. Ekofisk-xy-patch1

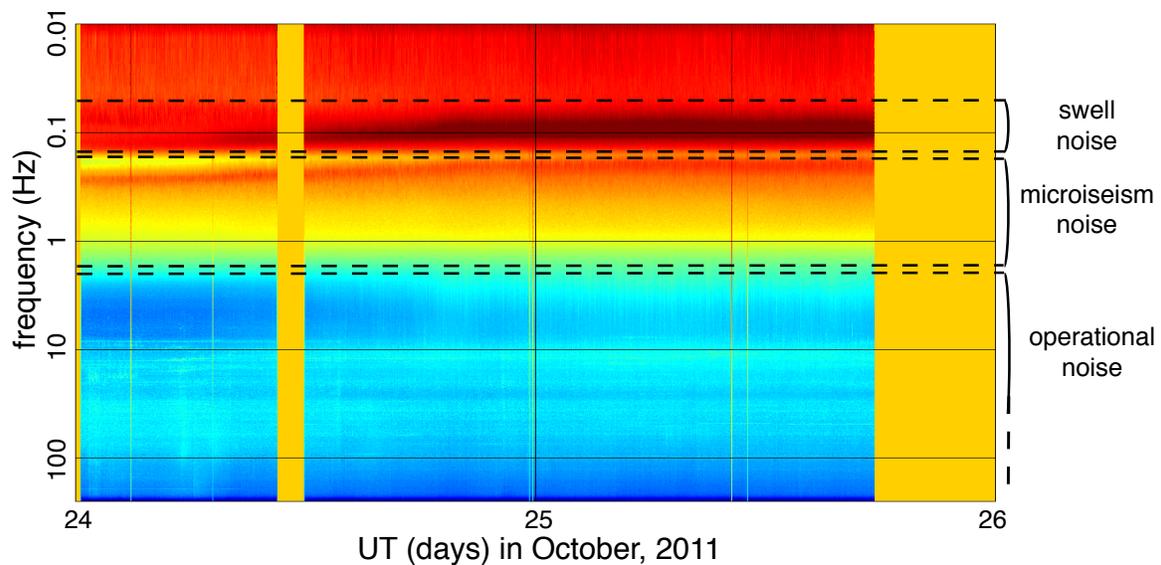
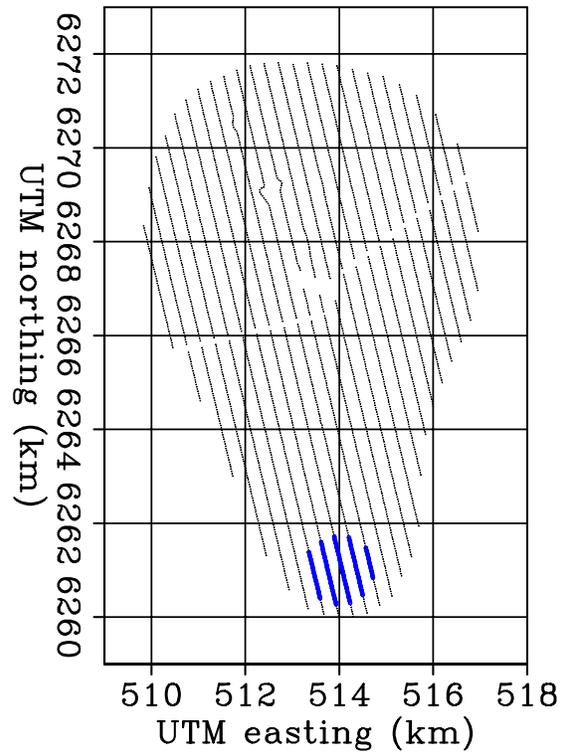


Figure 2: Spectrogram showing spectral amplitudes versus time for the duration of the entire recording used in this study. [CR] sjoerd1/. spectrogram

$|\mathbf{s}| = 0.19$  ms/m, corresponding to a velocity of 525 m/s. These are the Scholte waves that travel in many directions composing the microseism noise of the ambient seismic field. Notice how the noise is remarkably omnidirectional.

## VIRTUAL SOURCES FROM PASSIVE SEISMIC INTERFEROMETRY

Passive seismic interferometry is a technique that cross-correlates seismic recordings at two stations to form a signal, called the estimated Green's function (EGF), as if one of the stations were a seismic source (Claerbout, 1968; Rickett and Claerbout, 1999; Schuster et al., 2004; Wapenaar and Fokkema, 2006). Cross-correlations of passive seismic recordings between all possible station pairs in an array then creates a full virtual seismic survey,  $\mathbf{D}(\omega, x_r, x_s)$ :

$$\mathbf{D}(\omega, x_r, x_s) = \mathbf{r}(\omega, x_r) \mathbf{r}(\omega, x_s)^\dagger, \quad (1)$$

where  $\mathbf{r}(\omega, x)$  is a vector containing passive seismic recordings at all stations, and  $\dagger$  denotes complex conjugation and transposition. By cross-correlating different components of particle velocity we can retrieve all elements of the Green's matrix (Wapenaar and Fokkema, 2006). The cross-correlation signal contains the causal Green's function and its anti-causal reciprocal counter part. However, this technique is valid only when the energy in the ambient seismic field satisfies a condition known as energy equipartitioning. In practice, this requirement limits the application of seismic interferometry to certain frequency regimes. The ambient seismic field at low frequencies (0.5 – 1.3 Hz) is dominated by the double-frequency microseism peak, a source of seismic energy that satisfies the requirement of energy equipartitioning and can be utilized for seismic interferometry (Stewart, 2006; Dellinger and Yu, 2009). The causal Green's function and its anti-causal reciprocal counter part should be an even signal (the cross-correlation signal should be purely real-valued). The anti-symmetric part of the cross-correlation signal reflects deviations from the assumption of energy equipartitioning.

The recorded data in the hydrophone component was filtered using a frequency domain taper with a flat response for 0.4-1.3 Hz, and as a Hann-taper extending from 0.35 Hz to 1.35 Hz. The data was then cross-correlated in 10 blocks of 4 hours. These cross-correlations were then stacked to result in a virtual seismic survey with sources and receivers at all stations in Ekofisk's LoFS array. Figure 4 contains an example of a virtual seismic source in vertical component of particle velocity. Figures 4a to e contain the symmetric part of the cross-correlation signal, while Figures 4f to j contain the asymmetric part. There is very little coherent energy in the asymmetric part and the correlation fluctuations in the asymmetric part are of the order of the background fluctuations in the symmetric part. This indicates that the microseism noise at Ekofisk is very suitable for seismic interferometry.

A dispersion image is calculated as the amplitude in the Radon ( $\omega - p$ ) domain, balanced over frequencies. The dispersion image calculated using all offsets with midpoints between UTM (513–514, 6269–6270) km is shown in Figure 5. The fundamental-mode Scholte waves are clearly visible. It is a dispersive wavemode, i.e. the wavespeed varies with frequency.

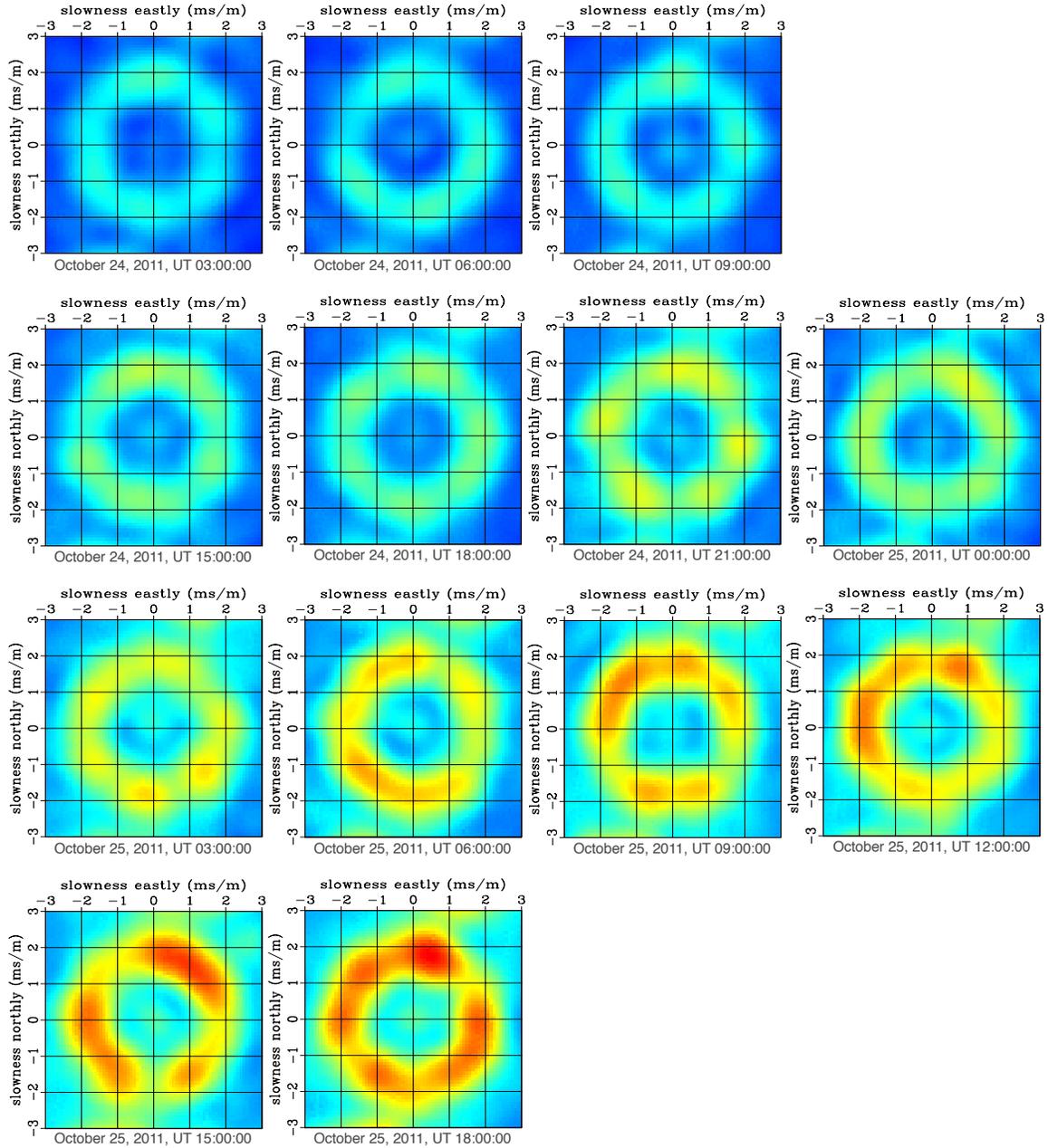


Figure 3: Beamform experiments for recorded data between 0.55 – 0.65 Hz. Frames are generally 3 hours apart (except for missing data) and they indicate the slowness and azimuth of incoming surface wave energy. [CR] [sjoerd1/. ekofisk-beams](#)

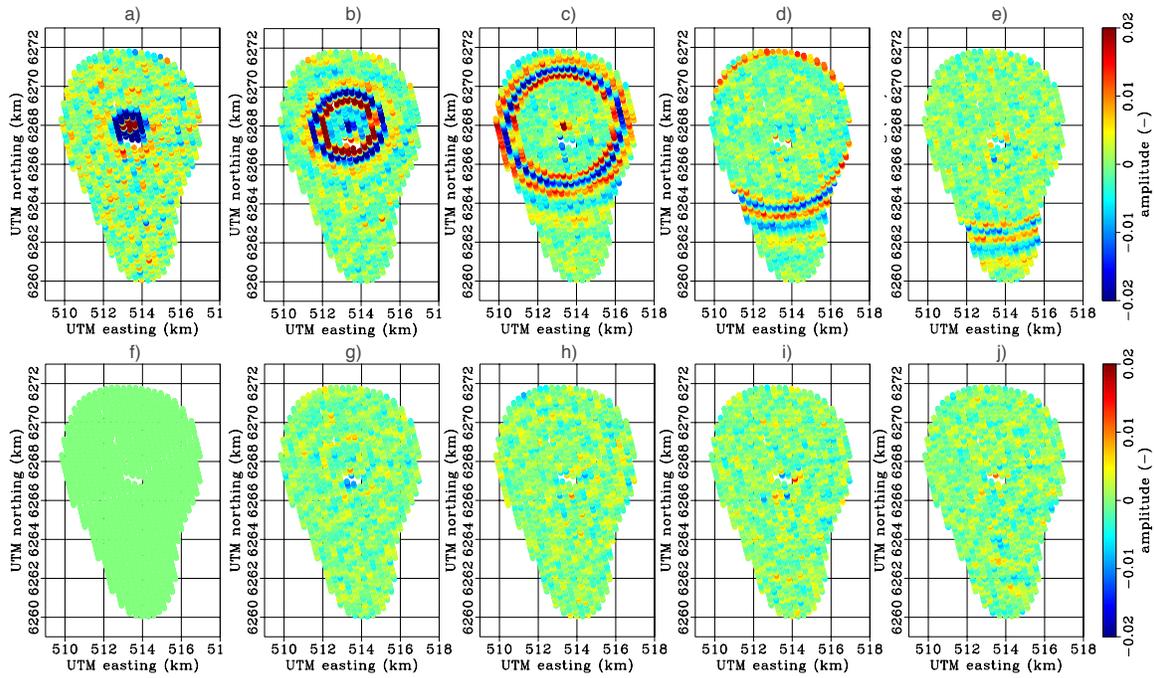
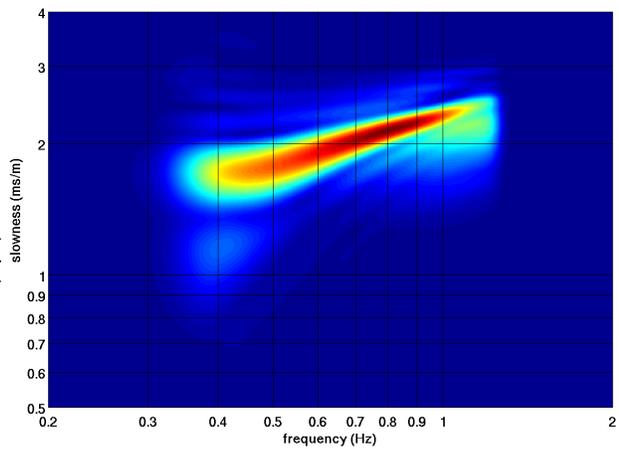


Figure 4: Snap shots for even (a-e) and odd (f-j) part of the virtual seismic sources. At correlation lags: 0s (a and f), 4s (b and g), 8s (c and h), 12s (d and i), and 16s (e and j).

[CR] sjoerd1/. ekofisk-shots

Figure 5: Dispersion image for the virtual seismic sources. [CR] sjoerd1/. TAPES-cables-compP-sif-cmpsg.



## AZIMUTHAL-ANISOTROPIC EIKONAL TOMOGRAPHY FOR PHASE VELOCITY USING THE INSTANTANEOUS PHASE

The surface waves emitted by the virtual seismic sources propagate along the seafloor with wavelengths between 1100 m at 0.5 Hz and 375 m at 1.2 Hz (from Figure 5). These long wavelengths provide sensitivity away from the seabed (thus in depth), and different frequencies will be sensitive differently with depth. Therefore, we are interested in creating phase velocity maps, showing the velocities as a function of space for a given frequency. We thus need to measure the apparent traveltime for a Scholte-wave for a given frequency; phase-delay times.

Aki and Richards (2002) showed that the fundamental mode Scholte wave Green's functions in the far field behave like:

$$G(\mathbf{x}_r, \mathbf{x}_s) \propto \exp \left[ -i \left( \frac{\omega}{c(\omega)} |\mathbf{x}_s - \mathbf{x}_r| + \pi/4 \right) \right]. \quad (2)$$

If we take the instantaneous phase of the cross-correlations and add a  $\frac{\pi}{4}$  phase shift, we find a phase-function that should be proportional to  $\omega \frac{|\mathbf{x}_s - \mathbf{x}_r|}{c(\omega)}$ . Thus it makes sense to define the phase-delay time that given a phase velocity  $c(\omega)$  is a linear function of source-receiver distance,  $\tau(\omega) = \frac{|\mathbf{x}_s - \mathbf{x}_r|}{c(\omega)}$ . This phase-delay time represents a travel-time for Scholte waves at a particular frequency.

Eikonal tomography is a technique that evaluates the Eikonal equation on travel-time surfaces to directly find slowness estimates. Both Lin et al. (2009) and Mordret et al. (2013) use an isotropic Eikonal equation to estimate the azimuthal-anisotropic local slowness of wave-propagation. A disadvantage of that method is that the observed azimuth of fast and slow directions cannot easily be regularized over space. Here I propose a different approach.

I assume the fundamental mode Scholte waves exhibit elliptical anisotropy, such that velocity can be expressed as:

$$V^2(\phi) = V_f^2 \cos^2(\phi - \alpha) + V_s^2 \sin^2(\phi - \alpha), \quad (3)$$

where  $V_f$  and  $V_s$  are the maximum and minimum velocities (for  $V$  over all angles  $\phi$ ) and  $\alpha$  is the fast direction of velocity. This velocity is plugged into the dispersion relation:

$$-\omega^2 = V^2(\phi) [k_x^2 + k_y^2], \quad (4)$$

defining a frequency domain equation:

$$-\omega^2 U(\mathbf{x}, \omega) = [(V_f^2 - V_s^2) \cos^2(\alpha) + V_s^2] \partial_x^2 U(\mathbf{x}, \omega) + \quad (5)$$

$$2 [(V_f^2 - V_s^2) \cos(\alpha) \sin(\alpha)] \partial_x \partial_y U(\mathbf{x}, \omega) + \quad (6)$$

$$[(V_f^2 - V_s^2) \sin^2(\alpha) + V_s^2] \partial_y^2 U(\mathbf{x}, \omega). \quad (7)$$

To derive an appropriate Eikonal equation, we plug the test function,  $U(\mathbf{x}, \omega) = A(\mathbf{x}, \omega) \exp \{-i\omega T(\mathbf{x}, \omega)\}$ , into equation 7 and collect the leading terms (in  $\omega^2$ ) to find:

$$1 = \begin{bmatrix} \partial_x T(\mathbf{x}, \omega) & \partial_y T(\mathbf{x}, \omega) \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix} \begin{bmatrix} \partial_x T(\mathbf{x}, \omega) \\ \partial_y T(\mathbf{x}, \omega) \end{bmatrix}. \quad (8)$$

For the matrix elements we have

$$M_{11} = (V_f^2 - V_s^2)\cos^2(\alpha) + V_s^2 \quad (9)$$

$$M_{12} = (V_f^2 - V_s^2)\cos(\alpha)\sin(\alpha) \quad (10)$$

$$M_{22} = (V_f^2 - V_s^2)\sin^2(\alpha) + V_s^2. \quad (11)$$

The eigenvalues of the matrix  $\mathbf{M}$  are  $V_f^2$  and  $V_s^2$ , and the tangent of the eigenvectors indicate the fast and slow directions. One mayor advantage of this approach is that the elements of this matrix can be regularized over space.

The easterly and northerly derivatives are calculated by simple centered finite difference on a phase-delay time map interpolated by splines (Sandwell, 1987) to a grid of 100 m by 100 m. We invert for the matrix  $\mathbf{M}$  using an  $L_1$  norm of the logarithm of equation 8 (after discarding outliers of measurements of derivatives). The matrix  $M$  is regularized over space using a Laplacian. Figure 6 provides an schematic overview of the input of Eikonal tomography for two sources.

Phase-delay times are plotted in Figures 6a and 6d. Easterly derivatives of the interpolated phase-delay times are shown in Figures 6b and 6e and northerly derivatives of the interpolated phase-delay times are shown in Figures 6c and 6f.

We construct maps of elliptical anisotropic Scholte-wave phase-velocities for each frequency. Figure 7 contain maps of Scholte-wave phase velocity at 0.5 Hz to 1.3 Hz. The colors depict the isotropic component,  $V_0 = (V_f + V_s)/2$ . The azimuth of the dashes denote the fast direction of anisotropy. The length of the dashes indicate the magnitude of anisotropy,  $dV = (V_f - V_s)/2$ , as a fraction of the isotropic component. The dash in the upper-right corner of each plot denotes a anisotropic magnitude of 10%. For comparison the bathymetry is included through blue contour lines with a 2 m interval. For some frequencies it is hard to distinguish an interpretable pattern in the anisotropy. However, for the Scholte wave phase velocities at 1.0 Hz we observe a pattern that relates to the seafloor subsidence. In addition, we generally find higher Scholte-wave velocities where the sea-floor curvature is low, and lower Scholte-wave velocities where the sea-floor curvature is high. The edge of high velocities at the northern edge of the recovered map, may be an extrapolation artifact but may also be the rim of the subsidence pattern.

For dispersive interface waves, lower frequencies generally relate to deeper regions (Aki and Richards, 2002). When we arrange maps of phase velocity for a set of frequencies as a cube with frequency as vertical axis, we find a phase-velocity cube  $(\omega, x, y)$  that forms a first order image of Scholte waves in the top few 100 m of the near surface. Figure 8 contains three slices through this cube. Notice how the generally higher velocities in the center of the subsidence pattern reach to extend all the way to the lowest frequencies. We notice a significant high velocity region in the very center of the array at very low frequencies.

## SUMMARY

In this report it is shown that the low frequency ambient seismic field as recorded by the Ekofisk LoFS array is suitable for retrieval of Scholte waves by seismic interferometry. In fact, the microseism energy at Ekofisk in this recording is very uniformly distributed over azimuth, and thus ideal for seismic interferometry. However, different recordings have different

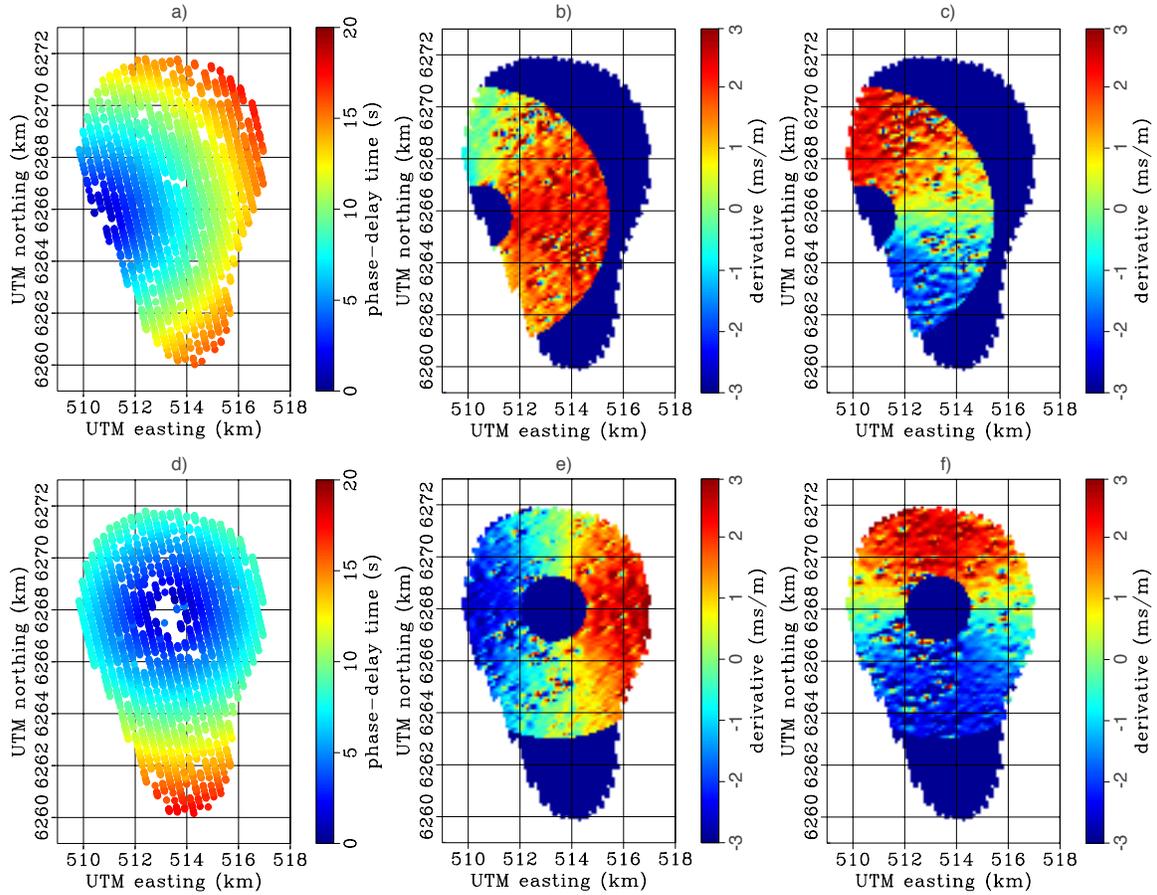


Figure 6: Schematic of frequency-domain Eikonal tomography. Phase delay times at 1.0 Hz for two virtual seismic sources (a and d). Easterly derivatives (b and e) and northerly derivatives (c and f) of the interpolated phase-delay surfaces. [CR] [sjoerd1/. anikonal-schematic](#)

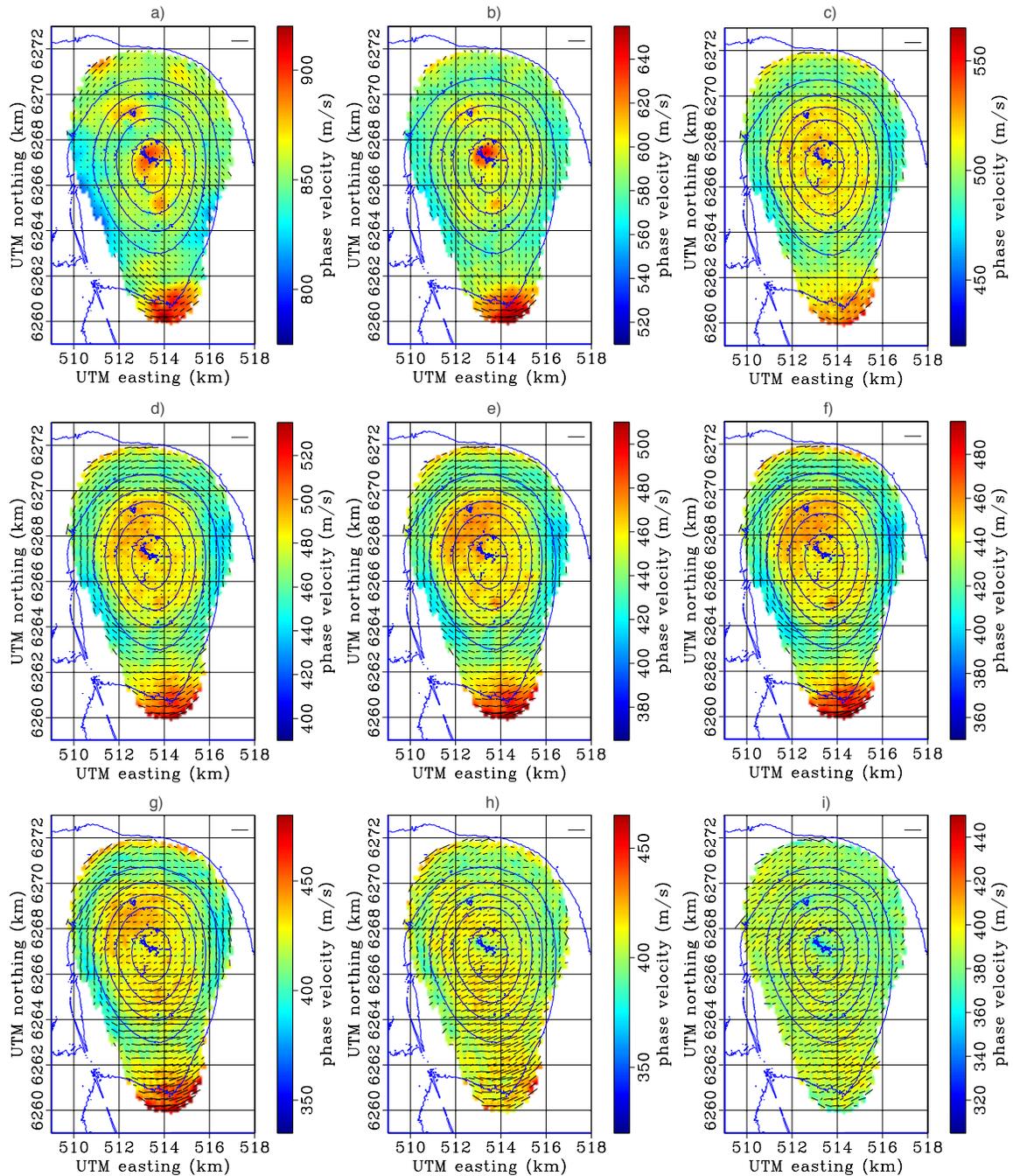


Figure 7: Scholte wave phase-velocities for 0.5 Hz (a), 0.6 Hz (b), 0.7 Hz (c), 0.8 Hz (d), 0.9 Hz (e), 1.0 Hz (f), 1.1 Hz (g), 1.2 Hz (h), 1.3 Hz (i). The colors depict the isotropic component. The azimuth of the dashes denote the fast direction anisotropy. The length of the dashes indicate the magnitude of anisotropy as a fraction of the isotropic component. The dash in the upper-right corner of each plot denotes a anisotropic magnitude of 10%. The bathymetry is included by blue contour lines with a 2 m interval. [CR]

sjoerd1/. Ekofisk-Meikon-al-maps

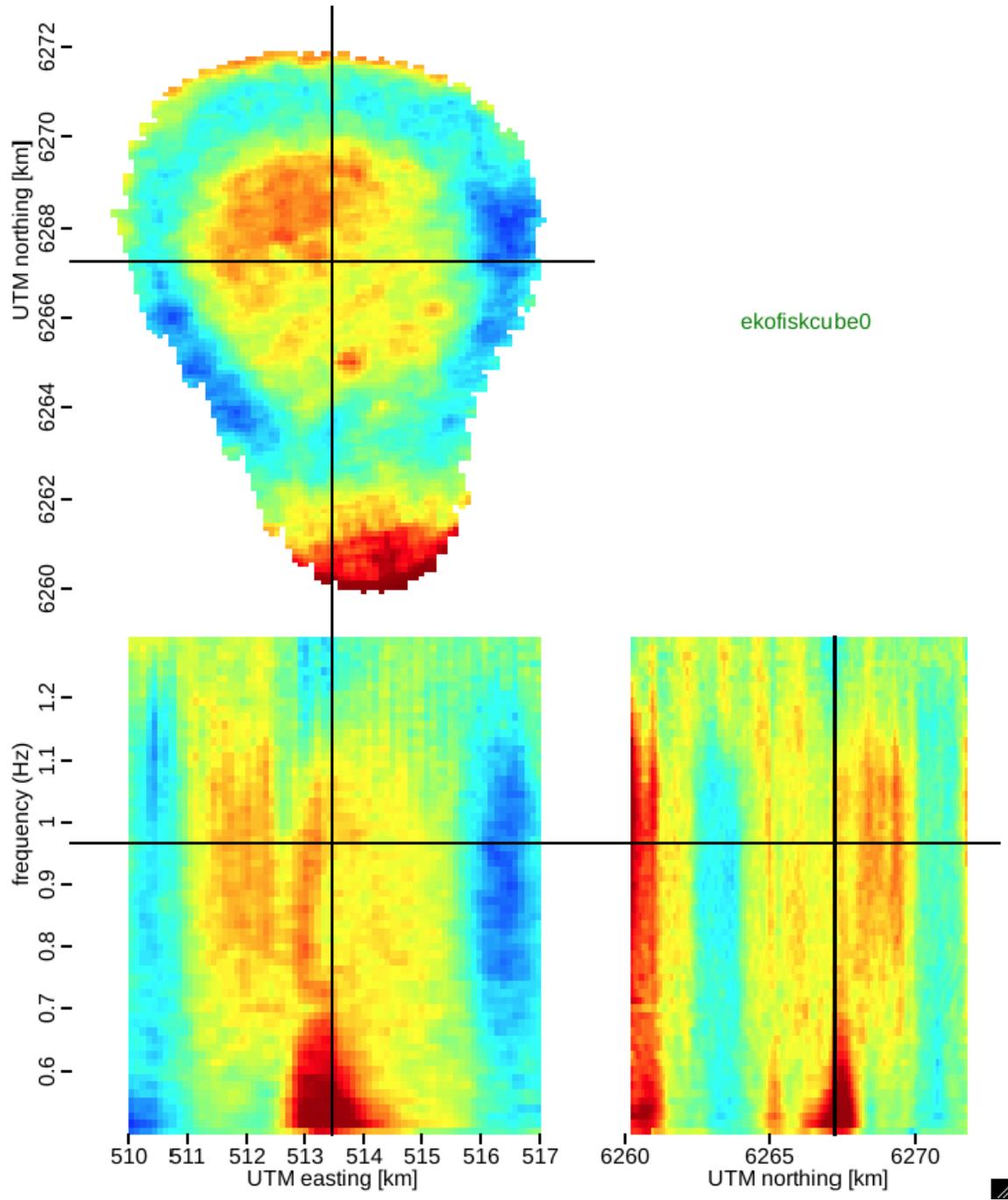


Figure 8: Cube of the isotropic part of Scholte wave velocities,  $(\omega, x, y)$ . [NR]

sjoerd1/. ekofiskcube

characteristics because noise characteristics are transient with time. Using a tomographic method based on evaluating the Eikonal equation over a phase delay-time surface, we recover a phase velocity map for 1.0 Hz. A high velocity anomaly is found in the center of the array, surrounded by a lower velocity region. Under the the southern end of the array we find higher velocities again. We retrieve azimuthal anisotropy that relates to the subsidence pattern.

## ACKNOWLEDGMENTS

ConocoPhillips Skandinavia AS and the PL018 Partnership (Total E&P Norge AS, ENI Norge AS, Statoil Petroleum AS and Petoro AS) for access to Ekofisk data, and permission to publish these results. Olaf Knoth, Ali Tura, Alexandre Bertrand and Roman Kazinnik for helpful discussions and suggestions.

## REFERENCES

- Aki, K., and P. G. Richards, 2002, *Quantitative Seismology - second edition*: University Science Books.
- Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: *Geophysics*, **33**, 264–269.
- Dellinger, J. A., and J. Yu, 2009, Low-frequency virtual point-source interferometry using conventional sensors: 71st Meeting, European Association of Geoscientists and Engineers, Expanded Abstracts, X047.
- Lin, F. C., M. H. Ritzwoller, and R. Snieder, 2009, Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array: *Geophys. J. Int.*, **177**, 1091–1110.
- Longuet-Higgins, M. S., 1950, A theory of the origin of microseisms: *Phil. Trans. R. Soc. Lond. A*, **243**, 135.
- Mordret, A., N. Shapiro, S. Singh, P. Roux, and O. Barkved, 2013, Helmholtz tomography of ambient noise surface wave data to estimate scholte wave phase velocity at valhall life of the field: *GEOPHYSICS*, **78**, WA99–WA109.
- Munk, W. H., 1950, Origin and generation of waves: *Proceedings 1st International Conference on Coastal Engineering*, 1–4.
- Rickett, J., and J. Claerbout, 1999, Acoustic daylight imaging via spectral factorization: *Helioseismology and reservoir monitoring: The Leading Edge*, **18**, 957–960.
- Sandwell, D. T., 1987, Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data: *Geophys. Res. Lett.*, **14**, 139–142.
- Schuster, G. T., J. Yu, J. Sheng, and J. Rickett, 2004, Interferometric/daylight seismic imaging: *Geophys. J. Int.*, **157**, 838–852.
- Stewart, P., 2006, Interferometric imaging of ocean bottom noise: *SEG Technical Program Expanded Abstracts*, **25**, 1555–1559.
- Wapenaar, K., and J. Fokkema, 2006, Green's function representations for seismic interferometry: *Geophysics*, **71**, SI33–SI46.



## Scholte-wave excitation

*Marine Denolle, Sjoerd de Ridder, Jason P. Chang, Eileen R. Martin, Taylor Dahlke, Humberto Arevalo-Lopez, Sr., and Stewart A. Levin<sup>1</sup>*

### ABSTRACT

We estimate the excitation of the Scholte waves using a new formulation of the surface-wave eigenproblem. We adapt the Rayleigh-wave case for solid media to accommodate the fluid shear-free condition and successfully calculate the Scholte-wave excitation. We detail here the derivation and numerical implementation, along with preliminary results for simple fluid-over-solid cases. We verify our results by comparing our phase velocity dispersion curve to the numerical solution of the dispersion relation for a fluid layer above an elastic half-space.

### INTRODUCTION

A modal representation of interface waves can be used to construct waveform solutions to the wave equation. Free air-to-solid interface waves are commonly called Love and Rayleigh waves, solid-to-solid interface waves are usually called Stoneley waves, and fluid-to-solid interface waves, the focus of our interest, are referred to as Scholte waves. The excitation of these interface waves can be reduced to a generalized eigenproblem in the frequency domain. Denolle et al. (2012) solve this problem using a Chebyshev collocation method and successfully define the Rayleigh- and Love-wave modes for the single- and multi-layer solid cases. In this work, we generalize the code from Denolle et al. to handle a fluid layer at the surface in order to construct solutions for Scholte waves. We first expand on the implementation of boundary conditions at the fluid-to-solid interface in the generalized eigenproblem case. We then show the resulting eigenvalues and eigenfunctions for a simple two-layer, fluid-solid medium, which we validate against known algorithms.

### INTRA-LIQUID FORMULATION

We first need to formulate the interface-wave eigenproblem with the proper boundary conditions to accommodate the fluid layer. Referring to the classic literature, e.g. Ewing et al. (1957), the behavior of waves in a fluid is correctly modeled by setting the shear modulus  $\mu$  to zero in the equations for an isotropic elastic medium. In cylindrical coordinates, this makes equation 3 in Denolle et al. reduce to

$$u = [r_1(k, z, \omega)\mathbf{S}_k^m(r, \phi) + ir_2(k, z, \omega)\mathbf{R}_k^m(r, \phi)] e^{-i\omega t} . \quad (3\text{Den})$$

---

<sup>1</sup> Chang, Dahlke and Martin are the main authors on the paper. Denolle and De Ridder led the summer mini seminar and authored the code and its modification respectively.

In this formulation,  $\mathbf{S}_k^m(r, \phi)$  is the gradient with respect to the radial and angular directions (holding depth constant) of  $J_m(kr)e^{im\phi}$ , and  $J_m(\cdot)$  is a  $m^{\text{th}}$  order Bessel function.  $\mathbf{R}_k^m(r, \phi)$  is equal to  $-J_m(kr)e^{im\phi}$ . In other words, we project our solution on those vectors that describe modal solutions of the wave equation in cylindrical coordinates. The scalar stress-displacement values for the Rayleigh waves are  $(r_1, r_2, r_3, r_4)$ , and  $(l_1, l_2)$  for the Love waves. For reference, explicit expressions for the Rayleigh wave  $r_3$  and  $r_4$  components are

$$r_3 = \sigma_{rz} = \mu \left( \frac{dr_1}{dz} - kr_2 \right) \quad \text{and}$$

$$r_4 = i\sigma_{zz} = k\lambda r_1 + (\lambda + 2\mu) \frac{dr_2}{dz} .$$

Equation 5 becomes

$$\begin{aligned} -\rho\omega^2 l_1 &= 0 \\ -\rho\omega^2 r_1 &= -k^2 \lambda r_1 - k\lambda \frac{dr_2}{dz} , \\ -\rho\omega^2 r_2 &= \frac{d}{dz} \left[ \lambda \frac{dr_2}{dz} + k\lambda r_1 \right] \end{aligned} \quad (5\text{Den})$$

and equation 9 reduces to

$$R_3 = k\lambda r_1 + \lambda \frac{dr_2}{dz} . \quad (9\text{Den})$$

Equation 8 yields

$$L_2 = [-i\sigma_{r\phi}] = 0 , \quad (8\text{Den})$$

and equation 10 becomes

$$R_4 = 0 . \quad (10\text{Den})$$

These last two equations are intuitive when coupled with  $l_1 = 0$  from equation 5, as ideal liquids do not support shear waves. As a result, all nondiagonal stresses (particularly  $\sigma_{r\phi}$  and  $\sigma_{rz}$ ) are thus zero. The free-shear condition also presumes that the Love waves do not exist in the upper fluid layer. In other words, ignore the fluid when solving for Love wave modes.

For the Rayleigh eigenproblem, where  $\rho$  is the medium density, and  $\lambda$  and  $\mu$  are the Lamé parameters, we can set  $\mu$  and  $R_4$  to zero from equation 12 to get

$$\begin{bmatrix} 0 & -\frac{d}{dz} & \frac{1}{\lambda} & 0 \\ \frac{d}{dz} & 0 & 0 & -\frac{1}{0} \\ \rho\omega^2 & 0 & 0 & \frac{d}{dz} \\ 0 & -\rho\omega^2 & -\frac{d}{dz} & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ R_3 \\ 0 \end{bmatrix} = k \begin{bmatrix} r_1 \\ r_2 \\ R_3 \\ 0 \end{bmatrix} . \quad (12\text{Den})$$

In order to have a system of three equations to solve for three unknowns, we eliminate a row of the matrix. Because of the singularity that arises in the second row of the matrix,

we eliminate this equation, leaving the 3x3 system:

$$\begin{bmatrix} 0 & -\frac{d}{dz} & \frac{1}{\lambda} \\ \rho\omega^2 & 0 & 0 \\ 0 & -\rho\omega^2 & -\frac{d}{dz} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ R_3 \end{bmatrix} = k \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ R_3 \end{bmatrix}. \quad (12Den)$$

## FLUID-SOLID BOUNDARY CONDITIONS

For the Scholte waves, the fluid couples to the top of the elastic model with the continuity of vertical displacement and normal stress (Ewing et al., 1957; Aki and Richards, 1980). Following Denolle et al., this translates to the continuity of  $r_2$  and  $(\lambda + 2\mu)\frac{dr_2}{dz} + k\lambda r_1$ , in addition to  $R_4 = 0$ . For the free surface,  $\lambda\frac{dr_2}{dz} + \lambda k r_1 = 0$  is the appropriate free shear boundary condition.

At the fluid-solid interface, we want continuity of  $(\lambda + 2\mu)\frac{dr_2}{dz} + \lambda k r_1$ , and can use equation 9 from Denolle et al. to translate this to a condition on  $R_3$ :

$$\begin{aligned} (\lambda + 2\mu)\frac{dr_2}{dz} + \lambda k r_1 &= \frac{\lambda + 2\mu}{\lambda} (R_3 - k(\lambda + 2\mu)r_1) + \lambda k r_1 \\ &= \left(1 + \frac{2\mu}{\lambda}\right) R_3 + k \left(\lambda - \frac{(\lambda + 2\mu)^2}{\lambda}\right) r_1. \end{aligned}$$

## IMPLEMENTATION

We strove to keep the implementation as simple as possible. To this end, our code alters the boundary conditions for the surface and fluid-solid interface as described above and selects only a subset of the equations within the liquid layer.

For both the Love- and Rayleigh-wave eigenvalue problems, we can write the eigenvalue system generally as  $A\mathbf{x} = kB\mathbf{x}$ , where  $B$  is mostly an identity matrix except for a few entries altered for boundary conditions. We create the matrices  $A$  and  $B$  as though modeling solid layers for which  $\mu$  just happens to equal zero in the top layer, as described in Denolle et al. (2012). Then we apply matrices to the left and right of these to select the relevant rows and columns so that we solve a new eigenvalue problem,  $L_r A L_c \mathbf{x} = k L_r B L_c \mathbf{x}$ .

For the Love waves, we remove both of the equations related to the water layer since we know the solution will be zero in a liquid. For example, consider a liquid layer above a solid layer with constant parameters within each layer. In this case, the matrices selecting rows and columns are

$$L_r = L_c^* = \left[ \begin{array}{cc|cc} I_{N_s} & 0_{N_s} & 0_{N_l} & 0_{N_l} \\ 0_{N_s} & I_{N_s} & 0_{N_l} & 0_{N_l} \end{array} \right],$$

where  $N_s$  and  $N_l$  are the number of points in the solid and liquid layer, respectively,  $0_n$  represents the zero matrix of size  $n \times n$ , and  $I_n$  represents the  $n \times n$  identity matrix.

For the Rayleigh wave we have a set of four equations. In the liquid layer, we set the shear wave speed to zero. Equation 12Den shows that we select only the first, third and fourth rows and the first, second and third columns of the matrices' liquid section. Looking

again at the example of a liquid layer above a single solid layer, we achieve this by choosing our row and column selection matrices as follows

$$L_r = \left[ \begin{array}{c|cccc} I_{4 \times N_s} & 0 & 0 & 0 & 0 \\ \hline 0 & I_{N_l} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{N_l} & 0 \\ 0 & 0 & 0 & 0 & I_{N_l} \end{array} \right], \quad L_c = \left[ \begin{array}{c|ccc} I_{4 \times N_s} & 0 & 0 & 0 \\ \hline 0 & I_{N_l} & 0 & 0 \\ 0 & 0 & I_{N_l} & 0 \\ 0 & 0 & 0 & I_{N_l} \\ 0 & 0 & 0 & 0 \end{array} \right].$$

## RESULTS

We display results from the modified code for a fluid-solid interface. The model parameters that we used are shown in Figure 1. For our calculations, both layers contain 50 collocation points. We assume that the solid is a Poisson medium when computing our associated Lamé parameters.

Figure 1: Profile of our 1D model of a fluid layer over a solid. Left: P-wave (red) and S-wave (blue) velocities with depth. Right: Density with depth. [CR] surfwaveminisem/. profile

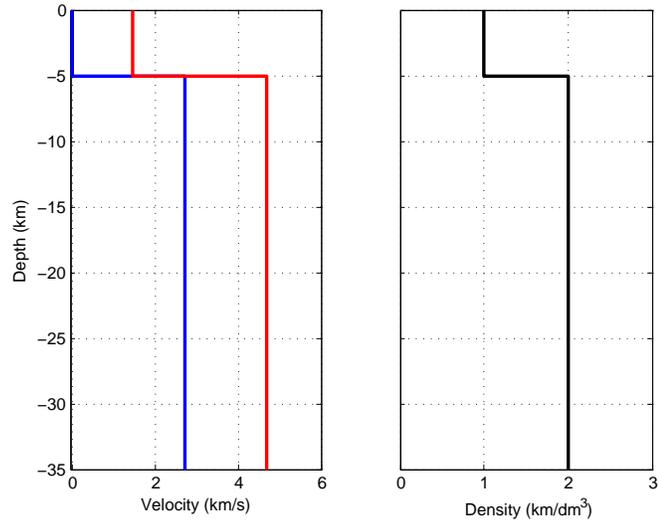


Figure 2 shows the computed fundamental-mode Scholte displacement eigenfunctions for the previously described 1D medium. We show the Scholte-wave eigenfunctions for three frequencies: 0.05, 0.10, and 1.00 Hz. We see that the horizontal displacement ( $r_1$ ) is discontinuous at the fluid-solid interface and that the vertical displacement ( $r_2$ ) is continuous, which are the boundary conditions that we imposed. We also see that the eigenfunction is more sensitive to the interface at higher frequencies than at lower frequencies. This makes sense, since lower frequencies have longer wavelengths that will be less affected by the interface. Figure 3, computed using the Haskell matrix codes of Herrmann (2010), further confirms the validity of our modifications.

Figure 4 shows the computed fundamental mode Scholte-wave stress eigenfunctions. Again, we show the eigenfunctions for three frequencies: 0.05, 0.10, and 1.00 Hz. We see that the shear stress ( $R_3$ ) is discontinuous at the interface, and that the normal stress ( $R_4$ ) is continuous across the interface. These observations are what we expect. Also note that much like with the displacement eigenfunctions, sensitivity of the stress eigenfunctions at the fluid-solid interface increases with frequency. Again, this is likely related to lower

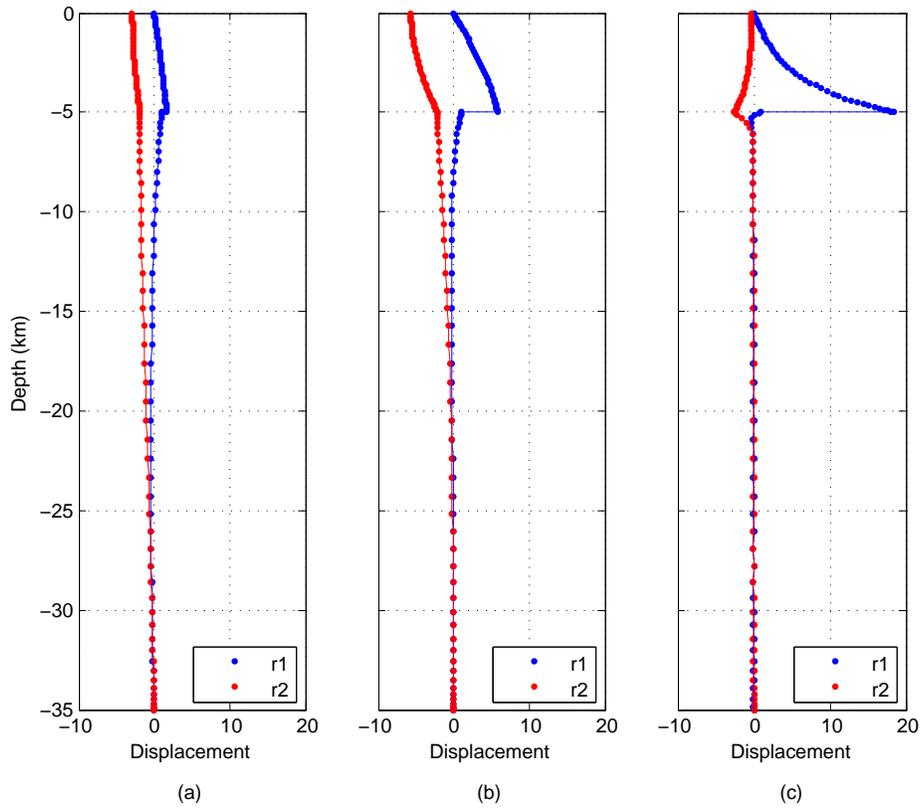
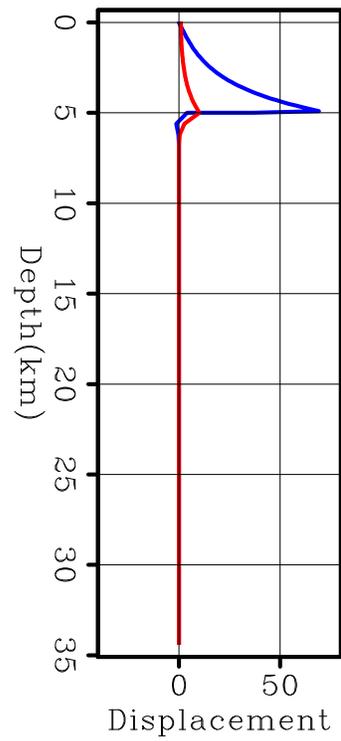


Figure 2: Displacement eigenfunctions for the described fluid-solid 1D medium. For Scholte waves: (a) 0.05 Hz, (b) 0.10 Hz, and (c) 1.00 Hz.  $r_1$  is the horizontal displacement and  $r_2$  is the vertical displacement. [CR] `surfwaveminisem/. disp-eigs`

Figure 3: 1.00 Hz displacement eigenfunctions for the described fluid-solid 1D medium computed using the Haskell matrix method. Except for the packaged choices for overall normalization and the sign for radial motion, this is in good agreement with Figure 2. [CR] `surfwaveminisem/. SRDER`



frequencies having longer wavelengths and hence being less sensitive to relatively shallow interfaces.

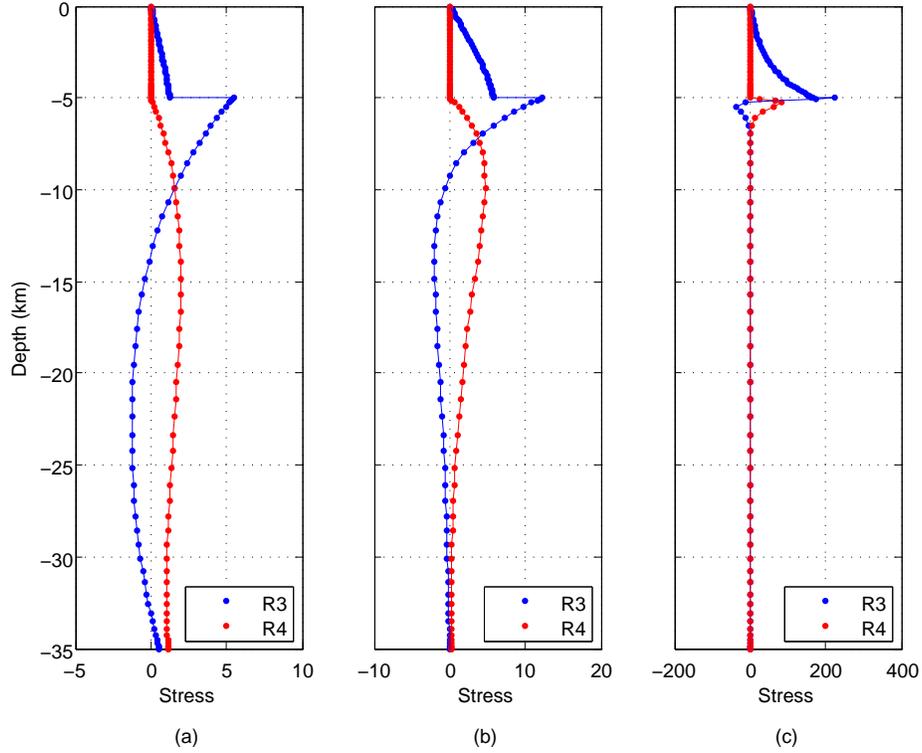


Figure 4: Scholte-wave stress eigenfunctions for the described fluid-solid 1D medium at (a) 0.05 Hz, (b) 0.10 Hz, and (c) 1.00 Hz. R3 is the shear stress and R4 is the normal stress. Note that the scales on stress differ between each frequency. **[CR]** surfwaveminisem/. stress-eigs

To view the eigenvalues (wavenumbers), we plot phase velocity dispersion curves for Scholte waves in Figure 5. We obtain these curves by solving

$$c_S = \frac{\omega}{k_S(\omega)}, \quad (1)$$

where  $c_S$  is the Scholte-wave phase velocity,  $\omega$  is angular frequency, and  $k_S$  is wavenumber as a function of frequency. To verify these results, we compare these dispersion curves to the numerically calculated Scholte wave phase velocity,  $c$ , which is the solution to the dispersion relation for a finite fluid layer over an elastic half-space (Biot, 1952). The relation is given as

$$4\sqrt{1 - \frac{c^2}{v_s^2}} - \frac{(2 - \frac{c^2}{v_s^2})^2}{\sqrt{1 - \frac{c^2}{v_p^2}}} = \frac{\rho_f}{\rho_s} \frac{\frac{c^4}{v_s}}{\sqrt{\frac{c^2}{v_f^2} - 1}} \tan \left[ kh \sqrt{\frac{c^2}{v_f^2} - 1} \right] \quad \text{for } \frac{c}{v_f} > 1, \quad (2)$$

$$4\sqrt{1 - \frac{c^2}{v_s^2}} - \frac{(2 - \frac{c^2}{v_s^2})^2}{\sqrt{1 - \frac{c^2}{v_p^2}}} = \frac{\rho_f}{\rho_s} \frac{\frac{c^4}{v_s}}{\sqrt{1 - \frac{c^2}{v_f^2}}} \tanh \left[ kh \sqrt{1 - \frac{c^2}{v_f^2}} \right] \quad \text{for } \frac{c}{v_f} < 1, \quad (3)$$

where  $k$  is wavenumber,  $h$  is the depth of the fluid layer,  $\rho_f$  is the fluid density, and  $\rho_s$  is the solid density. Furthermore,  $\eta_1 = \frac{c}{v_s}$ ,  $\eta_2 = \frac{c}{v_p}$ ,  $\eta_f = \frac{c}{v_f}$ ,  $v_p$  is P-wave velocity in the solid,

$v_s$  is the S-wave velocity in the solid, and  $v_f$  is the velocity in the fluid.

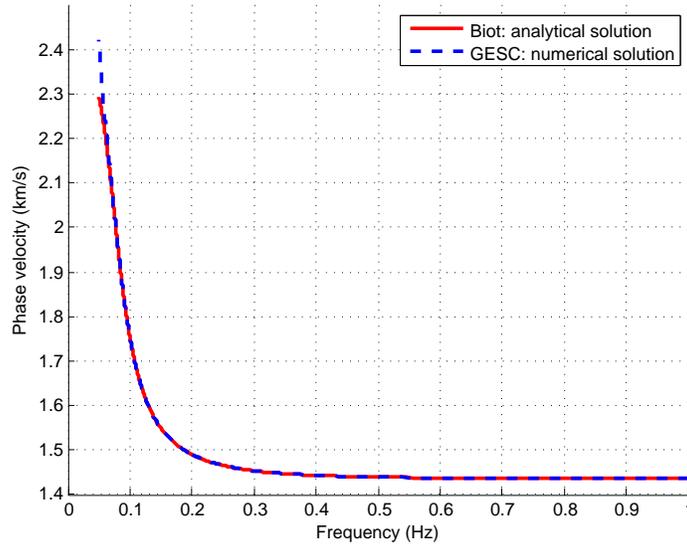


Figure 5: A comparison of phase velocity dispersion curves for a fluid-solid interface. Blue: our code. Red: numerical solution of the analytically-derived dispersion relation from Biot (1952). The solutions match up very well, suggesting that our modifications are correct. **[CR]** `surfwaveminisem/. dispersion-legend`

Figure 5 shows that the numerically calculated reference solution (red) and our solution (blue) to the Scholte-wave dispersion relation match very well. The deviation between the two solutions at frequencies below 0.1 Hz reflects our approximation of the half-space with a thick layer which has a rigid bottom boundary condition. Therefore, at these lower frequencies, we are approaching wavelengths that are longer than the domain of the model, causing our phase velocities to deviate from the numerical solution. Regardless, there is a clear match for the higher frequencies, which suggests that our calculated eigenfunctions and eigenvalues are indeed correct for Scholte waves.

## CONCLUSIONS

Denolle et al. (2012) obtained the Rayleigh- and Love-wave solutions to the elastic wave equation by posing the problem in the generalized matrix eigenvalue framework. Here we expanded this framework to also solve for Scholte waves, and added this functionality to their code. We verified our theory and its implementation by finding that our phase velocity dispersion curves match well with the numerical solution to the dispersion relation for a finite liquid layer over an elastic half space of Biot (1952). We also found that our displacement eigenfunctions match well with those from the Haskell matrix codes of Herrmann (2010).

## ACKNOWLEDGMENTS

We would like to thank Eric Dunham for helpful discussions and suggestions. We would also like to thank Gabe Lotto for providing code to verify our results.

**REFERENCES**

- Aki, K. and P. G. Richards, 1980, *Quantitative seismology*, volume **1424**: Freeman San Francisco.
- Biot, M. A., 1952, The interaction of Rayleigh and Stoneley waves in the ocean bottom: *Bulletin of the Seismological Society of America*, **42**, 81–93.
- Denolle, M. A., E. M. Dunham, and G. C. Beroza, 2012, Solving the surface-wave eigenproblem with Chebyshev spectral collocation: *Bulletin of the Seismological Society of America*, **102**, 1214–1223.
- Ewing, W. M., W. S. Jardetzky, F. Press, and A. Beiser, 1957, *Elastic waves in layered media*: *Physics Today*, **10**, 27.
- Herrmann, R. B., 2010, *Computer programs in seismology*, version 3.33.



## Equivalent accuracy at a fraction of the cost: Tackling spatial dispersion

*Huy Le, Robert G. Clapp, and Stewart A. Levin*

### ABSTRACT

To reduce numerical spatial dispersion and find an optimal set of finite difference coefficients for a given frequency bandwidth and a range of velocities, we minimize the weighted sum of the squared error between the finite difference operator and the continuous operator. We reformulate the optimization problem in terms of frequency and velocity, which allows us to weight our cost function according to the frequency content of our injected source and to the velocity distribution present in our model. We show that our method gives promising results on a constant velocity model and a constant-thickness, linearly-increasing velocity model. However, without selecting the appropriate portion of the domain on which we optimize, the error at mid-range frequencies may be increased as a trade-off for reducing the error at high frequencies. This problem has been noted in previous work but not emphasized strongly enough. In this paper, we show numerical examples demonstrating this critical point.

### INTRODUCTION

The method of finite differences (FD) is commonly used to solve the wave equation in seismic modeling, migration, and inversion. In general, FD will introduce deviations that make parts of the wavefield travel at a different velocity than the true medium velocity. This numerically-induced dispersion is more severe in the presence of high frequencies (Dablain, 1986). Reduction of the numerical dispersion can be achieved by using a finer spatial sampling or by using higher orders of approximation, both of which increase the computational cost.

A number of approaches have been taken to reduce numerical dispersion without requiring either a fine grid or higher orders. Holberg (1987) tried to correct for the error in group velocity caused by the differencing scheme. Etgen (2007) revisited the problem and corrected for phase-velocity error, taking a view that spatial and temporal dispersion could be used to compensate for each other. Zhang and Yao (2013) treated spatial dispersion separately and optimized the FD spatial derivative operator coefficients by reducing the maximum absolute error over some wavenumber range. Their approach assumes that temporal dispersion is either insignificant, possibly due to the use of a very small time step or a higher order of temporal approximation, or can be corrected separately. Correcting separately for the temporal dispersion is investigated in a companion article in this report (Li et al., 2013). We follow a path similar to that of Zhang and Yao, but parameterize our calculations in terms of velocity and frequency to take advantage of data-dependent values and limits for these physical parameters with aim of obtaining more accurate results at equivalent FD computational cost.

## OPTIMIZATION OF FD COEFFICIENTS

In the Fourier domain, the error between the finite-difference and the continuous, second-order spatial differentiation operators has the form:

$$\mathcal{E} = \mathcal{O}_{FD}(k) - \mathcal{O}_{cont}(k) = \left( \frac{1}{\Delta^2} \sum_{n=-M/2}^{M/2} a_n e^{ink\Delta} \right) - (-k^2), \quad (1)$$

where  $\Delta$  is the spatial discretization size,  $M$  is the order of approximation,  $a_n$  are constant coefficients, and  $k$  is the wavenumber. Because the same set of coefficients is used for all spatial axes, here  $k$  can be considered the wavenumber in any direction.

Following Zhang and Yao (2013), we require our FD constant coefficients to satisfy a certain set of conditions:

$$\text{symmetry:} \quad a_{-n} = a_n, \quad (2)$$

$$\text{zero mean:} \quad \sum_{n=-M/2}^{M/2} a_n = 0, \quad (3)$$

$$\text{monotonically decreasing amplitude:} \quad |a_n| > |a_{n+1}|, \quad (4)$$

$$\text{alternating signs:} \quad a_n a_{n+1} < 0. \quad (5)$$

The first two conditions, equations 2 and 3, lead to a relation for the center coefficient  $a_0 = -2 \sum_{n=1}^{M/2} a_n$ . As a result, the FD operator simplifies to

$$\mathcal{O}_{FD} = \frac{2}{\Delta^2} \sum_{n=1}^{M/2} a_n [\cos(nk\Delta) - 1]. \quad (6)$$

Another condition the FD coefficients should satisfy, which was not acknowledged in Zhang and Yao's work, comes from the fact that as the spatial step size is reduced toward zero, the FD operator should become a better and better approximation of the continuous operator. Mathematically, this means:

$$\lim_{\Delta \rightarrow 0} \frac{2}{\Delta^2} \sum_{n=1}^{M/2} a_n [\cos(nk\Delta) - 1] = -k^2, \quad (7)$$

which implies that

$$\sum_{n=1}^{M/2} n^2 a_n = 1. \quad (8)$$

The conventional FD coefficients obtained from Taylor series approximation satisfy this set of conditions. These conditions reduce the number of free coefficients to  $M/2 - 1$ .

Using the exact dispersion relation  $k = 2\pi f/v$ , we form our cost function as a weighted sum of squares of the error function (equation 1):

$$\mathcal{C} = \sum_{f,v} W_1(f) W_2(v) \left( 2 \sum_{n=1}^{M/2} a_n \left[ \cos\left(n \frac{2\pi f}{v} \Delta\right) - 1 \right] + \Delta^2 \left( \frac{2\pi f}{v} \right)^2 \right)^2, \quad (9)$$

where  $W_1(f)$  and  $W_2(v)$  are the weighting functions corresponding to the contributions of frequency and velocity in the cost function. In equation 9, the cost function has been multiplied by  $\Delta^2$  to make it dimensionless (Zhang and Yao, 2013).

In this work, we have used a simulated annealing algorithm to find the minimum of the cost function. As did Zhang and Yao, we chose to use simulated annealing because of its ability to easily handle the non-linear constraints in equations 4, 5, and 8. We end the search when the cost stabilizes at a value that is significantly smaller (by one order of magnitude) than those calculated by Zhang and Yao’s coefficients for a large number of iterations.

## RESULTS AND DISCUSSION

The parameters for our modeling examples are presented in Table 1.

Model	Velocity (m/s)	Temporal order	$\Delta t$ (ms)	Spatial order	$\Delta x$ (m)
Constant velocity	2400	4th	1	8th	5
Constant thickness	1500-3300	4th	1	8th	5

Table 1: Modeling parameters we use in our examples.

Since numerical dispersion is most problematic at high frequencies, for all of the following examples we chose a Ricker wavelet with fundamental frequency of 80 Hz as our source. The corresponding weighting function,  $W_1(f)$  in equation 9, is its amplitude spectrum. For the case of constant velocity, we applied our method with  $v = 2400$  m/s. Figure 1 shows the difference between the FD operator and the continuous operator ( $\mathcal{E}$  in equation 1) as a function of frequency for different sets of coefficients. In this example, we have experimented with different percentages of the source bandwidth on which we compute and minimize the cost function. From this figure, we observe that the conventional FD coefficients are good up to approximately 80 Hz, while with our coefficients optimized on 50% of the bandwidth (pink curve in Figure 1), the error starts to become significant at about 110 Hz. Using Zhang and Yao’s coefficients, this limit is pushed a little further to 120 Hz. As we enlarge the optimized domain, we start to introduce more error in the mid-range frequencies as a trade-off for reducing the error at the high end of the source bandwidth. This was also observed by Etgen (2007) and by Zhang and Yao (2013).

Figures 2 and 3 compare the wavefields and wavelets obtained by modeling with different sets of coefficients, for which the errors are shown in Figure 1. In both Figure 2 and Figure 3, going clockwise from the top-left quadrant, as the tailing dispersion (inside the true wavefront) is reduced, the leading dispersion (outside the true wavefront) becomes worse. This can be explained by the behavior of the error function we observed in Figure 1. The tailing dispersion is caused by the high-frequency waves, which have high wavenumber and travel more slowly, while the leading dispersion is caused by the mid-range frequency waves, which have lower wavenumber and travel faster (Dablain, 1986).

We also tested our optimization scheme on a constant-thickness 11-layer medium with velocity increasing linearly from 1500 m/s at the top to 3300 m/s at the bottom. Figures 4 and 5 compare the error between using conventional, Zhang and Yao’s coefficients, and

coefficients optimized on 100% of the source bandwidth a range of velocities. We can observe from both of these figures that the error is significantly large at high frequencies and low velocities. Figure 5b shows how the error changes with frequency for 11 velocities in our model. As the velocity changes, the position of the "error bumps" also changes. Unfortunately, as we shall see shortly, this is a red herring, that is, an artifact of the display. In fact, while we have shown there is value in *designing* the operator coefficients in the  $f - v$  domain, those coefficients remain coefficients of a single spatial finite difference operator.

Comparisons of the wavefields and wavelets for the layered-medium example are shown in Figures 6 and 7. Similar to the case of constant velocity, as the optimized domain increases, we observe the reduction of the tailing dispersion and amplification of the leading dispersion. Here we did not apply any weighting to the velocity part of the cost function,  $W_2(v) = 1$ , because the layer thicknesses were constant.

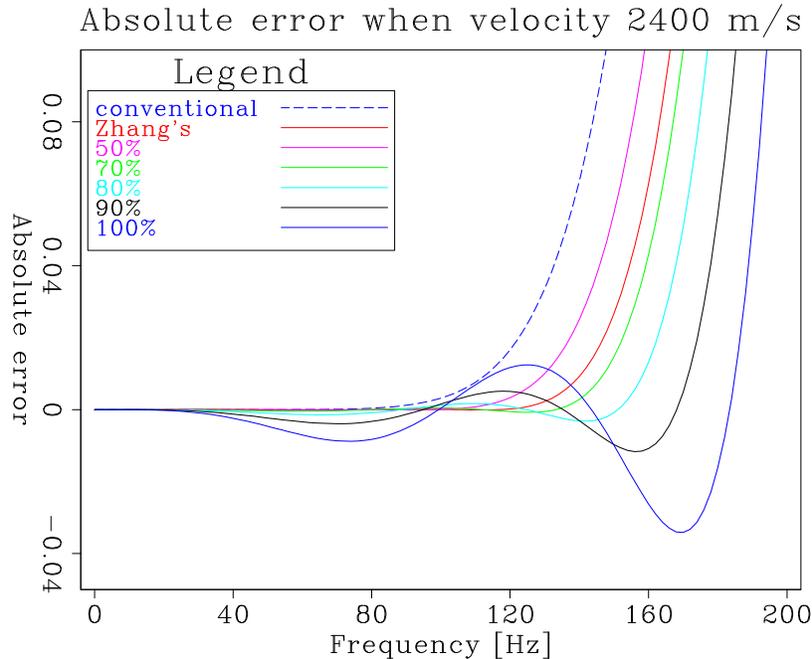


Figure 1: Error between the FD operator and the continuous operator ( $\mathcal{E}$  in equation 1) as a function of frequency for different stencils. Percent notation indicates the coverage of the bandwidth on which we try to minimize the cost function. Increasing the coverage of the bandwidth boosts the mid-range-frequency error as a trade-off for reducing the error at high frequency. [ER] huyle/. err2400

### No cancellation without temporal dispersion

One topic of discussion has been how we could achieve partial cancellation of spatial dispersion error in the presence of multiple velocities (Figure 5b); in particular, would we need to design separate spatial second derivative approximations for each of a range of velocities, or would a single approximation yield cancellation by itself.

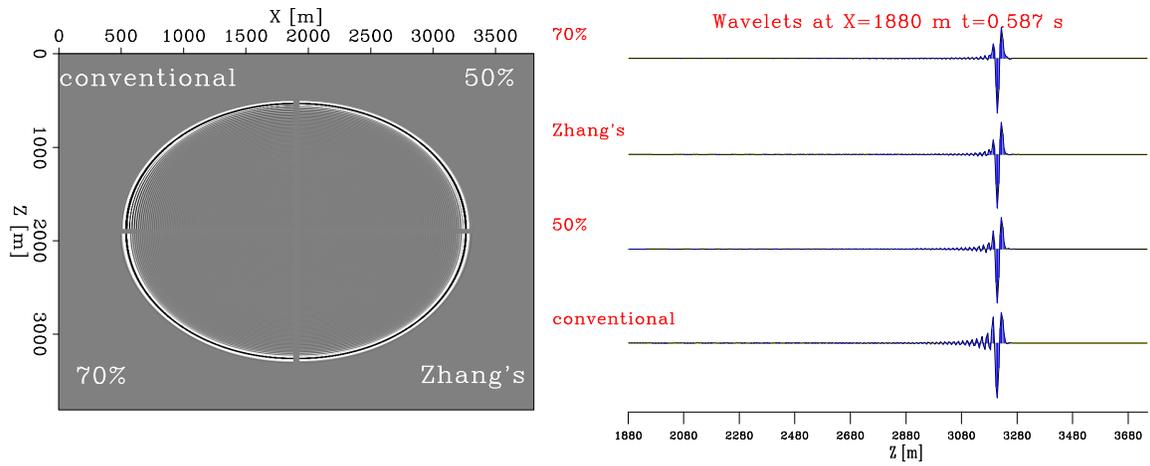


Figure 2: Comparison of wavefields (left) and of the wavelets (right) modeled with constant velocity using different sets of coefficients: conventional, Zhang and Yao's, optimized on 50%, and 70% of the bandwidth. [ER] huyle/. compare12

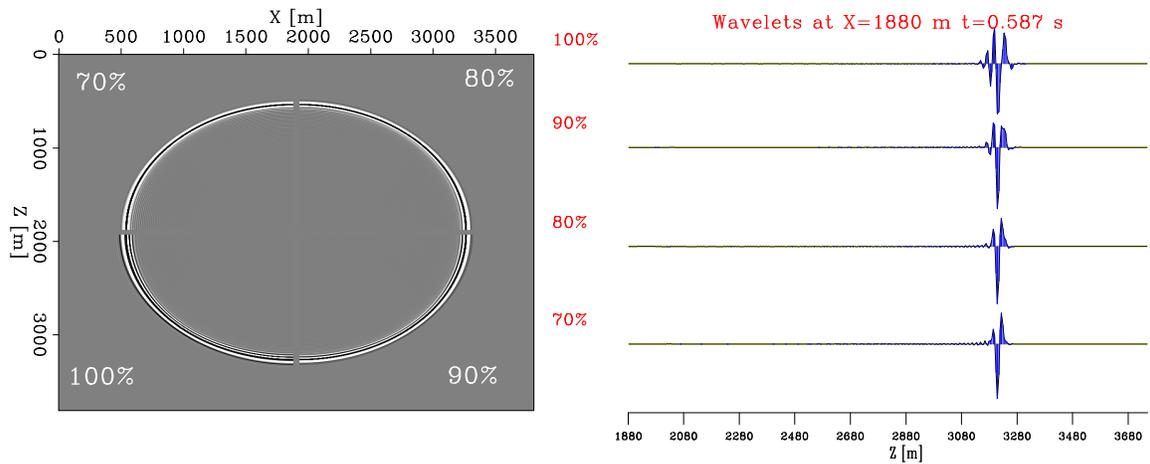


Figure 3: Comparison of wavefields (left) and of the wavelets (right) modeled with constant velocity using different sets of coefficients optimized on 70%, 80%, 90%, and 100% of the bandwidth. [ER] huyle/. compare22

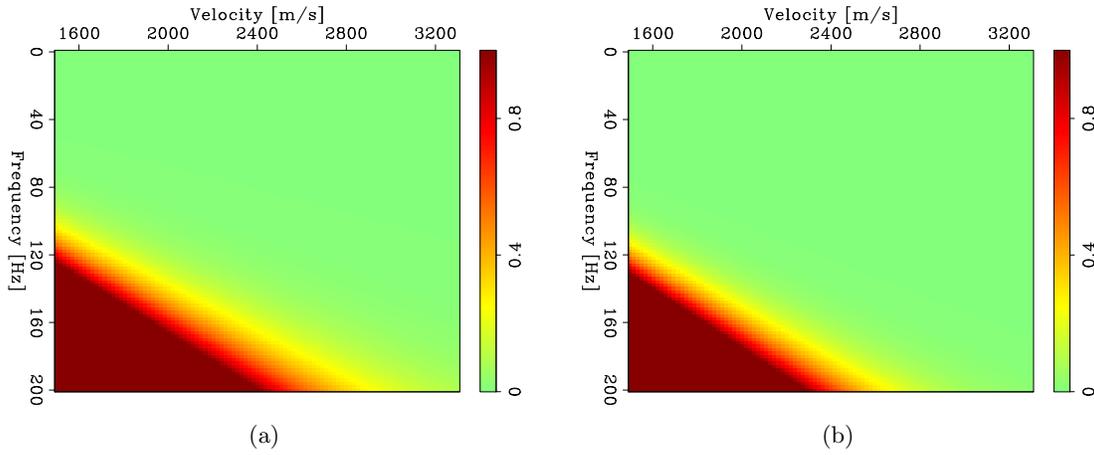


Figure 4: Error between the FD operator and the continuous operator as a function of frequency and velocity for: (a) conventional coefficients and (b) Zhang and Yao's coefficients. Carefully observed around the 120-Hz regions, Zhang and Yao's coefficients do a slightly better job than the conventional FD coefficient [ER] huyle/. conven.err,Zhang.err

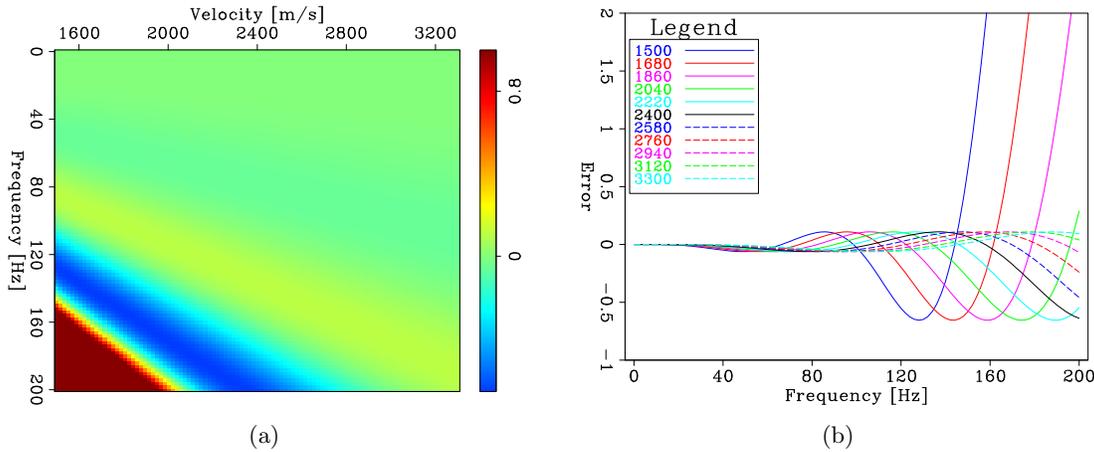


Figure 5: Error between the FD operator and the continuous operator as a function of frequency and velocity for coefficients optimized on 100% of bandwidth and: (a) a range of velocities from 1500-3300 m/s and (b) 11 velocities that are used for wavefield modeling later. Notice the movement of the "bumps" in the error plane and curves as velocity changes. [ER] huyle/. my.err,err

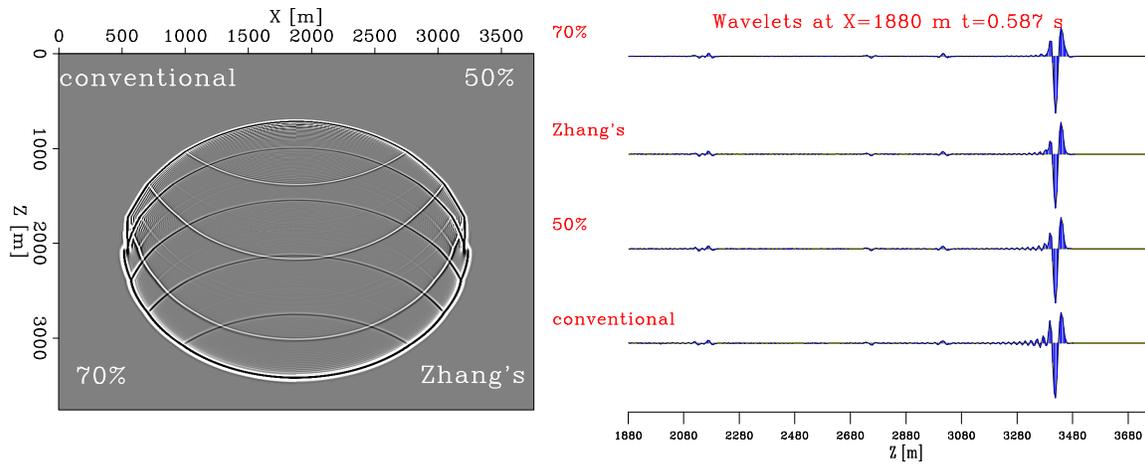


Figure 6: Comparison of wavefields (left) and of the wavelets (right) modeled in a constant-thickness layered medium using different sets of coefficients: conventional, Zhang and Yao's, optimized on 50%, and 70% of the bandwidth. [ER] huyle/. compare32

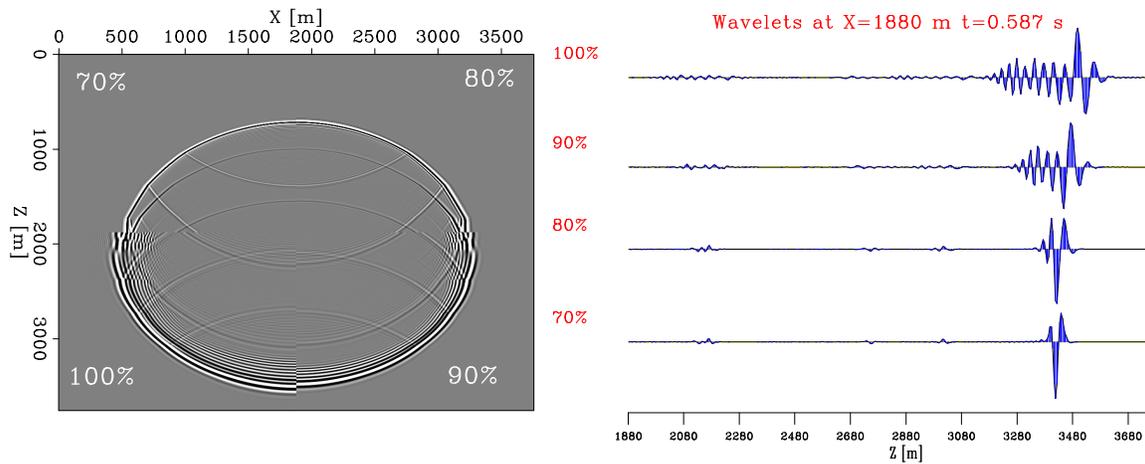


Figure 7: Comparison of wavefields (left) and of the wavelets (right) modeled in a constant-thickness layered medium using different sets of coefficients: 70%, 80%, 90%, and 100% of the bandwidth. [ER] huyle/. compare42

The key step was to understand how to separate spatial dispersion from temporal dispersion in our analysis. In essence, we asked how our wave extrapolation behaves when we treat  $t$  as a continuous variable. In this setting, we can simply replace  $\partial_t^2$  with  $-\omega^2$  in designing our spatial operator. If we let  $\kappa$  be the approximation to the true wavenumber  $k$  in our finite difference, then the wave equation we solve is

$$P_{tt} = -v^2(\kappa_x^2 + \kappa_z^2)P = -v^2|\vec{\kappa}|^2P \quad (10)$$

with solution being a linear combination of

$$e^{iv|\vec{\kappa}|t} \quad \text{and} \quad e^{-iv|\vec{\kappa}|t} .$$

Given this, taking steps of  $\Delta t_1$  at velocity  $v_1$ ,  $\Delta t_2$  at velocity  $v_2$ , etc. yields the cumulative phase shift

$$\phi_{tot} = \pm|\vec{\kappa}| \sum_i v_i \Delta t_i , \quad (11)$$

which differs from the continuous solution phase shift by the factor  $|\vec{\kappa}|/|k|$ .

*Conclusion:* The bumps and wiggles in the error of a single spatial second-derivative operator sum in phase in the absence of temporal dispersion.

### No cancellation with temporal dispersion

Having established that a single spatial second-derivative approximation would not yield phase error cancellation without temporal dispersion, we included temporal dispersion by examining the phase error as a function of one (large) time step over a range of medium velocities.

Appendix B of Le and Levin (2013) provides the appropriate framework for the analysis. Substituting the scalar operator  $-v^2|\vec{\kappa}|^2$  for  $\mathbf{L}$  in that appendix, we find that for stable time steps our eigenvalues are the complex conjugate pair

$$\lambda = \frac{2 - v^2\Delta t^2|\vec{\kappa}|^2 \pm \sqrt{(2 - v^2\Delta t^2|\vec{\kappa}|^2)^2 - 4}}{2} \quad (12)$$

on the unit circle with corresponding eigenvectors (ignoring normalization)

$$\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \lambda_2 \\ -1 \end{pmatrix} . \quad (13)$$

Looking now at the all-important phase  $\phi$  of the eigenvalues (equation 12), we have

$$\cos \phi = \frac{2 - v^2\Delta t^2|\vec{\kappa}|^2}{2} \quad (14)$$

$$\sin \phi = \pm \frac{v\Delta t|\vec{\kappa}| \sqrt{4 - v^2\Delta t^2|\vec{\kappa}|^2}}{2} . \quad (15)$$

For a practical comparison, we took the 16<sup>th</sup> order approximation of Zhang and Yao (2013) for a spatial 10-meter discretization and calculated the phase error after a time

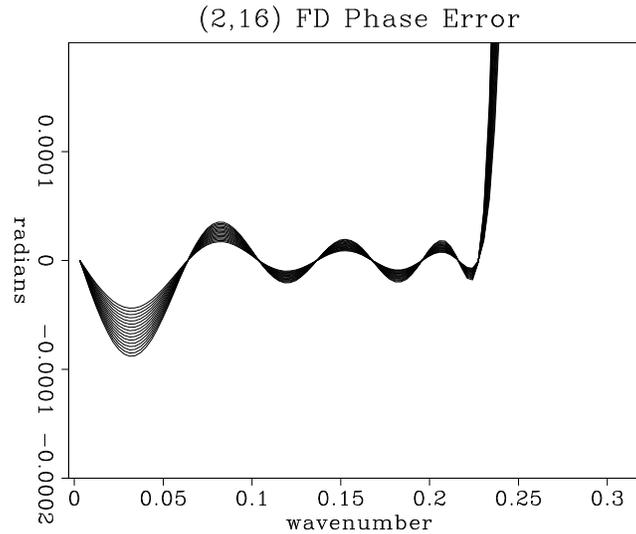


Figure 8: Finite-difference phase error after one time step of 2 ms on a 10-meter spacing grid using an explicit (2,16) finite difference scheme and varying the propagation velocity from 1500 to 3000 m/s. Quite clearly the phase errors for different velocities will not even partially cancel out. [ER] huyle/. phasediff

step of 2 ms, nearly the largest stable time step in this setting. The plot of the phase error is shown in Fig. 8. Even the most cursory examination of that result shows that the velocity-dependent errors are always in the same direction.

*Conclusion:* The bumps and wiggles in the error of a single spatial second derivative operator sum in phase in the presence of temporal dispersion.

*Conclusion:* We need to design multiple spatial operator approximations in order to obtain some cancellation of spatial dispersion as we step in time. This will be a significant focus of our future work.

## CONCLUSIONS

We have reformulated the optimization problem for the FD constant coefficients in terms of velocity and frequency. This allows us to weight the cost function according to the frequency content present in our data and the velocity distribution in our model. As a result, we gain the flexibility to design an optimal FD coefficients that are suitable for a particular imaging or inversion problem. We have demonstrated that our method works for both a constant-velocity model and a constant-thickness, linearly increasing velocity model. However, without carefully selecting an appropriate portion of the optimization domain, we may boost the low- and mid-range frequency dispersion. Using a better minimum-searching algorithm or higher-order stencils might help increase this optimization coverage. Although in this work we use the standard FD grid, our optimization design can be straightforward

to apply for more complicated FD schemes, such as staggered schemes, which have been recently used for modeling in anisotropic media (Chu, 2012).

## ACKNOWLEDGMENTS

I would like to thank my colleagues, especially Elita (Yunyue) Li, Mandy Wong, and Ohad Barak, and professors at the Stanford Exploration Project for their great help and valuable advice and discussions.

## REFERENCES

- Chu, C., 2012, A hybrid finite difference method for acoustic wave propagation in tilted orthorhombic media: 82nd SEG Ann. Internat. Meeting, Expanded Abstracts, 1–5, Soc. of Expl. Geophys.
- Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: *Geophysics*, **51**, 54–66.
- Le, H. and S. A. Levin, 2013, Practically stable unstable orthorhombic finite differences: SEP-Report, **149**, 243–250.
- Li, Y. E., M. Wong, and R. Clapp, 2013, Equivalent accuracy at a fraction of the cost: Overcoming temporal dispersion: SEP-Report, **150**, 139–148.
- Zhang, J.-H. and Z.-X. Yao, 2013, Optimized finite-difference operator for broadband seismic wave modeling: *Geophysics*, **78**, A13–A18.

## Equivalent accuracy at a fraction of the cost: Overcoming temporal dispersion

*Yunyue (Elita) Li, Mandy Wong, and Robert Clapp*

### ABSTRACT

Numerical dispersion in finite difference modeling produces coherent artifacts, severely constraining the resolution of advanced imaging and inversion schemes. Conventionally, we deal with this by increasing the order of accuracy of the finite difference operators and resign ourselves to paying the high computational cost that incurs. But is there a way to reduce such dispersion without increasing cost or, conversely, decrease the cost without increasing numerical dispersion? To tackle this, we separate the finite difference numerical dispersion into pure time and pure space dispersion and address them independently. In this article, we focus on time dispersion. We show that finite difference time dispersion is virtually independent of the medium velocity and the spatial grid for propagation, and only depends on the time stepping scheme and the propagation time. Based on this, we devise post-propagation filtering to collapse the time dispersion effect of the finite difference modeling. Our dispersion correction filters are designed by comparing the input waveform with dispersive waveforms obtained by 1D propagation of that waveform forward in time. These filters are then applied on multi-dimensional shot records to eliminate the time dispersion by two schemes: (1) stationary filtering plus interpolation and (2) non-stationary filtering. We show with both 1D and 2D examples that the time dispersion is effectively removed by our post-propagation filtering at nearly no additional cost.

### INTRODUCTION

Finite difference (FD) modeling for wave propagation has been widely used for advanced imaging techniques such as least squares reverse time migration (Lambare et al., 1992; Nemeth et al., 1999; Dai et al., 2010; Wong et al., 2011) and inversion schemes such as waveform impedance inversion (Kelly et al., 2010; Plessix and Li, 2013) and full waveform inversion (Tarantola, 1987; Virieux and Operto, 2009). In these methods, wavefields modeled by finite difference are compared to recorded data which, of course, contain no computational numerical dispersion. As a result, a great deal of the cost of these methods arises from the effort to make these comparisons meaningful by reducing finite difference dispersion.

Much effort has been put into suppressing the numerical dispersion of the FD methods. Kosloff and Baysal (1982) used the spatial Fourier transform to eliminate all errors from FD approximation of the spatial derivatives, and chose a small enough time step to limit the numerical time dispersion. However, this method requires many Fourier transforms at each time step and also a rather small time step, and hence has been considered prohibitively expensive in practice. Therefore, many authors (Holberg, 1987; Fornberg, 1998; Etgen,

2007) set up an optimization problem to generate FD coefficients that minimize the misfit between the numerical phase velocity and the theoretical phase velocity for some range of frequency and velocity. Nonetheless, with computational cost controlled by the number of coefficients that are optimized, achieving both accuracy and efficiency at the same time is still challenging.

Recently, Stork (2013) proposed to separate the temporal and the spatial FD dispersion. Here we follow this approach, albeit only tackling the temporal dispersion in this article. For discussion of the spatial dispersion component, please refer to a companion paper Le et al. (2013). In this paper, we first analyze the source of the time dispersion and show that the time dispersion is independent of the spatial sampling and the velocity of the medium. To eliminate the space error associated with propagation we use a Fourier modeling method and propagate a wavelet through an arbitrary medium. Then we design the filters using 1D Fourier modeling results at discrete propagation times. These filters are applied to both 1D and 2D shot records to remove the time dispersion. We test two different filtration schemes: (1) stationary filtering plus interpolation and (2) non-stationary filtering. The results show that both filtration schemes can eliminate the time dispersion on a shot record with nearly no additional cost.

## THEORY

Assuming constant density and a source free medium, the acoustic wave equation may be written

$$c^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)P = \frac{\partial^2}{\partial t^2}P, \quad (1)$$

where  $P$  is the pressure field and  $c(x, y)$  is the velocity.

Both sides of equation 1 are approximated by numerical discretization: the left-hand side in space and the right-hand side in time. Using notation similar to Kosloff and Baysal (1982), the finite difference equation we are solving is

$$c^2LP^n(i, j) = \frac{1}{\Delta t^2} [P^{n+1}(i, j) - 2P^n(i, j) + P^{n-1}(i, j)], \quad (2)$$

where  $P^n(i, j)$  represents the value of the pressure field at time  $t = n\Delta t$  at spatial location  $x = x_0 + (i - 1)\Delta x$ ,  $y = y_0 + (j - 1)\Delta y$ . The term  $c^2LP^n(i, j)$  represents the numerical approximation of the left-hand side.

Equation 2 represents an explicit, second-order time differencing scheme that is widely used in finite difference codes. Both the left-hand and the right-hand side of equation 2 contain errors with respect to equation 1. We call the error from the left-hand side approximation spatial dispersion, and from the right-hand side temporal dispersion.

Conventionally, temporal dispersion can be reduced either by decreasing  $\Delta t$  or increasing the order of the finite difference (Figure 1). Both options significantly increase the computational cost. Stork (2013) suggests that temporal dispersion is independent of the medium velocity and the spatial grid for propagation. We verify this idea by numerical tests. Figure 2(a) shows the 1-D modeling results at different propagation times when the second-order time stepping is used. The spatial derivative is computed in the Fourier space to avoid any spatial dispersions. In this example,  $\Delta t = 2\text{ms}$ ,  $c = 2000\text{m/s}$ ,  $\Delta x = 10\text{m}$ . It is clear that

the dispersion effects get stronger with the longer propagation time. Figure 2(b) shows the same modeling results as in figure 2(b), only with  $c = 1000\text{m/s}$ ,  $\Delta x = 6\text{m}$ . Despite the differences in the spatial sampling and the medium velocity, the temporal dispersion remains the same as long as the  $\Delta t$  and the propagation time are the same. Therefore, inverse filters can be designed to remove the velocity-independent dispersions after propagation.

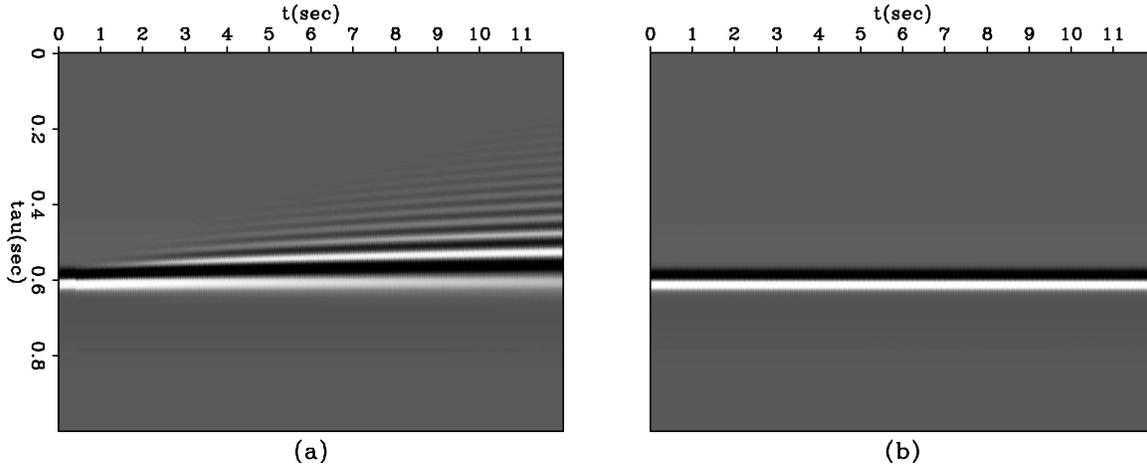


Figure 1: Wavefield from 1-D modeling using the Fourier method with (a) second-order time stepping and (b) fourth-order stepping. Time step is 2ms in both cases. Severe time dispersions are removed at twice the computation cost. [ER] `elita2/. showboth`

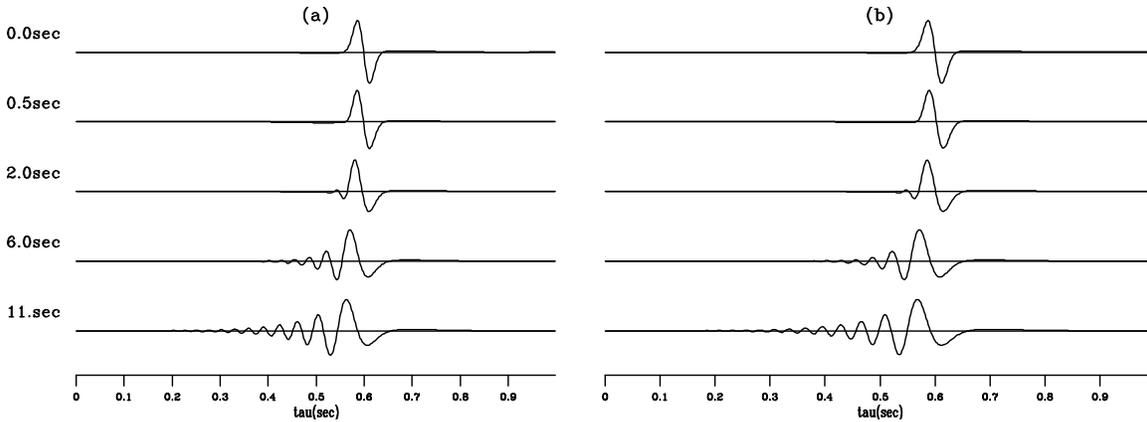


Figure 2: 1-D modeling results by Fourier method with different model parameters. In (a)  $c = 2000\text{m/s}$ ,  $\Delta x = 10\text{m}$ . In (b)  $c = 1000\text{m/s}$ ,  $\Delta x = 6\text{m}$ . The dispersion effects are the same. [ER] `elita2/. dispindp`

We estimate the inverse filters  $f(\tau, t)$  by comparing the waveform  $s(\tau, t)$  with the original waveform  $s(\tau, 0)$  in the Fourier space:

$$F(\omega, t) = \frac{S(\omega, 0)}{S(\omega, t) + \epsilon}, \tag{3}$$

where  $F(\omega, t)$  and  $S(\omega, t)$  are the Fourier representation of  $f(\tau, t)$  and  $s(\tau, t)$ , respectively. A small number  $\epsilon$  is added to stabilize the division. The final inverse filters are band-limited

within the frequency range of the original waveform. We only estimate the inverse filters at discrete propagation times. To apply these discrete filters on 2-D or 3-D continuous data record, we can choose from the following two schemes: stationary filtering plus interpolation, or non-stationary filtering.

### Stationary filtering plus interpolation (SFPI)

Given a record  $d(t, x)$  and inverse filters  $f(\tau, t_i)$ , we can first apply each of the inverse filters to the whole record:

$$d_i(t, x) = d(t, x) * f(\tau, t_i); \quad (4)$$

this is a trace-by-trace operation. On each filtered record  $d_i(t, x)$ , only the waveforms around  $t = t_i$  are correctly filtered. The other parts of the record are either over or under compensated.

We then interpolate among the filtered records to obtain the temporal dispersion free record  $\hat{d}(t, x)$ :

$$\hat{d}(t, x) = \sum_i h_i d_i(t, x), \quad (5)$$

where  $h(i)$  are the interpolation weights for each filtered records. In this paper, we use simple linear interpolation weights.

### Non-stationary filtering (NSF)

In a non-stationary filtering scheme, we use a moving window, which has the same length as the filter, to select the data patch to convolve with the filter defined at the center of the moving window. We overlap the moving windows to ensure smooth transitions across the data patch. Mathematically, the filtering process can be formulated as follows:

$$\hat{d}(t, x) = \frac{1}{N} \sum_i (W_i d(t, x)) * f(\tau, t_i^w), \quad (6)$$

where  $W_i$  is the  $i^{th}$  window acting on the data record,  $t_i^w$  is the center of this window, and  $N$  is the number of overlaps before moving to a new data patch.

Assuming the dispersion varies smoothly in time, we can build the filters at any propagation time by interpolating the estimated inverse filters as follows:

$$f(\tau, t_i^w) = \sum_i k(i) f(\tau, t_i), \quad (7)$$

where  $k$  is the set of interpolation weights. The same linear interpolation scheme is used.

More advanced non-stationary filtering schemes (Margrave, 1997; Fomel, 2009) can be adapted in practice; however, in the examples we present in this paper, this simple patching method yields satisfactory results.

## EXAMPLES

We test the proposed post-propagation filtering scheme on both 1-D and 2-D examples. All spatial derivatives in the numerical modeling are performed in the Fourier space to avoid the spatial dispersion. Figures 3(a) and 3(b) show the second-order and the fourth-order modeling results with  $\Delta x = 10\text{m}$  and  $c = 2000\text{m/s}$ . Dispersion is eliminated at twice the cost of second-order time stepping. Figure 3(c) and (d) show the second-order time stepping results after dispersion correction by SFPI and NSF, respectively. After the filtering, dispersion is eliminated at nearly no additional cost.

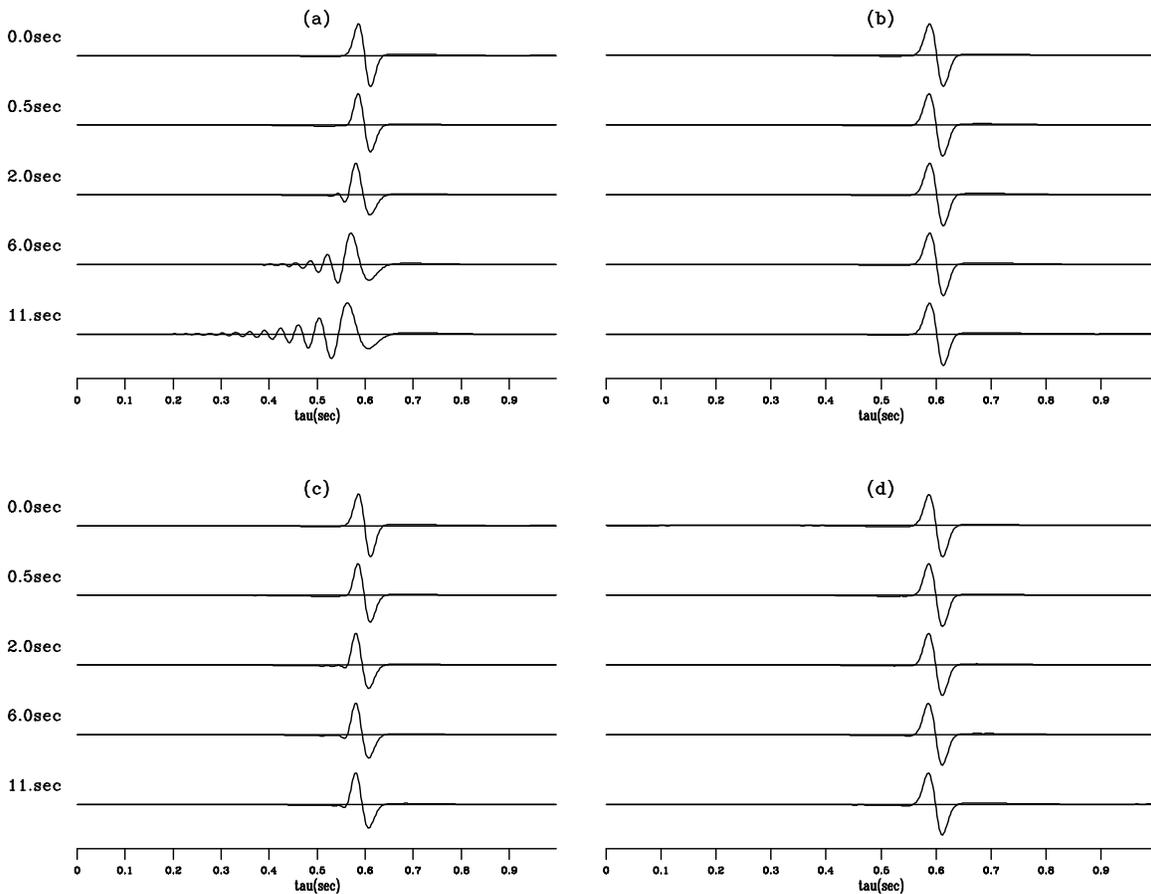


Figure 3: One-D modeling and filtering results in TK domain with 10m spacing and 2000m/s velocity. Waveforms are shown after being propagated for 0sec, 0.5sec, 2sec, 6sec and 11sec. (a): second-order time stepping results. Leading time dispersion gets greater with propagation time. (b): fourth-order time stepping results. Dispersion is eliminated at twice the cost of second-order time stepping. (c): second-order time stepping results after dispersion correction by SFPI. (d): second-order time stepping results after dispersion correction by NSF. In (c) and (d), dispersion is eliminated at nearly no additional cost. [ER]

**elita2/. 1D10mTK**

Figure 4 repeats the same modeling and filtering process with a different spatial sampling (6m) and a different velocity (1000m/s). Figure 4(a) shows the same dispersion effects as figure 3(a). We use the inversion filters estimated from 3(a) to process the data. Figures

4(c) and (d) show that the dispersion correction filters are effective as long as the time step  $\Delta t$  and the propagation time remain the same.

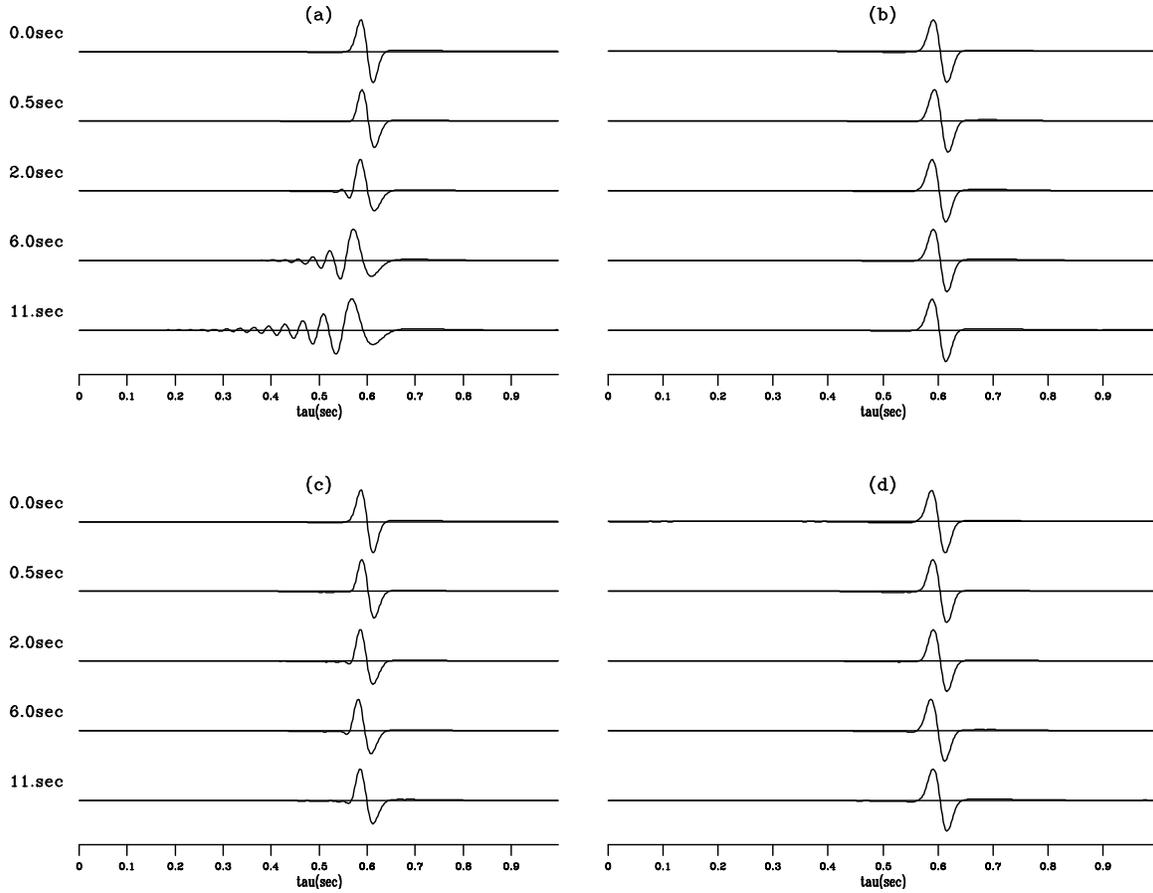


Figure 4: One-D modeling and filtering results in TK domain with 6m spacing and 1000m/s velocity. Waveforms are shown after being propagated for 0sec, 0.5sec, 2sec, 6sec and 11sec. (a): second-order time stepping results. Dispersion gets greater with propagation time. The dispersion effect is the same as in figure 3(a). (b): fourth-order time stepping results. Dispersion is eliminated at twice the cost of second-order time stepping. (c): second-order time stepping results after dispersion correction by SFPI. (d): second-order time stepping results after dispersion correction by NSF. In (c) and (d), the same set of filters applied in figures 3(c) and (d) are applied here. [ER] elita2/. 1D6mTK

Figure 5 shows the modeling and filtering results on a 2D record. Panels (c) and (d) show that both filtering schemes can remove the dispersion effects on the shot record without increasing the FD order in time. Figure 6 shows a zoom-in view of the shot record in figure 5. All the parallel events leading the main lobe are temporal dispersions. The post-propagation filterings remove them all and perfectly match the fourth-order modeling results.

Figure 7 shows the later waveform at receiver  $x = 2000\text{m}$ . The wave packet between 9s and 9.5s is severely dispersed so that any waveform-based inversion scheme would fail. Post-propagation filterings restore the phase of the waveform at no additional cost.

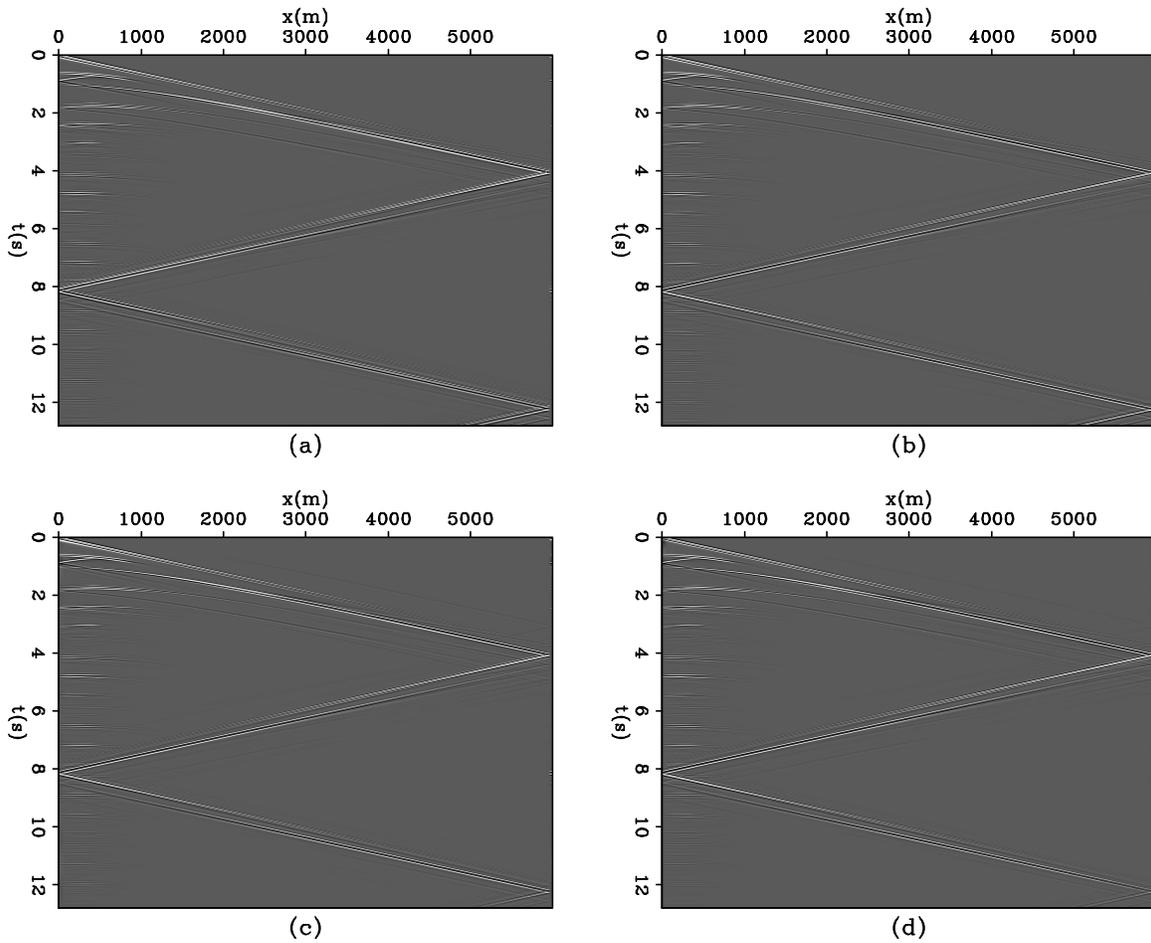


Figure 5: Shot record by two-dimensional TK domain modeling. (a) Shot record with 2nd-order modeling. Time dispersion gets greater with time. (b) Shot record with 4th-order modeling. Dispersion was eliminated by twice the cost of 2nd-order time stepping. (c) Shot record (a) after dispersion correction by SFPI. (d) Shot record (a) after dispersion correction by NSF. [ER] `elita2/. 2D10mTK`

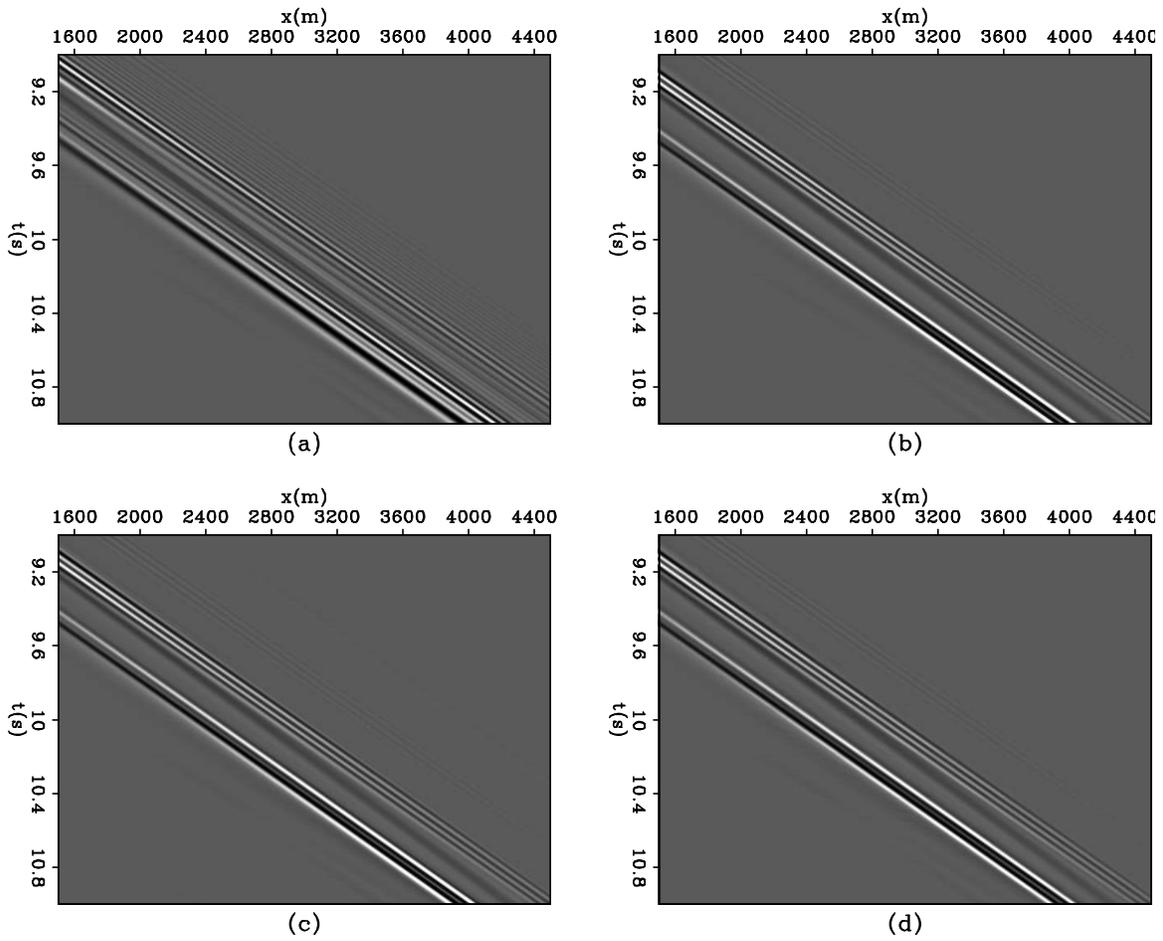


Figure 6: Zoom-in view of a shot record by two-dimensional TK domain modeling. (a) Shot record with 2nd-order modeling. Time dispersion gets greater with time. (b) Shot record with 4th-order modeling. Dispersion was eliminated at twice the cost of 2nd-order time stepping. (c) Shot record (a) after dispersion correction by SFPI. (d) Shot record (a) after dispersion correction by NSF. [ER] `elita2/. 2Dzoom10mTK`

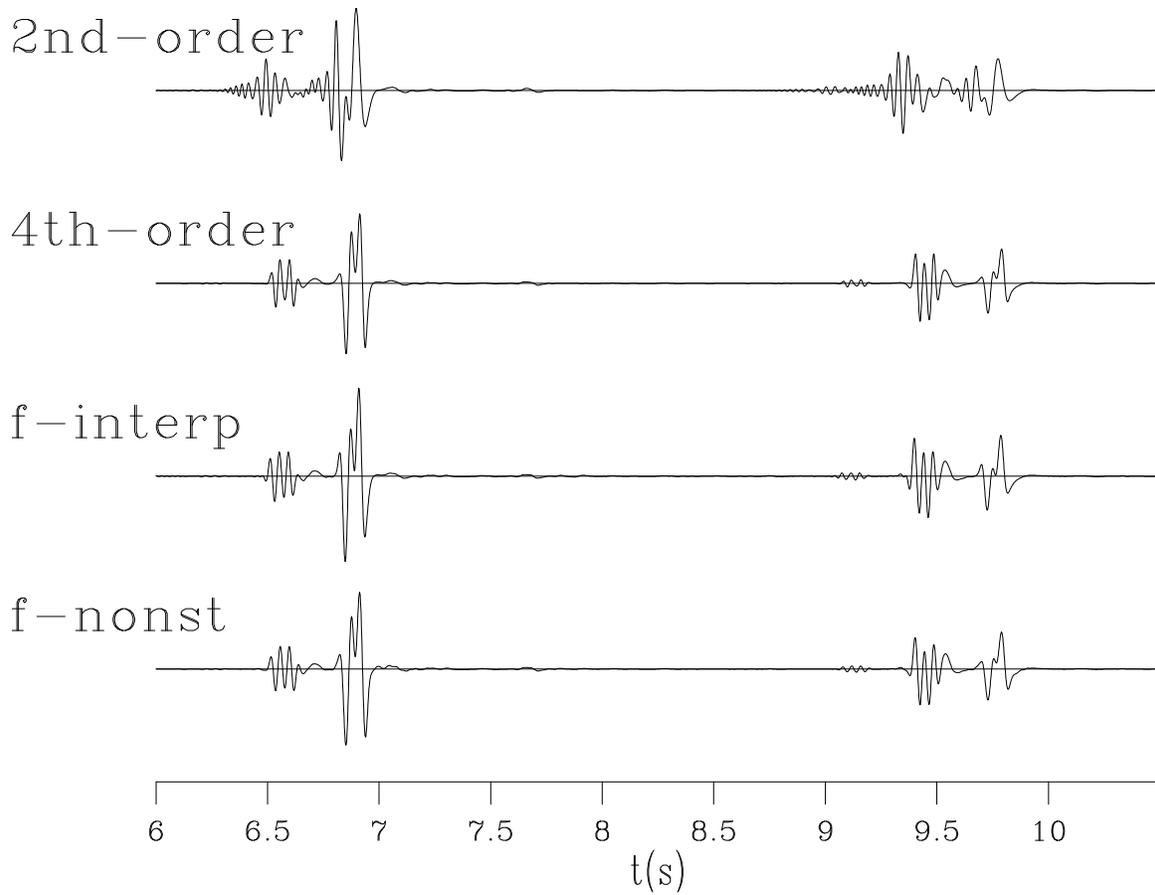


Figure 7: Full waveform at receiver  $x = 2000\text{m}$  by two-dimensional Fourier domain modeling. Dispersion gets greater in time in the 2nd-order time stepping. The energy packet between 9s and 9.5s is severely distorted by dispersion in the 2nd-order modeling result. Post-propagation filtering can remove the dispersion at nearly no additional cost. [ER]

`elita2/. 2D10mTKgraph`

## CONCLUSIONS

We have developed two post-propagation filtering schemes to remove the temporal dispersion caused by the inaccuracy of the second-order FD approximation to the time derivatives when solving the wave equation. We show that the temporal dispersion is independent of the medium velocity and spatial sampling, which is the reason why post-propagation filtering is possible. We design two different filtering schemes: stationary filtering plus interpolation and non-stationary filtering. The filtering results on both 1-D and 2-D second-order modeling show that we can successfully remove the dispersion artifacts at nearly no additional cost.

Although the filtering results of both schemes are similar, the SFPI scheme involves many (the number of filters) passes of convolution across the data record, whereas the NSF scheme involves only one convolutional pass. Furthermore, the SFPI scheme interpolates among multiple copies of the data record, whereas the NSF scheme interpolates among the filter coefficients. Therefore, considering both the computation and memory requirements, the non-stationary filtering scheme is more suitable for large scale computations.

## REFERENCES

- Dai, W., C. Boonyasiriwat, and G. T. Schuster, 2010, 3D multi-source least-squares reverse time migration: SEG Expanded Abstracts, 3120–3124.
- Etgen, J., 2007, A tutorial on optimizing time domain finite-difference schemes: “Beyond Holberg”: SEP report, **129**, 33–43.
- Fomel, S., 2009, Adaptive multiple subtraction using regularized nonstationary regression: Geophysics, **74**.
- Fornberg, B., 1998, Calculation of weights in finite difference formulas: SIAM Review, **40**, 685–691.
- Holberg, O., 1987, Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena: Geophysical Prospecting, **35**, 629–655.
- Kelly, S., J. Ramos-Martinez, B. Tsimelzon, and S. Crawley, 2010, Application of an impedance-based full-waveform inversion method for dual-sensor single streamer field recordings: 72nd EAGE Conference & Exhibition Extended Abstract A020.
- Kosloff, D. D. and E. Baysal, 1982, Forward modeling by a Fourier method: Geophysics, **47**, 1402–1412.
- Lambare, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.
- Le, H., R. G. Clapp, and S. A. Levin, 2013, Equivalent accuracy at a fraction of the cost: Tackling spatial dispersion: SEP-Report, **150**, 129–138.
- Margrave, G., 1997, Nonstationary filtering: review and update: CREWES Research Report, **9**.
- Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.
- Plessix, R. E. and Y. Li, 2013, Waveform acoustic impedance inversion with spectral shaping: Geophysical Journal International.
- Stork, C., 2013, Eliminating dispersion from FD modeling is crucial for inversion: SEAM Workshop: Geoscience Advances from the use of SEAM Data.

- Tarantola, A., 1987, *Inverse problem theory: Methods for data fitting and model parameter estimation*: Elsevier Science Publication Company, Inc.
- Virieux, J. and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **64**, WCC1–WCC26.
- Wong, M., S. Ronen, and B. Biondi, 2011, Least-squares reverse time migration/inversion for ocean bottom data: A case study: *SEG Expanded Abstracts*, **30**, 2369–2373.



# Stochastic rock physics modeling for seismic anisotropy with two different shale models

*Yunyue (Elita) Li, Dave Nichols, and Gary Mavko*

## ABSTRACT

Anisotropic model building using surface seismic data is a well-known underdetermined and nonlinear problem. To stabilize the inversion, we need a priori anisotropic model to precondition the model space. In this paper, we build the anisotropic prior model using stochastic rock physics modeling for sandy-shale anisotropy. We investigate two different methodologies to combine sand (quartz) and shale (clay): suspension model and lamination model. Anisotropic differential effective medium model is used to model the quartz suspension, and Backus average model is used to model the sand/shale lamination. The modeling results from both methodologies show greater differences for  $\delta$  than for  $\epsilon$ . By taking compaction and mineral transition into account, we then perform more realistic modeling at a well location where the shale content and porosity are available from the well log measurements. Both the deterministic and the stochastic model results from these two approach have similar trend but different span over the  $\epsilon$ - $\delta$  space. The combined distribution will provide looser constraints to the anisotropic parameter estimation.

## INTRODUCTION

Anisotropic model building tries to resolve more than one parameter at each grid point of the subsurface. This number could be three for a vertical transverse isotropic (VTI) media, and increases to five for a tilted transverse isotropic (TTI) media. Any inversion scheme based on surface seismic data only becomes ill-posed and highly underdetermined due to the rapidly increasing model space with the increasing complexity of the subsurface (Bakulin et al., 2009, 2010b,a).

To constrain the multi-parameter inversion, a local cross-parameter covariance is needed to better describe the subsurface (Li et al., 2011; Yang et al., 2012). One source of the cross-parameter covariance comes from rock physics studies (Hornby et al., 1995; Sayers, 2004, 2010; Bachrach, 2010b). Many authors (Dræge et al., 2006; Bandyopadhyay, 2009; Bachrach, 2010a) have built averaged depth trends serving for seismic processing. In particular, Bachrach (2010a) develops both deterministic and stochastic modeling schemes based on the rock physics effective media models for compacting shale and sandy shale. When building the rock from the minerals, Bachrach (2010a) made the choice of using the suspension model, where sand (quartz) are modeled as spherical inclusions inside the background shale.

Li et al. (2013) build spatially varying (deterministic and stochastic) anisotropic models using Backus average (Backus, 1962) for sand and shale laminated medium. Both suspension and lamination scenarios are common in a sedimentary basin. In this paper, we

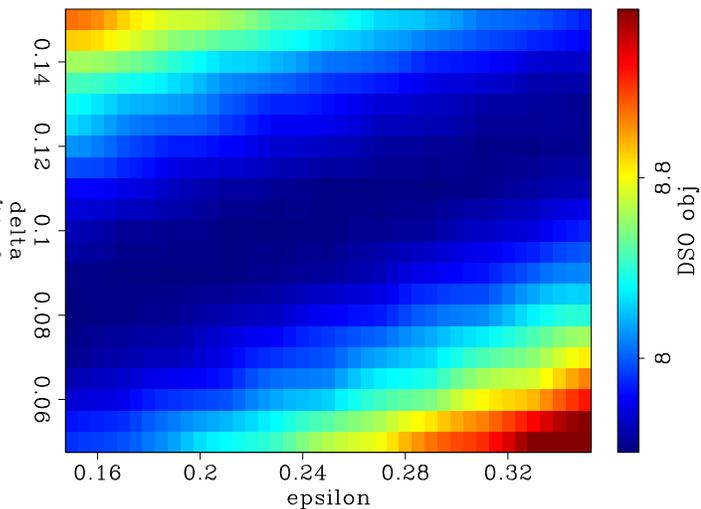
first investigate the differences between these two methodologies by comparing the elastic properties of the dry rock, assuming clay (smectite) particles are perfectly aligned. We use different end-member properties for clay and show the modeling response to them. In a more realistic test, we include the compaction and the mineral transition effect using well log measurements from the Gulf of Mexico. By varying the key parameters for rock physics modeling, we generate the stochastic distributions for the anisotropic parameters from both methodologies. The stochastic results have similar variance but different mean value. Therefore, a combined distribution is necessary to include both geological scenarios when constraining the anisotropic parameter estimation.

## WAVE-EQUATION MIGRATION VELOCITY ANALYSIS FOR ANISOTROPY

Anisotropic wave-equation migration velocity analysis (WEMVA) aims at building an anisotropic Earth model that minimizes the residual image from the surface seismic data (Li and Biondi, 2011). One of the most commonly used objective function in WEMVA is the differential semblance optimization (DSO) (Shen, 2004) objective function. It relates the unfocused energy in the subsurface-offset domain common image gathers to the inaccuracy in the subsurface models.

Figure 1: Modeled topography of the DSO objective function using a 1.5D single layer synthetic example.

[ER] `elita1/.objed`



We test the DSO objective function on a 1.5D single layer synthetic model. The true anisotropic model is  $\epsilon = 0.25$  and  $\delta = 0.1$ . We plot in figure 1 the topography of the DSO objective function while  $\epsilon$  and  $\delta$  are perturbed within the range from  $-50\%$  to  $50\%$ . In general, the DSO objective function has a better resolution in the direction of  $\epsilon - \delta$  than in  $\epsilon + \delta$ . Nonetheless, the flat-bottom of the topography indicates a very low resolution of the DSO objective function to the anisotropic parameters.

To constrain the null space and stabilize the inversion, a regularization term is needed in addition to the anisotropic WEMVA objective function. The topography of the model regularization objective function can be estimated by stochastic rock physics modeling. We will discuss the process of the rock physics modeling in detail in the next section.

## TWO MODELS FOR SHALE ANISOTROPY

Two workflows to model shale anisotropy has been proposed by Bachrach (2010a) and Li et al. (2013). These two workflows are similar to each other, except for the last step: Bachrach (2010a) models quartz as inclusions in the clay background, and Li et al. (2013) model quartz and clay as a laminated system.

To analyze the difference between these two methodology, we first study the elastic property of the dry-rock with pure clay and pure quartz, ignoring other geological and mineralogical effects. Since pure clay is very fragile, the elastic property of clay mineral is very difficult to measure and hence has very high uncertainty. Therefore, we repeat the rock physics modeling using three sets of elastic properties for clay: isotropic clay, weakly anisotropic clay and strongly anisotropic clay. Quartz is considered isotropic in all three tests. To model quartz as an inclusion in the clay background, we use the anisotropic differential effective medium (DEM) (Bandyopadhyay, 2009) method. To model the fine layering of clay and quartz, we use the Backus averaging method (Backus, 1962).

The rock physics modeling results with respect to the quartz content are shown in figure 2. In each plot, the blue curve show the modeling result by lamination model and the red curve by inclusion model. The left column shows the  $\epsilon$  model, and the right shows the  $\delta$  model. The top, middle and bottom rows show the modeling result assuming isotropic clay, weakly anisotropic clay and strongly anisotropic clay, respectively.

In both modeling schemes, anisotropy of the rock decreases with the increasing amount of isotropic quartz in the rock. However, the lamination model predicts different apexes for  $\epsilon$  and  $\delta$  when clay is isotropic. It also predicts negative  $\delta$  values when clay is weakly anisotropic. These predictions will point to different correlation directions in the stochastic modeling results. From figure 2, we can see that  $\epsilon$  estimates from both modeling results are very close to each other except when clay is assumed isotropic. However, the estimates for  $\delta$  are significantly different from each other in all three cases. These differences show the value of complementing one model with another to include more possible geological scenarios.

### Rock physics modeling using well log inputs

The following shows the workflow we adapt to model the anisotropy at a well location.

- Compute the percentage of illite in the rock given a temperature model.
- Compute the average stiffness coefficients over a orientation distribution for smectite and illite, given a porosity model.
- Compute the volumetric percentage for each of the mineral phase, given a volumetric percentage of shale.
- Compute the stiffness coefficients for the inclusion model or the lamination model.

Figure 3 shows one instance of the modeling results by both modeling schemes. Figure 3(a) and 3(c) shows the  $\epsilon$  and  $\delta$  model using the lamination model. Figure 3(b) and 3(d)

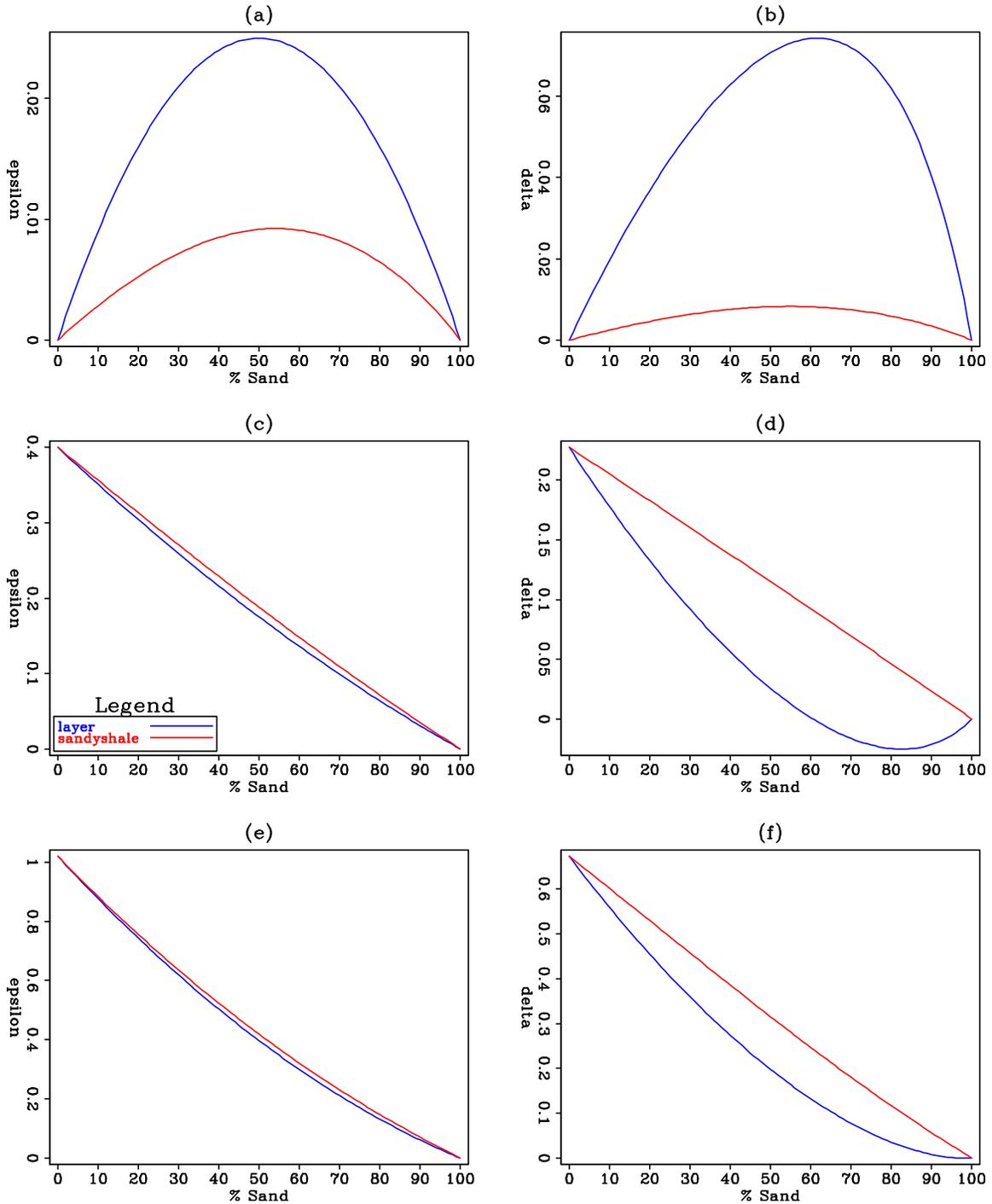


Figure 2: Thomsen parameters  $\epsilon$  (left column) and  $\delta$  (right column) modeled by two methods. Blue curve denotes the layering model. Red curve denotes the inclusion model. Clay is isotropic in (a) and (b), weakly anisotropic in (c) and (d), strongly anisotropic in (e) and (f). [ER] elita1/. tstmodel

shows the  $\epsilon$  and  $\delta$  model using the inclusion model. In general, anisotropy predicted by both models correlated with the shale content in the well log. Due to the Backus averaging effect, anisotropic profiles from the inclusion model contains more frequency content towards the higher end. We also observe greater differences in  $\delta$  than in  $\epsilon$ .

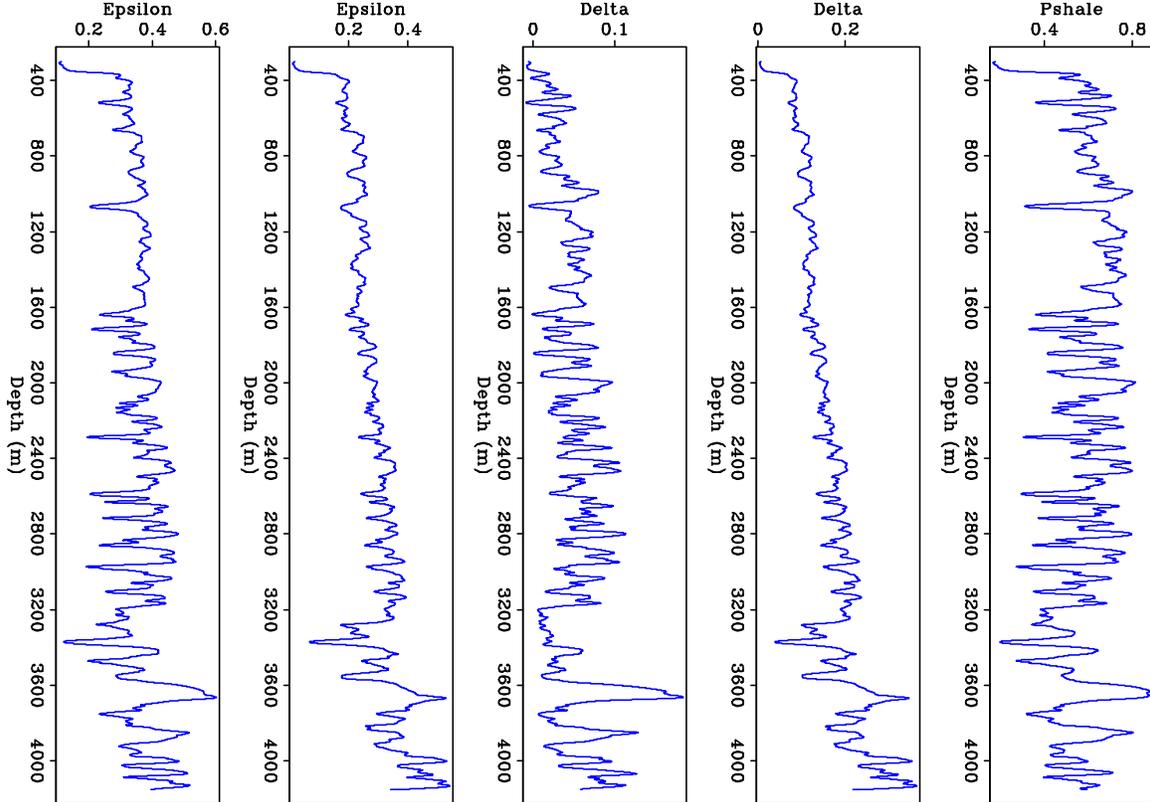


Figure 3: One instance of the rock physics modeling experiment. From left to right, panels are  $\epsilon$  profile from the lamination model and inclusion model;  $\delta$  profile from the lamination model and inclusion model; Shale content at the well derived from the Gamma-ray measurements. [ER] `elita1/. cmpmodel`

Finally, we vary the key parameters in the rock physics modeling workflow (Li et al., 2013) to approximate the distribution of the anisotropic parameters at the well location. Figure 4 shows that the resulting distributions from both modeling are similar in shape, but have different span in the  $\epsilon$ - $\delta$  space. The combined distribution of these two models allows larger variations in both  $\epsilon$  and  $\delta$ .

## CONCLUSIONS AND DISCUSSION

In this paper, we study the topography of the image-space DSO objective function with respect to anisotropic parameters. We show that due to the lack of constraints on the anisotropic parameters, other sources of information are needed to regularize the inversion.

We compare two rock physics modeling scheme to combine clay and quartz minerals: the inclusion model and the lamination model. The modeling responses show larger difference in  $\delta$  than in  $\epsilon$ . In a more realistic test, we add the compaction and mineral transition effects

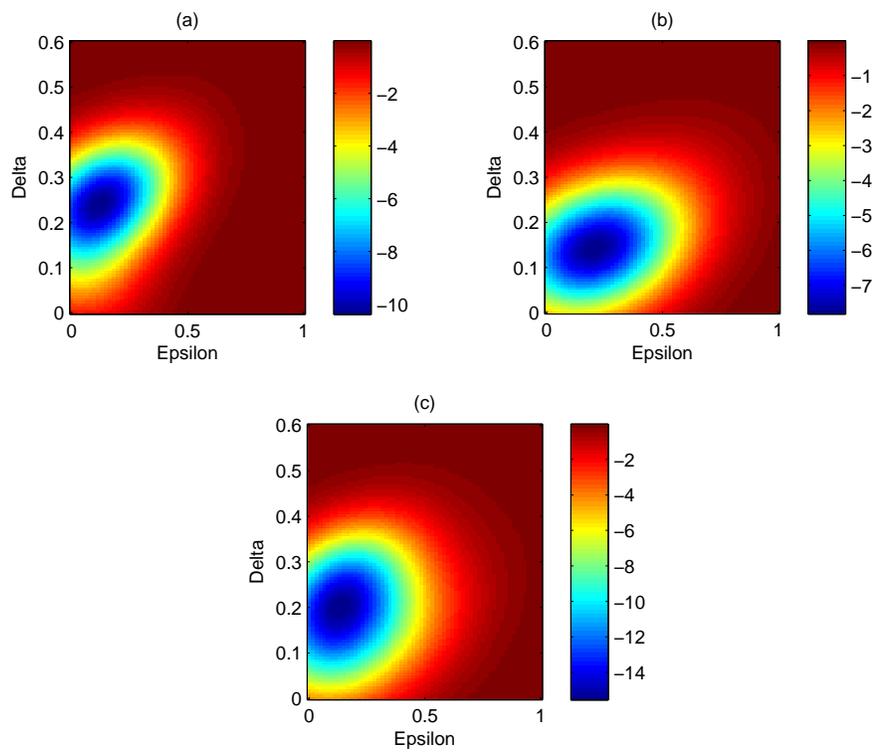


Figure 4: Stochastic rock physics modeling results for anisotropy. [ER] elita1/. combinepdf

prior to combing the clay and quartz minerals. Both modeling results show high correlation between predicted anisotropy with the shale content. By stochastic rock physics modeling, we show that the resulting distributions from both modeling are similar in shape, but have different span in the  $\epsilon$ - $\delta$  space. The combined distribution leads to looser constraints on  $\epsilon$  and  $\delta$ .

Finally, it is worth noting that the DEM method is valid when the quartz content is between 0% and 60%. Therefore, a smooth transition from the inclusion model to the layering model, which may translate into a smooth weighting function between the two distributions, is necessary to properly describe the covariance of the subsurface.

## ACKNOWLEDGMENTS

The authors thank Schlumberger-WesternGeco for the field dataset. This paper includes data supplied by IHS Energy Log Services; Copyright (2013) IHS Energy Log Services Inc.

## REFERENCES

- Bachrach, R., 2010a, Applications of deterministic and stochastic rock physics modeling to anisotropic velocity model building: SEG Expanded Abstracts, **29**, 2436–2440.
- , 2010b, Elastic and resistivity anisotropy of compacting shale: Joint effective medium modeling and field observations: SEG Expanded Abstracts, **29**, 2580–2584.
- Backus, G., 1962, Long-wave elastic anisotropy produced by horizontal layering: Journal of Geophysical Research, **76**, 4427–4440.
- Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2009, Can we distinguish TTI and VTI media?: SEG Expanded Abstracts, **28**, 226–230.
- , 2010a, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, 27–36.
- , 2010b, Localized anisotropic tomography with well information in VTI media: Geophysics, **75**, 37–45.
- Bandyopadhyay, K., 2009, Seismic anisotropy: geological causes and its implications: PhD thesis, Stanford University.
- Dræge, A., M. Jakobsen, and T. A. Johansen, 2006, Rock physics modeling of shale diagenesis: Petroleum Geoscience, **12**, 49–57.
- Hornby, B., D. Miller, C. Esmersoy, and P. Christie, 1995, Ultrasonic-to-seismic measurements of shale anisotropy in a North Sea well: SEG Expanded Abstracts, **14**, 17–21.
- Li, Y. and B. Biondi, 2011, Migration velocity analysis for anisotropic models: SEG Expanded Abstract, **30**, 201–206.
- Li, Y., B. Biondi, D. Nichols, G. Mavko, and R. Clapp, 2013, Stochastic rock physics modeling for seismic anisotropy: SEP report, **149**, 201–220.
- Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011, Anisotropic tomography using rock physics constraints: 73<sup>rd</sup> EAGE Conference & Exhibition.
- Sayers, C., 2004, Seismic anisotropy of shales: What determines the sign of Thomsen’s delta parameter?: SEG Expanded Abstracts, **23**, 103–106.
- , 2010, The effect of anisotropy on the Young’s moduli and Poisson’s ratios of shales: SEG Expanded Abstracts, **29**, 2606–2611.

- Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.
- Yang, Y., K. Osypov, R. Bachrach, M. Woodward, O. Zdraveva, Y. Liu, A. Fournier, and Y. You, 2012, Anisotropic tomography and uncertainty analysis with rock physics constraints: Green Canyon case study: SEG Expanded Abstract, 1–5.

# Wave equation migration velocity analysis for VTI models

*Yunyue Li, Biondo Biondi, Robert Clapp and Dave Nichols*

## ABSTRACT

Anisotropic models are recognized as more realistic representations of the subsurface where a complex geological environment exists. These models are widely needed by all kinds of migration and interpretation schemes. In this paper, we extend the theory of wave equation migration velocity analysis (WEMVA) to build vertical transverse isotropic (VTI) models. Because of the ambiguity between depth and  $\delta$ , we assume  $\delta$  can be accurately obtained from other sources of information, and invert for the NMO slowness and the anellipticity parameter  $\eta$ . We use a differential semblance optimization objective function to evaluate the focusing of the prestack image in the subsurface-offset domain. To regularize the multi-parameter inversion, we build a framework to adapt the geological and the rock physics information to guide the updates in both NMO slowness and anisotropic parameter  $\eta$ . This regularization step is crucial to stabilize the inversion and to produce geologically meaningful results. We test the proposed approach on a 2-D Gulf of Mexico dataset starting with a fairly good initial anisotropic model. The inversion result reveals a shallow anomaly collocated in NMO velocity and  $\eta$  and improves both the continuity and the resolution of the final stacked image.

## INTRODUCTION

When compared with isotropic models, anisotropic models are recognized as more realistic representations of the subsurface in complex geological environment (McCollum and Snell, 1932; Postma, 1955; Helbig, 1956). Many authors (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008) have developed migration and processing schemes for vertical transverse isotropic (VTI) and tilted transverse isotropic (TTI) media; however, the challenge of building a reliable anisotropic Earth model remains the bottleneck for the exploration workflow.

Anisotropic model building tries to resolve more than one parameter at each grid point of the subsurface. This number could be 3 for VTI medium, and increase to 5 for TTI medium. Surface seismic data inversion becomes ill-posed and highly underdetermined due to the rapidly increasing size of the model space with the increasing complexity of the subsurface (Woodward et al., 2008; Bakulin et al., 2010b,a). To reduce the size of the model space, Sarkar and Tsvankin (2003, 2004) and Takanashi and Tsvankin (2012) parametrize their model space using factorized blocks or layers, where the anisotropic parameters  $\epsilon$  and  $\delta$  are constant and the velocity variations are simple linear or quadratic functions of the space coordinates. These factorization methods dramatically decrease the number of unknowns

to be solved by the inversion, and hence stabilize the inversion. However, these models tend to over simplify the Earth where heterogeneities may occur at all scales.

In principle, grid-based tomographic methods handle both anisotropy and heterogeneity at the same time. Most of the existing grid-based anisotropic model-building schemes are based on ray-based tomography by measuring the non-hyperbolic moveout to flatten the common image gathers (CIG). The residual moveout is then back-projected along the ray paths to update the anisotropic models (Zhou et al., 2003, 2004; Yuan et al., 2006; Cai et al., 2009; Woodward et al., 2008). However, ray-based methods are prone to unrealistic results when multi-pathing exists in areas of complex overburden. Furthermore, due to the ill-posed and underdetermined nature of the anisotropic tomography, small amounts of noise in the residual moveout on a CIG gather can lead to significant variations in the inverted model. Therefore, we propose a wave-equation-based tomography method to build the anisotropic model.

In this paper, we first generalize the methodology of image-space WEMVA (Biondi and Sava, 1999; Sava and Biondi, 2004a,b) from an isotropic medium to an anisotropic medium. We use grid-based representations of NMO slowness,  $\eta$  and  $\delta$  to parametrize the VTI medium, but keep  $\delta$  fixed during the iterative WEMVA updates. We show that theoretically the gradient of the tomographic objective functional for an isotropic medium can be modified to describe an anisotropic medium by simply adding a term for additional parameters. To constrain the inversion, we precondition the model using geological and rock physics information. Finally, a field data test demonstrates that the proposed method can identify anomalies in the velocity and the anisotropic model, and hence produce a better migrated image with better spatial continuity and higher resolution.

## WAVE EQUATION MIGRATION VELOCITY ANALYSIS FOR ANISOTROPY

In this section, we extend the methodology of WEMVA to the anisotropic medium. We provide a more concise derivation of the WEMVA gradients than a previous work (Li and Biondi, 2011) using the Lagrangian augmented functional instead of perturbation theory. The interpretations of the adjoint-state equations suggest the same implementation as perturbation theory suggests.

### Governing equations

We parameterize the VTI subsurface using NMO slowness  $s_n$ , and Thomson parameter  $\delta$  (Thomsen, 1986) and the anellipticity parameter  $\eta$  (Alkhalifah and Tsvankin, 1995). In the shot-profile domain, both source wavefields  $\mathbf{D}$  and receiver wavefields  $\mathbf{U}$  are downward continued using the following one-way wave equation and boundary condition (Shan, 2009):

$$\begin{cases} \left( \frac{1}{\sqrt{1+2\delta}} \frac{\partial}{\partial z} + i\Lambda \right) D(\mathbf{x}, \mathbf{x}_s) = 0 \\ D(x, y, z = 0, \mathbf{x}_s) = f_s \delta(\mathbf{x} - \mathbf{x}_s) \end{cases}, \quad (1)$$

and

$$\begin{cases} \left( \frac{1}{\sqrt{1+2\delta}} \frac{\partial}{\partial z} - i\Lambda \right) U(\mathbf{x}, \mathbf{x}_s) = 0 \\ U(x, y, z = 0, \mathbf{x}_s) = d_{\mathbf{x}_s} \delta(\mathbf{x} - \mathbf{x}_r) \end{cases}, \quad (2)$$

where  $f_s$  is the source wavelet,  $d_{\mathbf{x}_s}$  is the shot record by a shot at  $\mathbf{x} = \mathbf{x}_s$ . We approximate the dispersion relation  $\Lambda$  of the VTI medium as follows:

$$\Lambda = \omega s_n \sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 s_n^2 - 2\eta|\mathbf{k}|^2}}, \quad (3)$$

with  $\mathbf{k} = k_x + k_y$ . Equations 1 and 2 can be summarized in matrix forms as follows:

$$\mathbf{L}\mathbf{D} = \mathbf{f} \quad (4)$$

and

$$\mathbf{L}^*\mathbf{U} = \mathbf{d}, \quad (5)$$

where

$$\mathbf{f} = f_s \delta(\mathbf{x} - \mathbf{x}_s, z), \quad (6)$$

$$\mathbf{d} = d_{\mathbf{x}_s} \delta(\mathbf{x} - \mathbf{x}_r, z), \quad (7)$$

and

$$\mathbf{L} = \frac{1}{\sqrt{1 + 2\delta}} \frac{\partial}{\partial z} - i\Lambda. \quad (8)$$

It is well known that parameter  $\delta$  is the least constrained by surface seismic data due to the lack of depth information. Therefore, we assume  $\delta$  can be correctly obtained from other sources of information (such as check shots and well logs) and keep it fixed through the WEMVA iterations.

We use an extended imaging condition (Sava and Formel, 2006) to compute the image cube with subsurface offsets:

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{D})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{U}), \quad (9)$$

where  $\mathbf{S}_{+\mathbf{h}}$  is a shifting operator which shifts the wavefield  $+\mathbf{h}$  in the  $\mathbf{x}$  direction. Notice that  $(\mathbf{S}_{+\mathbf{h}})^* = \mathbf{S}_{-\mathbf{h}}$ . Equations 4, 5 and 9 are state equations, and  $\mathbf{D}$ ,  $\mathbf{U}$  and  $I_{\mathbf{h}}$  are the state variables.

## DSO gradients for anisotropic parameters

To evaluate the accuracy of the subsurface model, we use a DSO objective function (Symes and Carazzone, 1991; Shen, 2004):

$$J = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle. \quad (10)$$

where  $\mathbf{h}$  is the subsurface offset. In practice, other objective functions (linear transformations of the image) can be used rather than DSO. To derive the gradient of the DSO objective function with respect to  $s_n$  and  $\eta$ , we follow the recipe provided by Plessix (2006).

First, we form the Lagrangian augmented functional:

$$\begin{aligned} \mathcal{L}(\mathbf{D}, \mathbf{U}, I_{\mathbf{h}}; \lambda, \mu, \gamma_{\mathbf{h}}; s_n, \eta) = & \quad (11) \\ & \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h} I_{\mathbf{h}}, \mathbf{h} I_{\mathbf{h}} \rangle \\ & + \langle \lambda, \mathbf{f} - \mathbf{L}(s_n, \eta) \mathbf{D} \rangle \\ & + \langle \mu, \mathbf{d} - \mathbf{L}^*(s_n, \eta) \mathbf{U} \rangle \\ & + \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}}, (\mathbf{S}_{+\mathbf{h}} \mathbf{D})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) - I_{\mathbf{h}} \rangle. \end{aligned} \quad (12)$$

Then the adjoint-state equations are obtained by taking the derivative of  $\mathcal{L}$  with respect to state variables  $\mathbf{D}$ ,  $\mathbf{U}$  and  $I_{\mathbf{h}}$ :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{D}} = -\mathbf{L}^*(s_n, \eta) \lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) \gamma_{\mathbf{h}} = \mathbf{0}, \quad (13)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = -\mathbf{L}(s_n, \eta) \mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^* (\mathbf{S}_{+\mathbf{h}} \mathbf{D}) \gamma_{\mathbf{h}} = \mathbf{0}, \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} = 0, \forall \mathbf{h}. \quad (15)$$

Equation 13, 14, and 15 are the adjoint-state equations. Parameters  $\lambda$ ,  $\mu$  and  $\gamma_{\mathbf{h}}$  are the adjoint-state variables, and can be calculated from the adjoint-state equations.

With the solutions to the equations above, we can now derive the gradients of the objective function 10 by taking the derivative of the augmented functional  $\mathcal{L}$  with respect to the model variables  $s_n$  and  $\eta$  as follows:

$$\nabla_{s_n} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \mathbf{s}_n} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \mathbf{s}_n} \mathbf{U} \right\rangle \quad (16)$$

$$\nabla_{\eta} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \eta} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \eta} \mathbf{U} \right\rangle. \quad (17)$$

If the forward modeling operator  $\mathbf{L}$  is isotropic, equation 16 reduces to the gradient of the isotropic DSO objective function with respect to the isotropic velocity. Therefore, the extension of WEMVA to the anisotropic medium is simply adding another term that is related to the additional anisotropic parameter.

## Preconditioning the DSO gradient

Velocity model building is a highly underdetermined and nonlinear problem. Therefore, prior knowledge of the subsurface is needed to define a plausible subsurface model. Assuming Gaussian distribution, Tarantola (1984) characterizes the prior information using the mean and the covariance of the model and includes it as a regularization term. We separate the covariance into two parts, a spatial covariance between the same parameter at two different locations and a cross parameter covariance between different parameters at

the same location. Physically, the spatial covariance describes the smoothness/roughness of the model parameter with respect to the space coordinates (Clapp, 2000; Woodward et al., 2008). In the multi-parameter anisotropic model estimation, spatial correlation lag for anisotropic parameters are considered longer than that for velocity (Woodward et al., 2008). Therefore, we can design different spatial smoothers for slowness and anisotropic parameters to reflect this prior information.

Moreover, given a certain lithology, velocity and anisotropic parameters are correlated with each other. Bachrach et al. (2011) shows that additional well log information also reduces the plausible range of correlations between the anisotropic parameters. Therefore, a local cross-parameter covariance is also necessary to include the prior lithological knowledge of the subsurface.

In this study, we assume the covariance of the model has two independent components: spatial covariance and local cross-parameter covariance (Li et al., 2011). In practice, instead of regularizing the inversion using Tarantola (1984), we use a preconditioning scheme (Claerbout, 2009) to speed up the convergence: smoothing filtering to approximate square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the preconditioning variable  $\mathbf{n}$  is related to the original model  $\mathbf{m}$  as follows:

$$\mathbf{m} = \mathbf{B}\Sigma\mathbf{n}. \quad (18)$$

In equation 18, the smoothing operator  $\mathbf{B}$  is a band-limited diagonal matrix:

$$\mathbf{B} = \begin{vmatrix} \mathbf{B}_s & 0 \\ 0 & \mathbf{B}_\eta \end{vmatrix}. \quad (19)$$

with different smoothing operators for NMO slowness and  $\eta$ , according to the geological information in the study area. The standard deviation matrix  $\Sigma$ :

$$\Sigma = \begin{vmatrix} \sigma_{ss}\mathbf{I} & \sigma_{s\eta}\mathbf{I} \\ \sigma_{s\eta}\mathbf{I} & \sigma_{\eta\eta}\mathbf{I} \end{vmatrix}. \quad (20)$$

can be obtained by rock-physics modeling and/or lab measurements (Bachrach et al., 2011; Li et al., 2011). In an ideal case, we should estimate matrix  $\Sigma$  at each subsurface location to reflect the local lithological information.

The initial preconditioning model  $\mathbf{n}_0$  is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \langle \mathbf{m}_0 - \mathbf{B}\Sigma\mathbf{n}_0, \mathbf{m}_0 - \mathbf{B}\Sigma\mathbf{n}_0 \rangle. \quad (21)$$

The gradient of the objective function (10) with respect to this preconditioning variable  $\mathbf{n}$  is

$$\begin{aligned} \nabla_{\mathbf{n}}J &= \left(\frac{\partial \mathbf{m}}{\partial \mathbf{n}}\right)^* \nabla_{\mathbf{m}}J \\ &= \Sigma^* \mathbf{B}^* \nabla_{\mathbf{m}}J, \end{aligned} \quad (22)$$

where  $\nabla_{\mathbf{m}}J = [\nabla_{s_n}J \ \nabla_{\eta}J]^T$ . We use a nonlinear conjugate gradient algorithm to minimize the objective function.

## FIELD DATA TESTS

Anisotropic parameters cannot be fully constrained by surface seismic data. Therefore, anisotropic models are better defined in well-developed and production areas, where various well log, check shots and VSP data are available. The Gulf of Mexico dataset used in this study is from a producing area in an offshore sedimentary basin close to a salt body. While vertical check shots and sonic logs best constrain the vertical velocity model, offset check shot (OCS) data, thanks to their long offset give information on both vertical and horizontal velocity. The current anisotropic model was built using an interactive visualization method (Bear et al., 2005) that integrates surface seismic, sonic logs, vertical check shots, and offset check shots.

We extract a 2-D line from the Gulf of Mexico dataset at a location where the salt body is far away. Source spacing is 100 m, and receiver spacing is 50 m. The maximum offset in this 2-D line is about 4 km. The initial stack image is shown on the top panel in Figure 4. Although the apparent dip is high due to the large vertical stretch (3 km in depth vs. 16 km on the horizontal axis), the reflectors in this 2-D line are fairly flat. We then estimate the dip field from the initial image and use it to precondition the gradient for both slowness and  $\eta$ . Due to the lack of rock physics information, a stationary matrix  $\Sigma$  with  $\sigma_{s\eta} = 0$  is applied in the preconditioning to scale the relative amplitude of the slowness and  $\eta$  gradient.

We can see many small-scale faults in this area on the top panel in Figure 4. Migration artifacts at  $x = 13$  km and  $z = 2500$  m are caused by a big vertical fault running from  $x = 14$  km on the top to the bottom of the section. The initial angle gathers are shown in the bottom row in Figure 4. Since this is a streamer geometry, the subsurface reflectors are only illuminated from positive angles. Although the gathers are close to flat, we can still see upward residual moveouts in the angle domain. Therefore, we have a chance to improve the model and the image by flattening the gathers.

The inverted anisotropic model after 8 iterations is shown in figure 2. Updates between the initial and the inverted velocity and  $\eta$  models are shown in figure 3(a) and 3(b), respectively. First, notice that we are able to resolve a localized shallow anomaly between 13 km and 15 km at around 800 m below surface. Also, notice that the updates in velocity are less than 10%, whereas the updates in  $\eta$  are around 25%. These positive updates in both velocity and  $\eta$  agree well with the negative travel time misfits in the previous OCS modeling results (Bear et al., 2005).

Comparing the initial stack image on the top panel in Figure 4 with the final stack image on the top panel of Figure 5, we can see improved continuity and signal strength in the area highlighted by the oval. The faulting in this area is also better defined in the final image. If we correlate the inverted models with the final migrated image, we may choose to change the interpretation of the major fault: the emerging location on the surface may be shifted from 14.2 km to around 16 km.

We can also verify the effect of the updates in velocity and  $\eta$  on the angle gathers at different CMP locations. The initial angle domain common image gathers (ADCIGs) are shown in the bottom row in Figure 4, and the final ADCIGs produced using the inverted models are shown in the bottom row in Figure 5. To better illustrate the effects of the model updates, the ADCIGs are sampled more densely between  $\text{CMP} = 13\text{km}$  and  $16\text{km}$  and sparsely outside of this range. In general, we can see improved flatness for all the

reflectors. Specifically, for the shallower events above 1km, most improvements happen at large angles over  $35^\circ$ . Therefore, we interpret the improvements for the shallow events primarily as the contribution of the improved  $\eta$  model.

For the deeper events at the same CMP location, both the depth and the flatness of the angle gather have been changed by inversion. The upward-curving events in the angle domain from the initial migration has been flattened by the improved velocity and  $\eta$  model.

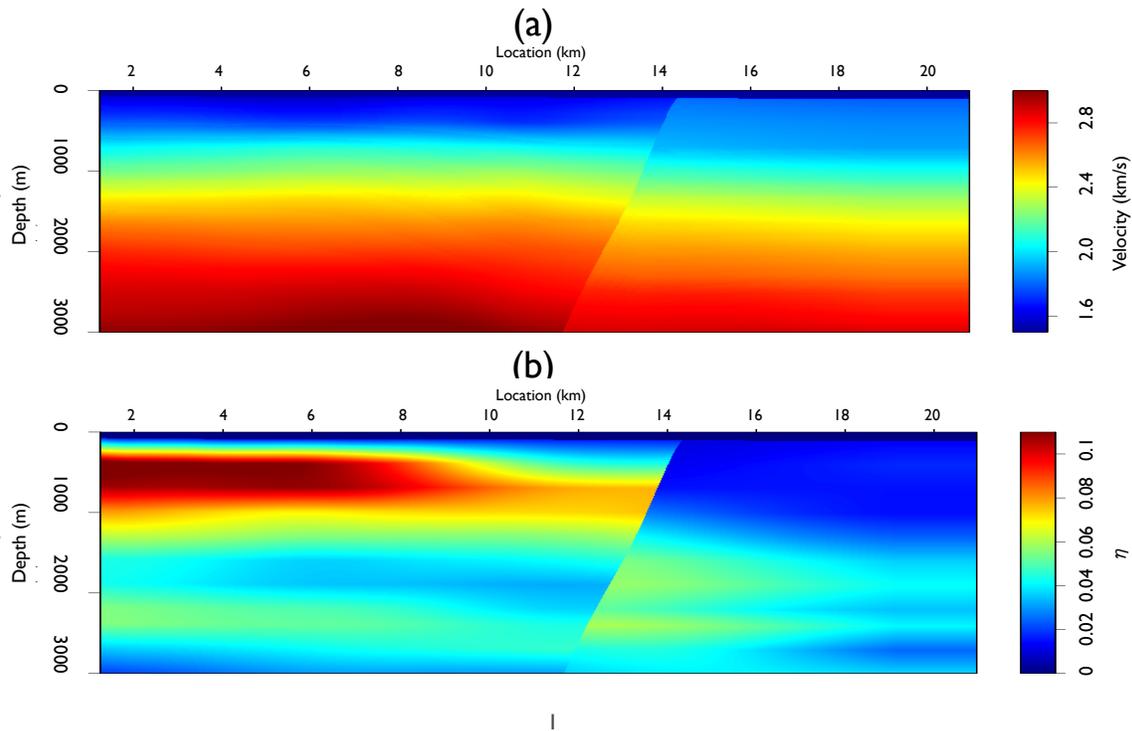


Figure 1: Initial velocity model (a) and initial  $\eta$  model (b). `elita3/. initmodel`

## CONCLUSIONS

We have presented a methodology of wave equation migration velocity analysis for anisotropic media. Our method is fully automated and picking-free. The grid-based parameterization for the velocity and anisotropic parameters provides the flexibility to reveal perturbations with different scales. The regularization scheme of our proposed method is crucial to obtain geological meaningful results for both velocity and anisotropic parameters. It is also the interface to include human interpretation and other prior knowledge. The test of our anisotropic WEMVA algorithm on a 2-D slice of a GoM field dataset produces improved stacked image with better continuity and higher resolution. Therefore, we believe our method provides a good model building tool for anisotropic media.

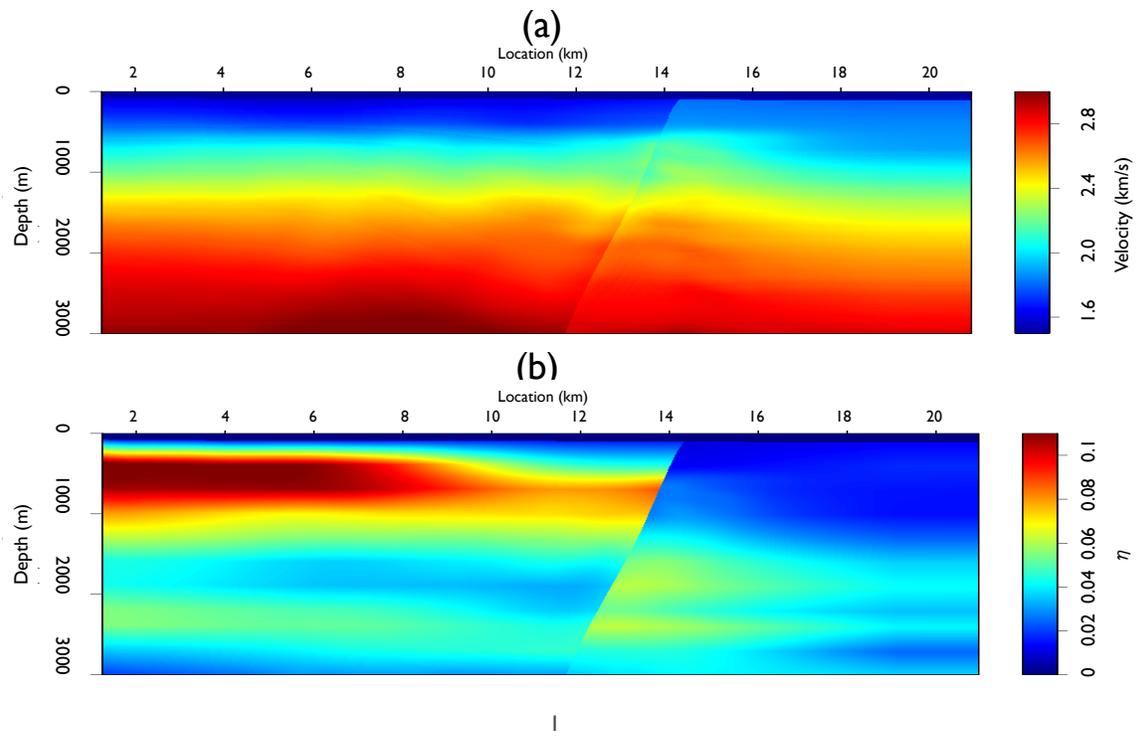


Figure 2: Inverted velocity model (a) and inverted  $\eta$  model (b). `elita3/. finalmodel`

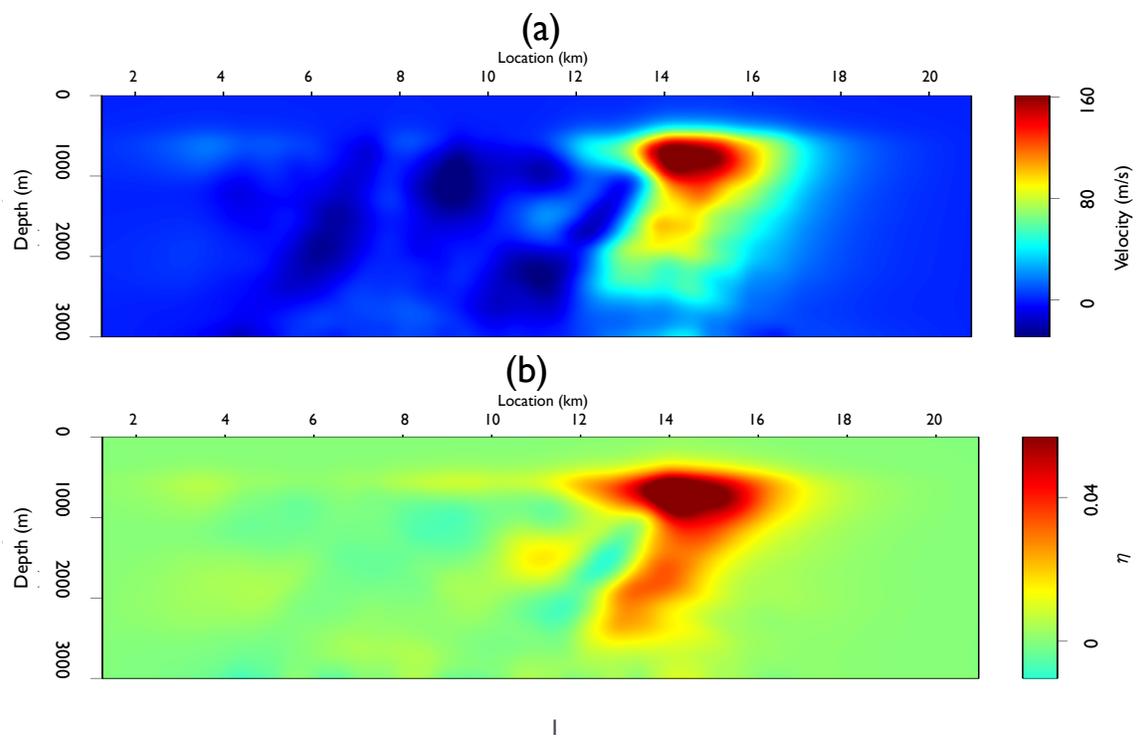


Figure 3: Updates in velocity model (a) and updates in  $\eta$  model (b) after inversion.

elita3/. updates

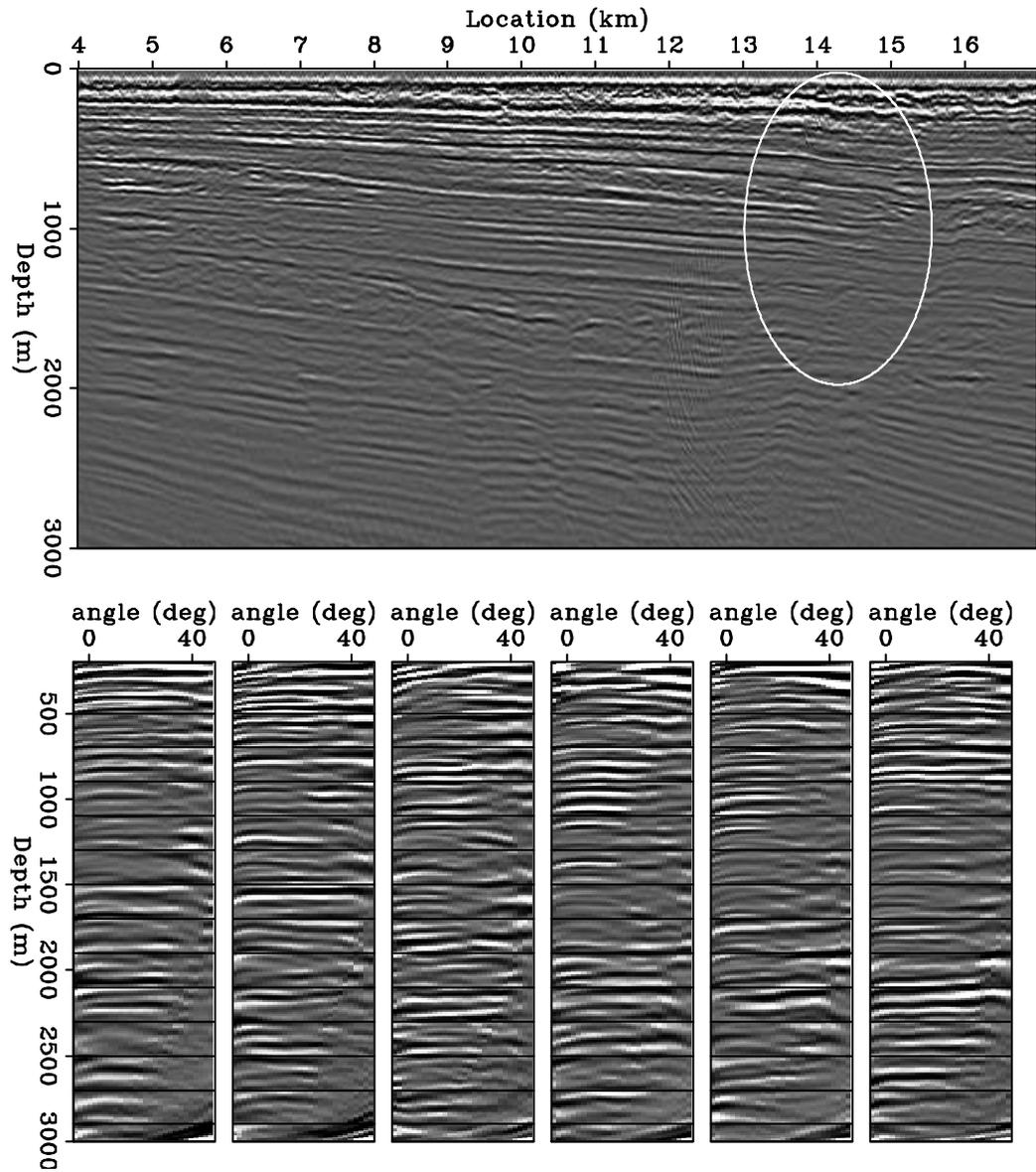


Figure 4: The initial stack image (Top panel) and initial angle domain common image gathers at CMP = 7, 10, 13, 14, 14.5, 15 km (Bottom row). `elita3/. image-init-an`

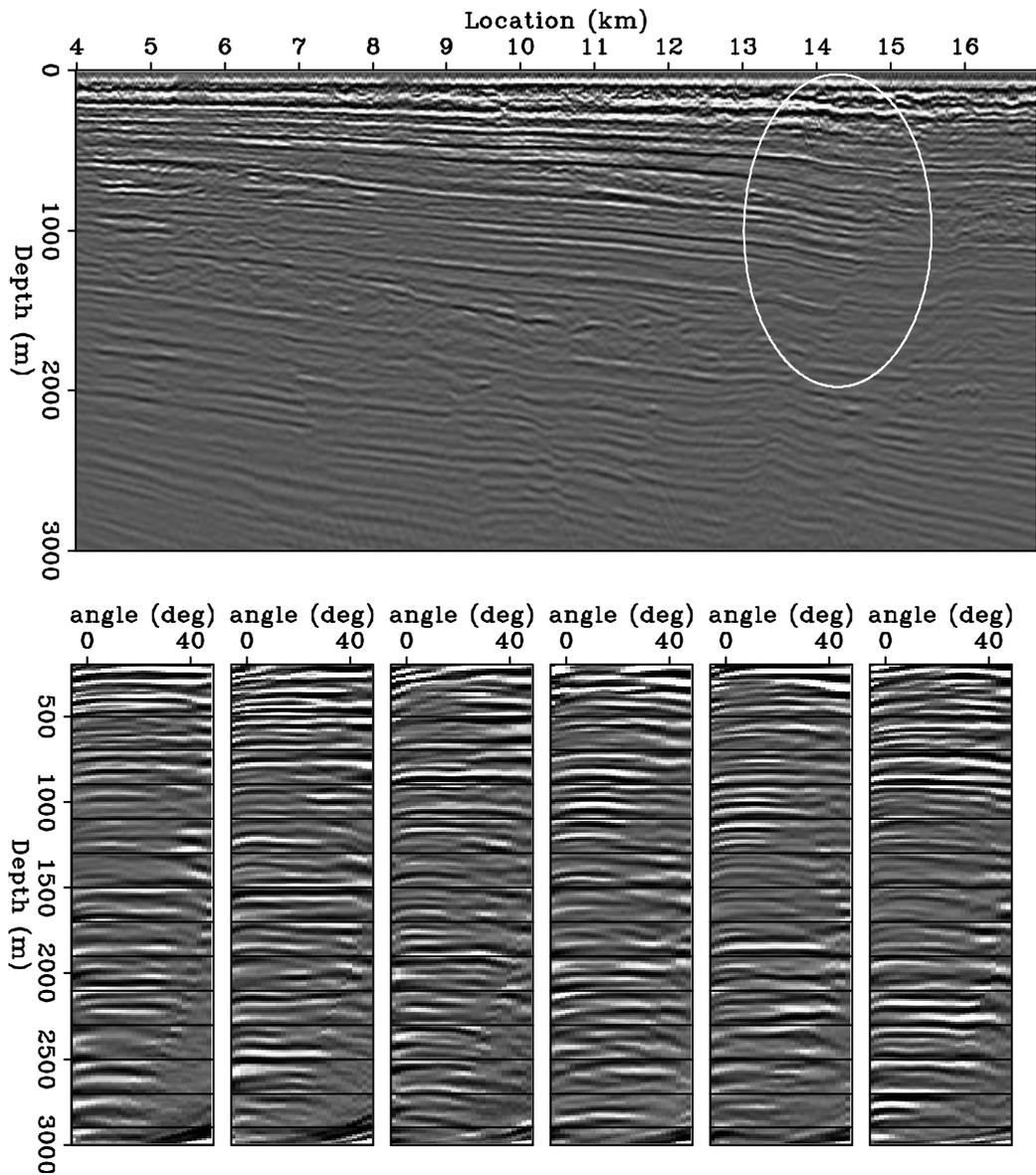


Figure 5: The final stack image (Top panel) and final angle domain common image gathers at CMP = 7, 10, 13, 14, 14.5, 15 km (Bottom row). Compared with Figure 4, improvements in continuity and enhancements in amplitude strength are highlighted by the oval.

elita3/. image-fnal-an

## ACKNOWLEDGEMENT

The authors thank ExxonMobil for providing the field data and the sponsors of the Stanford Exploration Project for their financial support.

## REFERENCES

- Alkhalifah, T. and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: *Geophysics*, **60**, 1550–1566.
- Bachrach, R., Y. K. Liu, M. Woodward, O. Zradrova, Y. Yang, and K. Osypov, 2011, Anisotropic velocity model building using rock physics: Comparison of compaction trends and check-shot-derived anisotropy in the gulf of mexico: *SEG Expanded Abstract*, **30**, 207–211.
- Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2010a, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: *Geophysics*, **75**, 27–36.
- , 2010b, Localized anisotropic tomography with well information in VTI media: *Geophysics*, **75**, 37–45.
- Bear, L., T. Dickens, J. Krebs, J. Liu, and P. Traynin, 2005, Integrated velocity model estimation for improved positioning with anisotropic PSDM: *The Leading Edge*, 622–634.
- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: *SEG Expanded Abstract*, **69**, 1723–1726.
- Cai, J., Y. He, Z. Li, B. Wang, and M. Guo, 2009, TTI/VTI anisotropy parameters estimation by focusing analysis, part i: theory: *SEG Expanded Abstracts*, **28**.
- Claerbout, J. F., 2009, *Image estimation by example*.
- Clapp, R., 2000, *Geologically constrained migration velocity analysis*: PhD thesis, Stanford University.
- Fei, T. W. and C. L. Liner, 2008, Hybrid fourier finite-difference 3D depth migration for anisotropic media: *Geophysics*, **73**, S27.
- Fletcher, R., X. Du, and P. J. Fowler, 2009, Stabilizing acoustic reverse-time migration in TTI media: *SEG Expanded Abstracts*, **28**.
- Helbig, K., 1956, Die ausbreitung elastischer Wellen in anisotropen Medien: *Geophys. Prosp.*, **04**, 70–81.
- Li, Y. and B. Biondi, 2011, Migration velocity analysis for anisotropic models: *SEG Expanded Abstract*, **30**, 201–206.
- Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011, Anisotropic tomography using rock physics constraints: *73<sup>rd</sup> EAGE Conference & Exhibition*.
- McCollum, B. and F. Snell, 1932, Asymmetry of sound velocity in stratified formations: *Physics (Journal of Applied Physics)*, **2**, 174–185.
- Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: *Geophysical Journal International*, **167**, 495–503.
- Postma, G. W., 1955, Wave propagation in a stratified medium: *Geophysics*, **20**, 780–806.
- Sarkar, D. and I. Tsvankin, 2003, Analysis of image gathers in factorized VTI media: *Geophysics*, **68**, 2016–2025.
- , 2004, Migration velocity analysis in factorized VTI media: *Geophysics*, **69**, 708–718.
- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis-I: Theory: *Geophysical Prospecting*, **52**, 593–606.

- , 2004b, Wave-equation migration velocity analysis-II: Examples: Geophysical Prospecting, **52**, 607–623.
- Sava, P. and S. Formel, 2006, Generalized imaging conditions for wave equation migration: CWP Report, **524**.
- Shan, G., 2009, Optimized implicit finite-difference and fourier finite-difference migration for VTI media: Geophysics, WCA189–WCA197.
- Shen, P., 2004, Wave-equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Symes, W. W. and J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.
- Takanashi, M. and I. Tsvankin, 2012, Migration velocity analysis for TI media in the presence of quadratic lateral velocity variation: Geophysics, **77**, U87–U96.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
- Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.
- Woodward, M. J., D. Nichols, O. Zdraveva, P. Whitfield, and T. Johns, 2008, A decade of tomography: Geophysics, **73**, VE5VE11.
- Yuan, J., X. Ma, S. Lin, and D. Lowrey, 2006, P-wave tomographic velocity updating in 3d inhomogeneous VTI media: SEG Expanded Abstracts, **25**.
- Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, **28**.
- Zhou, H., D. Pham, and S. Gray, 2004, Tomographic velocity analysis in strongly anisotropic TTI media: SEG Expanded Abstracts, **23**.
- Zhou, H., D. Pham, S. Gray, and B. Wang, 2003, 3-d tomographic velocity analysis in transversely isotropic media: SEG Expanded Abstracts, **22**.



## Integrated seismic image segmentation and efficient model evaluation: field data example

*Adam Halpert*

### ABSTRACT

Salt interpretation and model building in areas with complicated salt geology represent significant bottlenecks during large iterative imaging projects. Automated tools like image segmentation can help interpreters quickly identify salt bodies in 3D seismic volumes, reducing the need for time-consuming manual picking. In addition, a scheme to efficiently test multiple possible models without fully re-migrating the dataset is useful when more than one salt scenario is in play. Here, a 3D field data example demonstrates that a combination of these two computational interpretation tools can effectively generate and test alternative models. Re-migration with a preferred velocity model produces an improved subsalt image.

### INTRODUCTION

Recent increases in computing power have shifted model-building bottlenecks from computational tasks toward interpretation and similar human-intensive tasks. One approach to alleviate these bottlenecks is to develop computational interpretation tools, which can allow interpreters to take advantage of increased computational capabilities, while still allowing them to use their expertise to control the interpretation workflow. Two such tools are seismic image segmentation, and an efficient velocity model-evaluation method using synthesized wavefields. Here, I will use a 3D field data example from the Gulf of Mexico to demonstrate how these two tools can work together to effectively generate and test velocity models based on different salt scenarios, leading to an improved subsalt image.

Salt interpretation is crucial for successful model building. Because of the high velocity contrast between salt and sediments, mis-placed or mis-interpreted salt bodies can have a disproportionately negative impact on the quality of subsalt images. Seismic image segmentation is one tool interpreters can use to partially automate salt body delineation. Several individual algorithms have shown promise for this task, including the Normalized Cuts Image Segmentation scheme (Shi and Malik, 2000), which can be effectively adapted for use on 3D seismic images (Lomask, 2007; Lomask et al., 2007). In this report, I use the segmentation scheme of Felzenszwalb and Huttenlocher (2004), which can also be successfully applied to 3D seismic data (Halpert, 2010). Importantly, this method can also incorporate interpreter input and rely on the interpreter's discretion to identify the extent of salt bodies (Halpert, 2013). These capabilities allow us to quickly generate a range of possible velocity models based on different, but plausible, salt interpretations. A way to efficiently test these models is necessary in order to avoid costly and time-consuming re-migration of very large modern 3D datasets.

Quickly updating images based on discrete changes to the velocity model is another

problem with several proposed solutions. Many proposals rely on a fast, target-oriented variant of beam migration (Hill, 1990); for example, Wang et al. (2008) update post-stack images to qualitatively judge the effects of using different velocity models. An alternative is to use a similar approach but with wave-equation methods like reverse-time migration (RTM) (Wang et al., 2011); however, this remains an extremely expensive option. A different approach, first proposed by Halpert and Tang (2011), also employs wave equation imaging. However, by using an initial image and Born modeling (Stolt and Benson, 1986) to synthesize a new, targeted dataset (Tang and Biondi, 2010; Tang, 2011), the computational challenge is drastically reduced. Furthermore, if a pre-stack image with subsurface offset information is available, this method can make use of that information to identify and help correct errors in an initial velocity model.

In the following sections, I will briefly review the methods used for both the image segmentation and model evaluation parts of the computational interpretation workflow. I will then demonstrate how these tools can be applied to a 3D example from a Gulf of Mexico survey provided courtesy of WesternGeco (Figure 1(a)). Using image segmentation, I will generate a new velocity model based on a salt interpretation slightly different than the one used to create the model provided with the data (Figure 1(b)). In Figure 1, note that an inclusion within the salt body has not been assigned a velocity distinct from the rest of the salt, and that the interpretation of the base of salt is somewhat ambiguous. Both of these factors could contribute to a fading of the subsalt reflectors seen in Figure 1(a). Finally, I will compare these two models using the synthesized wavefield methodology mentioned above, and validate the comparison by showing a full re-migration of the data using the alternate model.

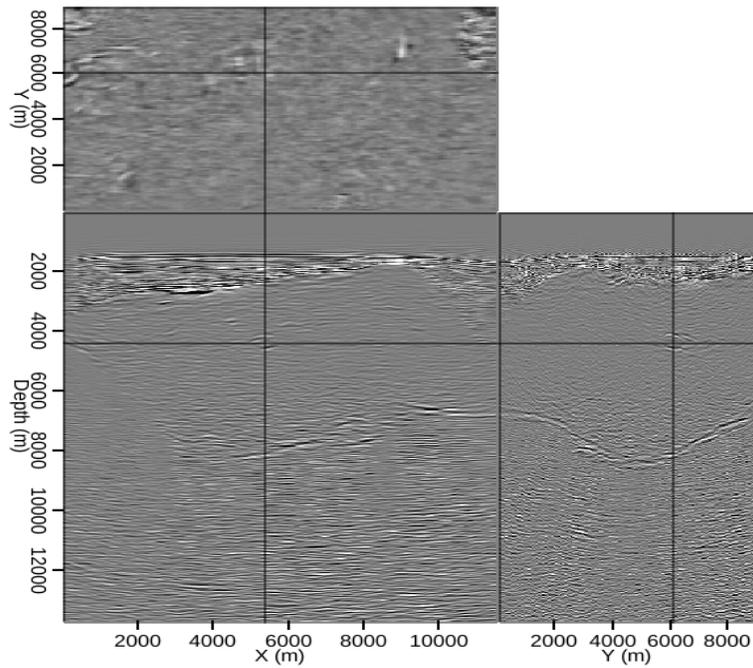
## IMAGE SEGMENTATION

The Pairwise Region Comparison (PRC) image segmentation algorithm is a graph-cut technique based on the method of Felzenszwalb and Huttenlocher (2004). Unlike other graph partitioning algorithms such as the Normalized Cuts method (Shi and Malik, 2000; Lomask et al., 2007), it operates extremely efficiently, requiring on the order of  $n \log n$  operations for an image with  $n$  pixels (compared to  $n^2$  operations for Normalized Cuts). When adapted for use with seismic images (Halpert, 2010), this method can help identify salt bodies or other regions by constructing *edges* between each pixel in the image and several surrounding pixels. Each edge is weighted according to the largest amplitude value that exists between the two end pixels  $i$  and  $j$ :

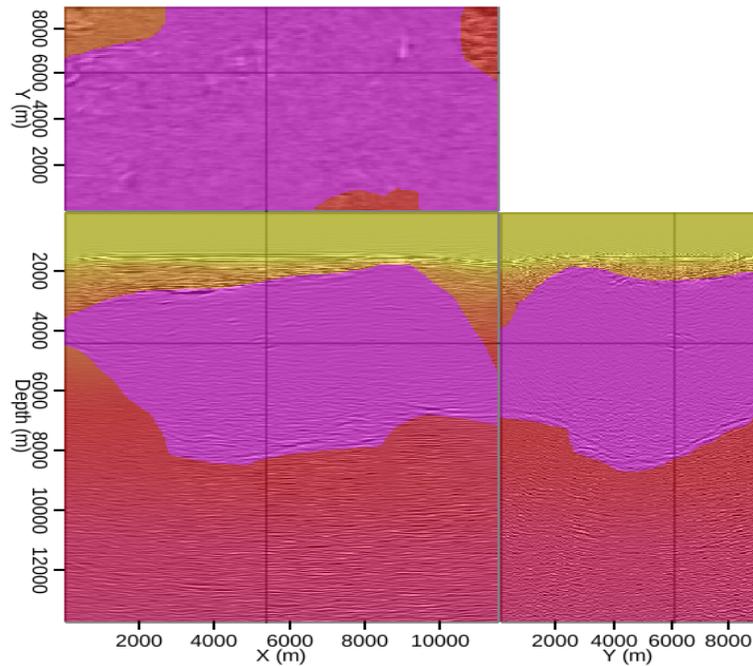
$$w_{ij} = \exp((\max I(\mathbf{p}_{ij}))^2) \exp(d_{ij}), \quad (1)$$

where  $\mathbf{p}_{ij}$  is the vector of all pixels between  $i$  and  $j$  and  $d_{ij}$  is simply the Euclidean distance (in samples) between the two pixels. This approach has the effect of identifying high-amplitude edges, often indicative of salt boundaries.

Once each of the edges is assigned a weight, the segmentation of the image can proceed as described in Felzenszwalb and Huttenlocher (2004). In summary, the process begins with each pixel as its own image segment; then individual pixels, and eventually, groups of pixels, are merged according to thresholding criteria. Segments can also be merged in post-processing if they are smaller than a “minimum segment size” parameter specified by the user. Furthermore, interpreter input may be included if a fully automatic segmentation



(a)



(b)

Figure 1: (a) A 3D image from the Gulf of Mexico (data courtesy of WesternGeco) obtained via one-way migration with the velocity model shown in (b). A prominent sediment inclusion within the salt body, and/or a misinterpreted base of salt, may contribute to the subsalt reflectors' loss of continuity. [CR] `adam1/. img-orig,vel-orig`

is not feasible. For a 3D segmentation, an interpreter can provide a limited manual interpretation on one or more 2D slices; this information is then “sprayed” into the third dimension to alter the input data prior to segmentation. In addition, interpreters maintain control of the process by choosing which segments to include within a salt body, for example. These chosen segments can be easily merged together, and used to create or modify an existing velocity model.

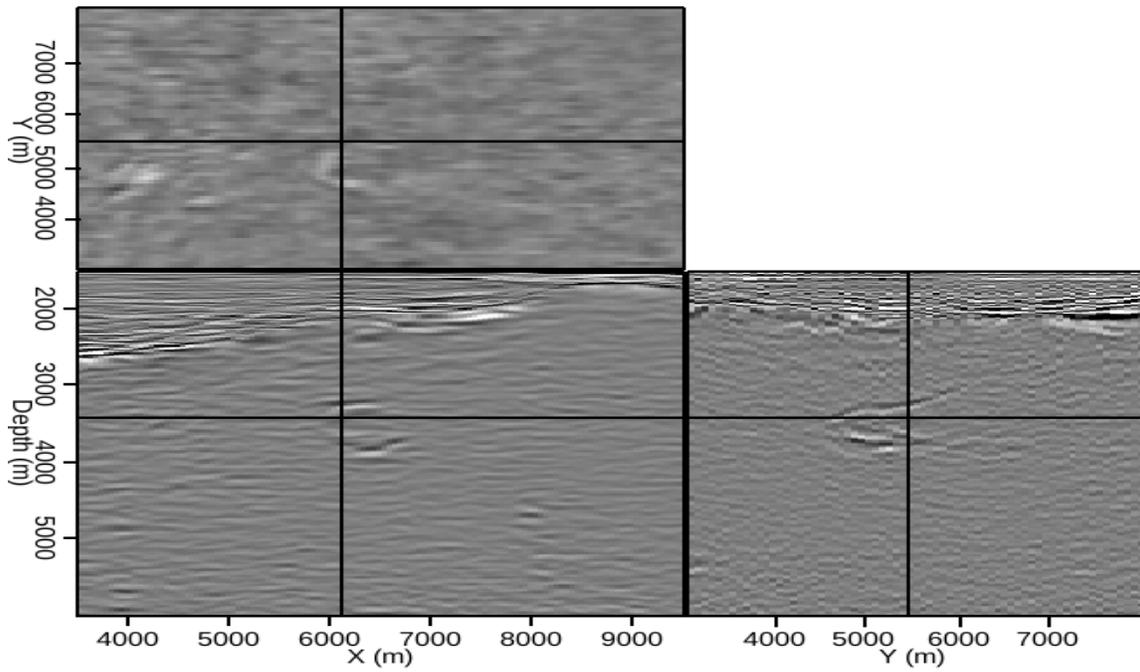
Figure 2(a) is a close-up image of the salt inclusion mentioned in the introduction. By isolating this smaller region for segmentation analysis, we are free to set the minimum segment size to a small number, allowing the automatic segmentation process to capture a higher degree of detail. Figure 2(b) is the interpreter-guided segmentation result for the salt inclusion. After following a similar process to define a new base-salt boundary, a new velocity model is produced by assigning appropriate velocities to the segmented regions. Figures 3(a) and 3(b) show the original and modified velocity models, respectively, for this region. In this case, replacement velocities were taken at appropriate depths from the background sediment velocities in areas without salt.

## MODEL EVALUATION

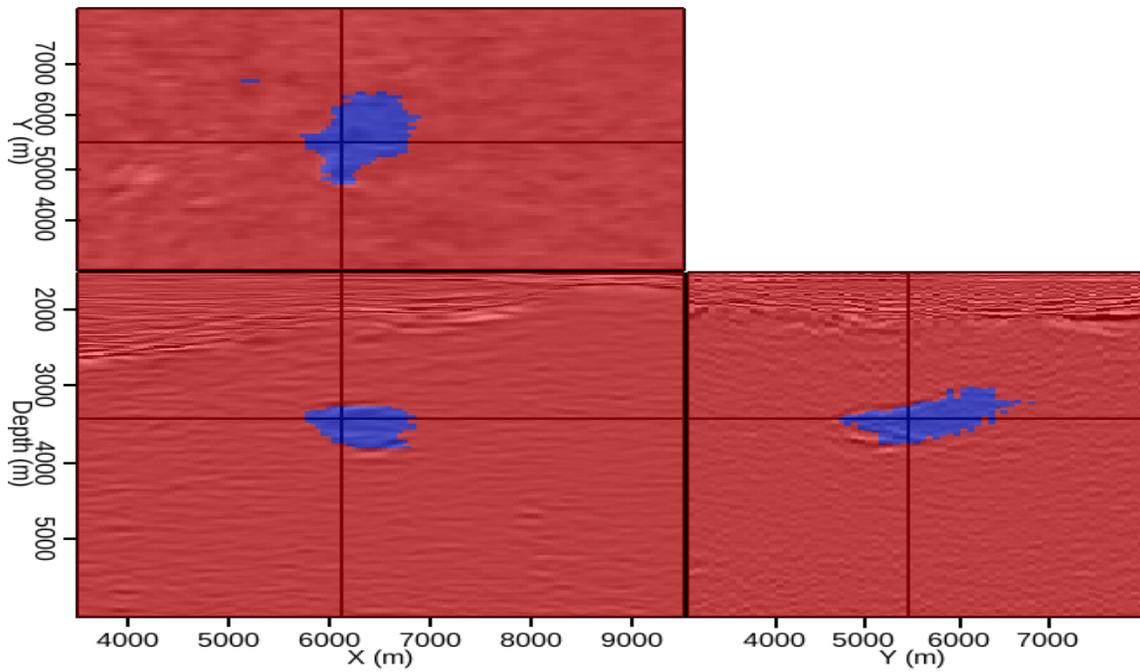
The efficient velocity model evaluation scheme used here is based on the concept of synthesized Born-modeled wavefields (Stolt and Benson, 1986; Tang and Biondi, 2010; Tang, 2011), and was first introduced by Halpert and Tang (2011). The method also uses the “prestack exploding reflector” concept (Guerra, 2010) to identify and correct errors in an initial velocity model and its associated prestack image. The model evaluation procedure consists of three major steps (for mathematical details, see Halpert and Tang (2011)):

1. Starting from an initial image, use a form of exploding reflector modeling to “record” an areal source function at a preferred datum. Crosstalk artifacts are reduced by first having an interpreter select a horizon of interest from the initial image; for example, the base salt reflector indicated in Figure 4. Then, isolated points from the reflector are selected for further analysis. As long as the selected locations are separated by at least twice the maximum subsurface half-offset used from the initial image, crosstalk issues are avoided.
2. Using the new source function and the initial image as a reflectivity model, synthesize a Born-modeled dataset (Tang, 2011) with arbitrary acquisition geometry suited to the imaging target(s). As long as the same velocity model used to create the initial image is used to model this new dataset, the new receiver wavefield will be kinematically invariant of the initial model. This property is what allows us to confidently test additional models in the next step.
3. Now, the synthesized source and receiver wavefields can be imaged conventionally, using any velocity model under consideration. New images can be produced using only a single shot, a substantial computational savings over re-migrations of the full dataset.

To test the models seen in Figure 3, we choose a single location from the base-salt reflector picked in Figure 4. According to the procedure outlined above, new areal source



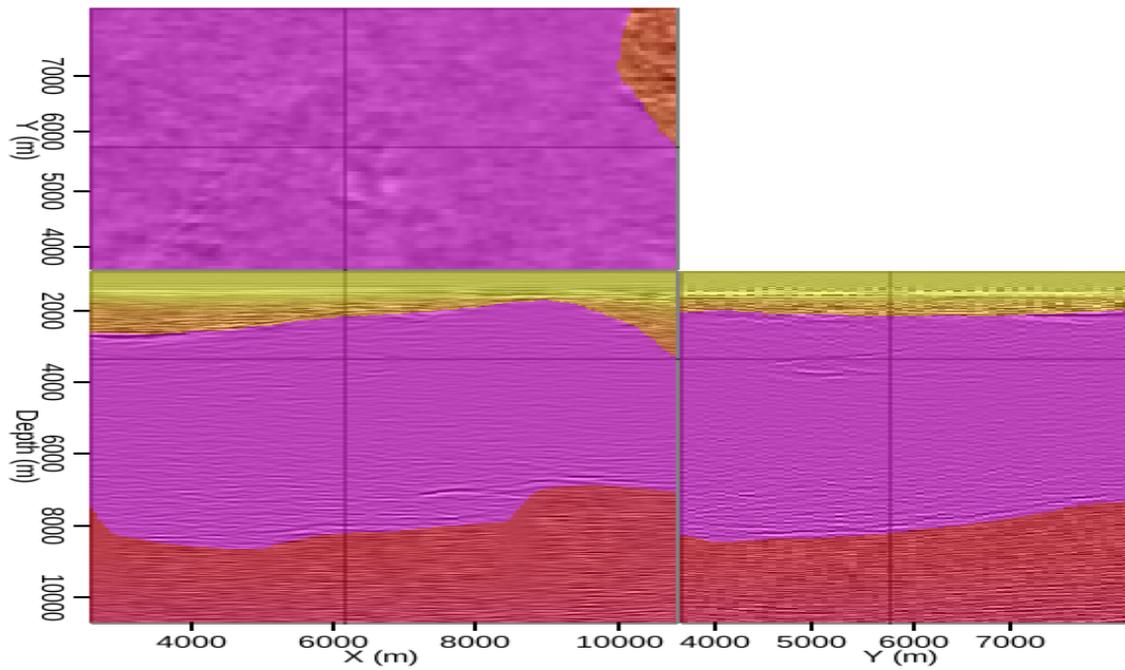
(a)



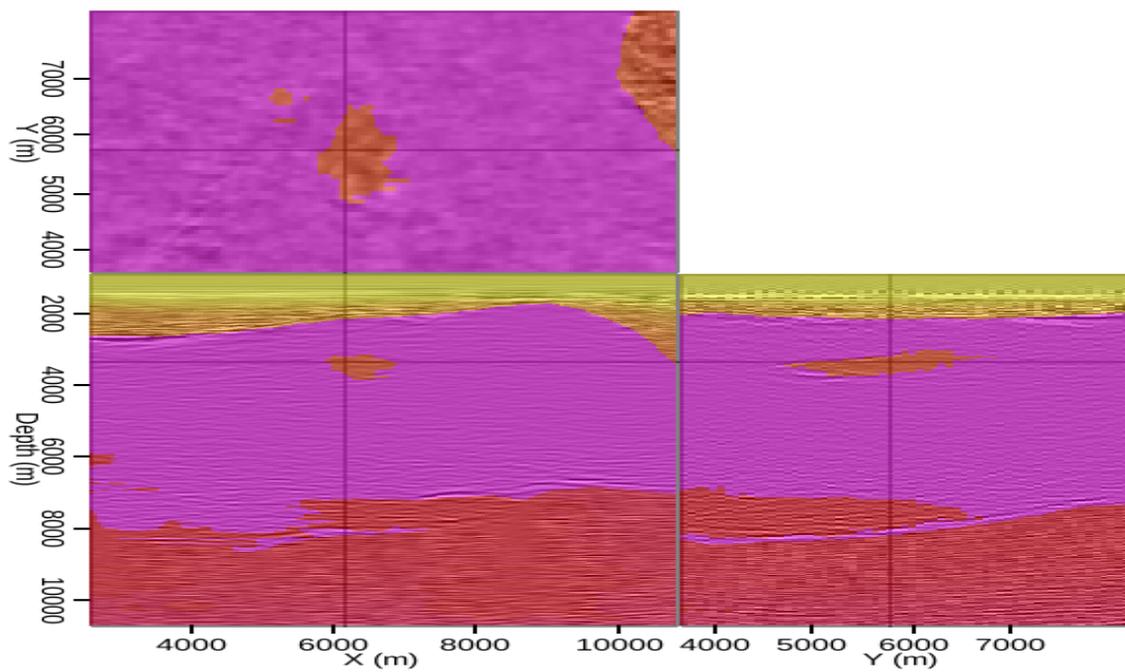
(b)

Figure 2: (a) Close-up view of the sediment inclusion first seen in Figure 1(a); (b) Interpreter-guided 3D segmentation of the image in (a). [CR]

adam1/. sizoom-img,sizoom-seg



(a)



(b)

Figure 3: (a) The velocity model provided with the data; (b) An updated model based on the segmentation result in Figure 2(b) (and another defining and alternative base-salt interpretation). [CR] `adam1/. vzoom-orig,vzoom-new`

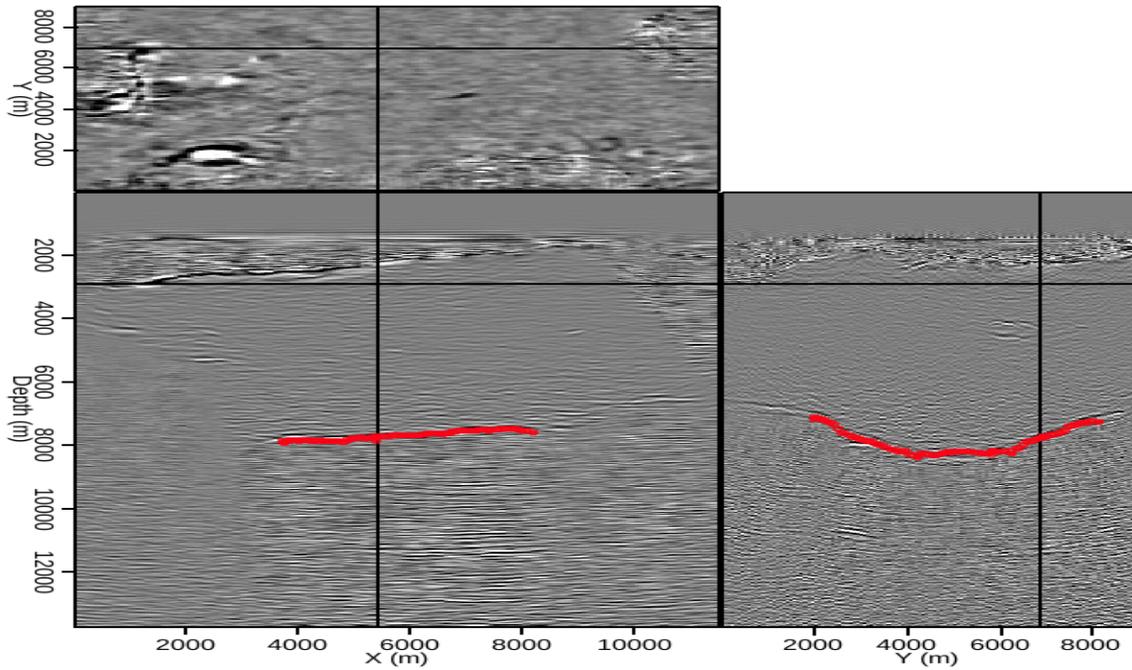


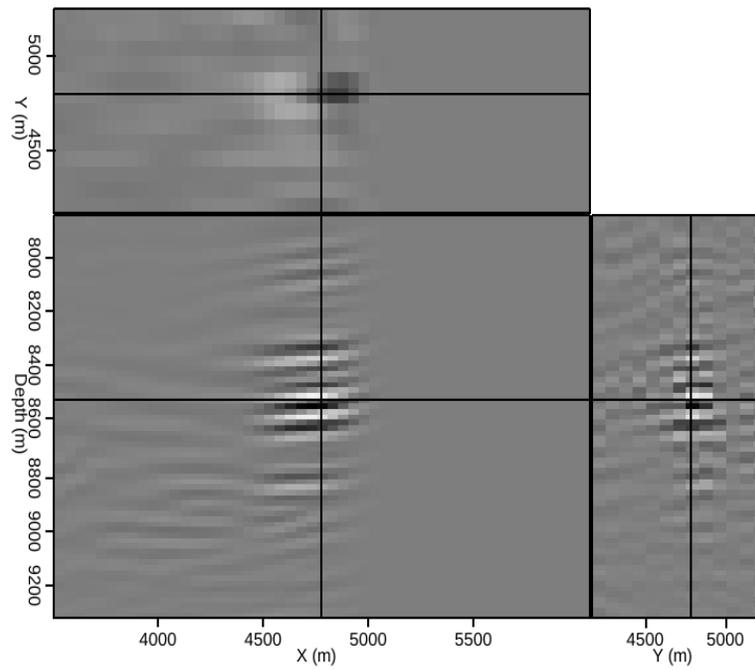
Figure 4: A manually-selected base-salt reflector that will be used to quickly evaluate the velocity models in Figure 3. `adam1/. base-picks`

and receiver wavefields were synthesized using the initial velocity model. Then, the location isolated from the picked reflector was imaged using the two different models. The results in Figure 5 are qualitatively extremely similar, although a slight increase in focusing may be detected along the crossline direction in Figure 5(b). A larger difference between the two images can be detected by examining the subsurface offset domain; Figures 6(a) and 6(b) show this information taken from the  $x, y, z$  location indicated in Figure 5. Now, a higher degree of focusing is more evident in the image obtained using the new model (Figure 6(b)). However, a quantitative measure of image focusing is still desirable, especially when the images are difficult to judge qualitatively, as is the case here. To measure image focusing, we use

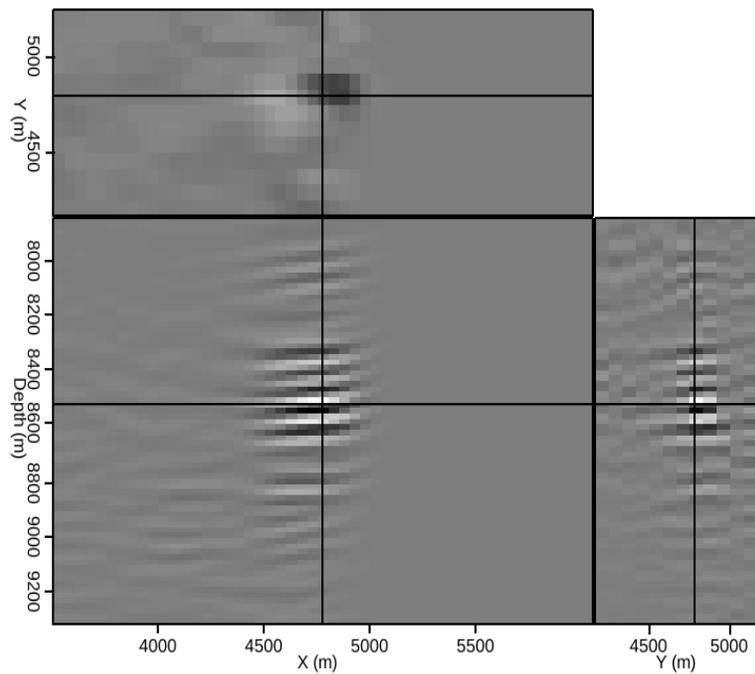
$$F = \frac{\sum_{i=\mathbf{p}} |A_i|}{\sum_{i=\mathbf{p}} |A_i| \exp(\alpha \frac{|h_i|}{h_{\max}})}, \quad (2)$$

where  $\mathbf{p}$  is the set of all image points,  $A_i$  is the amplitude at a given point,  $h_i$  is the subsurface offset at that point, and  $\alpha$  is an optional user-specified weighting parameter. Using this measure, a value of  $F = 1$  means that all energy is perfectly focused at zero offset; as  $F$  decreases toward zero the image becomes progressively less focused. In this case, the  $F$  value for the image obtained using the original velocity model was 0.727, while the new model yielded an  $F$  value of 0.736, indicating that the new model provides a better-focused image at that location.

To test the prediction of the model evaluation procedure, full migrations were performed using both of the tested models. Figures 7(a) and 7(b) are images produced using the original and new models, respectively. Both migrations were performed using a one-way, split-step Fourier method with interpolation, using identical parameters. There are clear



(a)



(b)

Figure 5: An isolated location from the reflector in Figure 4, quickly imaged using (a) the original velocity model; and (b) the updated model. [CR] `adam1/. born-orig,born-new`

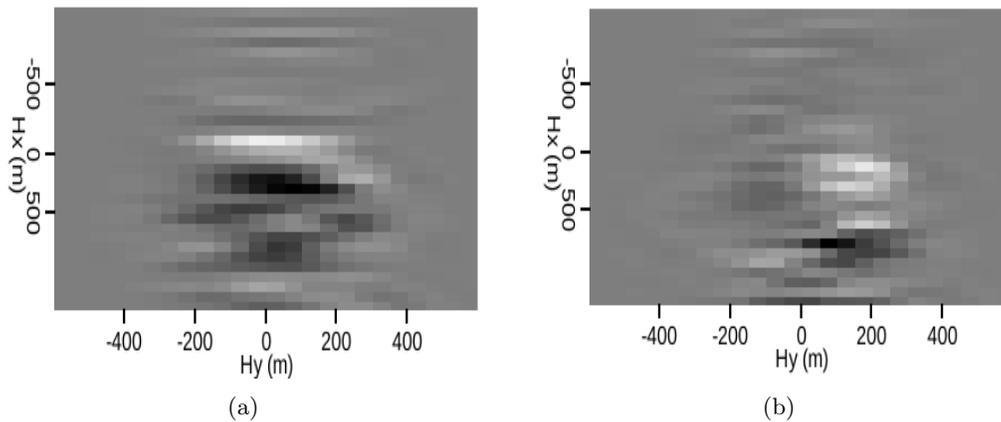


Figure 6: Subsurface offset panels at a single  $x, y, z$  location indicated in Figure 5 for the image corresponding to (a) the original velocity model; and (b) the updated model. [CR] adam1/. hxy-orig,hxy-new

differences between the two results, and in our subsalt target area, the new model has produced a superior result, especially at the indicated locations.

## CONCLUSIONS

Computational interpretation tools such as interpreter-guided image segmentation and efficient model evaluation using synthesized wavefields can effectively add automation to an interpreter-driven model building workflow. In this example, image segmentation was used to delineate a salt body inclusion and define a base salt slightly different from an original model. To test a new model derived from these segmentations, Born-modeled wavefields were synthesized and used to quickly (in a single shot) image an isolated location from a key reflector. Qualitative and quantitative analysis suggested that the new model would produce a better-focused image; full migrations using both models confirmed that the updated model produced improved continuity in subsalt reflectors. Creation and testing of additional models would be straightforward, and could yield even better results.

## ACKNOWLEDGMENTS

I am grateful to Yaxun Tang for his contributions to the Born wavefield modeling framework, and to WesternGeco for providing the data and initial model used in the example.

## REFERENCES

- Felzenszwalb, P. F. and D. P. Huttenlocher, 2004, Efficient graph-based image segmentation: *International Journal of Computer Vision*, **59**, 167–181.
- Guerra, C., 2010, Migration-velocity analysis using image-space generalized wavefields: PhD thesis, Stanford University.

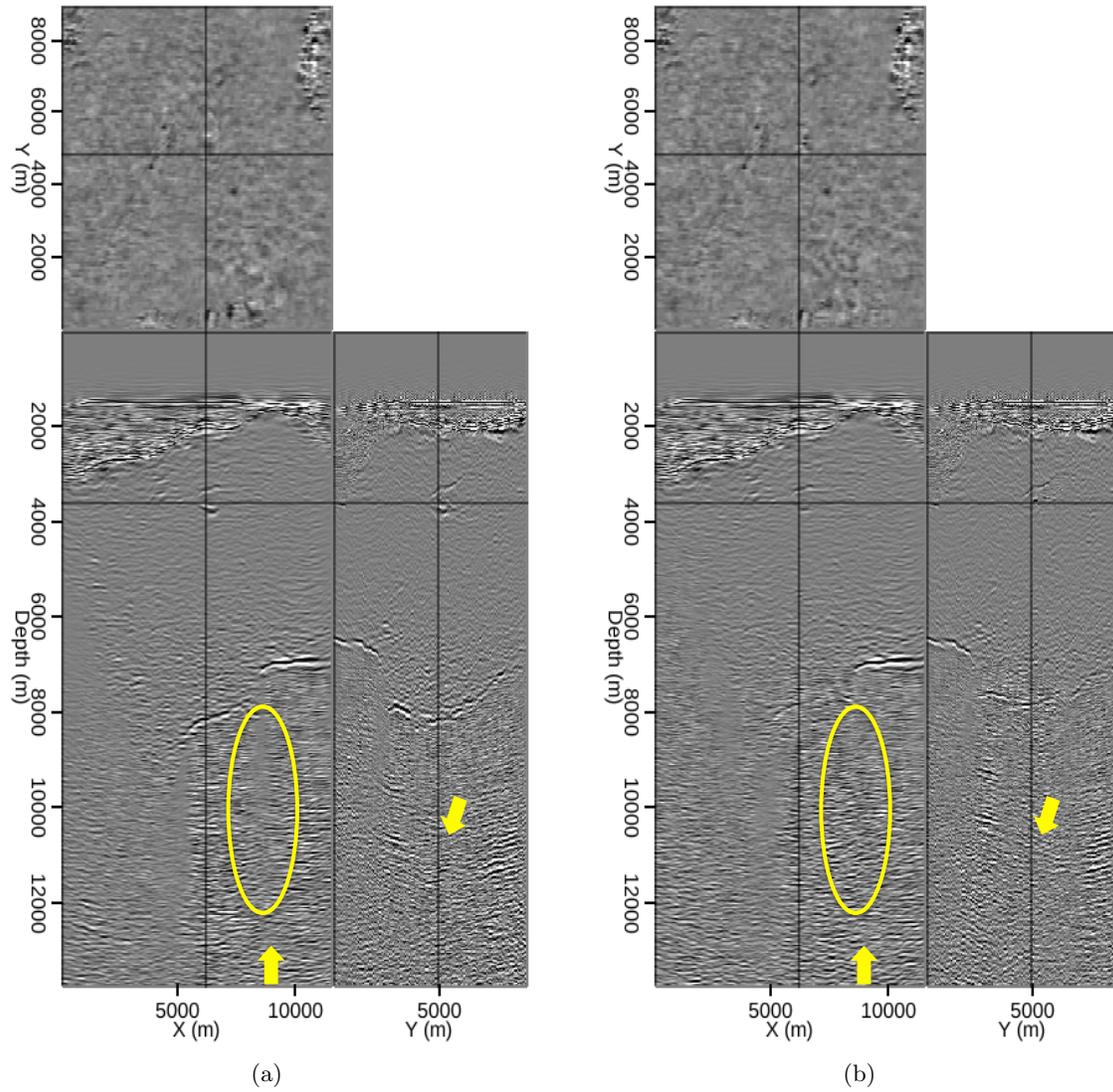


Figure 7: Full, one-way migrations with identical parameters using (a) the original velocity model; and (b) the updated model. Improvements in subsalt reflectors' continuity is apparent in (b), especially at the indicated locations. [CR] `adam1/. comp-origa,comp-newa`

- Halpert, A., 2010, A new method for more efficient seismic image segmentation: SEP-Report, **140**, 213–228.
- , 2013, Salt delineation via interpreter-guided 3D seismic image segmentation: SEP-Report, **149**, 107–120.
- Halpert, A. and Y. Tang, 2011, Velocity model evaluation through Born modeling and migration: a feasibility study: SEP-Report, **145**, 15–26.
- Hill, N. R., 1990, Gaussian beam migration: *Geophysics*, **55**, 1416–1428.
- Lomask, J., 2007, Seismic volumetric flattening and segmentation: PhD thesis, Stanford University.
- Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3D salt boundaries: *Geophysics*, **72**, P47–P56.
- Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: *Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence*, **22**, 838–905.
- Stolt, R. H. and A. Benson, 1986, *Seismic migration: Theory and practice*: Geophysical Press.
- Tang, Y., 2011, Imaging and velocity analysis by target-oriented wavefield inversion: PhD thesis, Stanford University.
- Tang, Y. and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEP-Report, **140**, 67–82.
- Wang, B., J. Ji, C. Mason, S. Gajawada, and Y. Kim, 2008, Beam-based interactive imaging for salt interpretation and salt model building: *SEG Technical Program Expanded Abstracts*, **27**, 3073–3077.
- Wang, B., C. Mason, K. Yoon, J. Ji, J. Cai, S. Suh, and Z. Li, 2011, Complex salt model building using a combination of interactive imaging and layer-stripping RTM: *First Break*, **29**, 47–54.



# Distance Regularized Level Set Salt Body Segmentation

*Taylor Dahlke*

## ABSTRACT

Segmentation of seismic images using a Distance Regularized Level Set Evolution (DRLSE) scheme maintains numerical stability of our implicit surface without the expense and accuracy issues associated with reinitialization approaches. In this work I apply the DRLSE algorithm to the Sigsbee salt model as well as an offshore salt data set. I then apply a modified energy functional which includes a Frobenius norm term that further improves the segmentation results. These applications of DRLSE demonstrate promising results using a very simplified energy functional.

## INTRODUCTION

Delineating the boundaries of salt bodies is important to sub-salt depth imaging as well as inversion schemes such as Full Waveform Inversion (FWI). Explicit parametrization approaches such as “Snakes” (Kass et al., 1988) and “Intelligent Scissors” (Mortensen and Barrett, 1998) methods encounter difficulty when changes of topology are necessary, or if there is a need to adapt to sharp corners and cusps on an object boundary. The implicit approach of the level set method avoids these issues (Lee, 2005), but requires a mathematically regular implicit surface to be successful (no sharp edges or “creases”). Standard level set evolution schemes lack an intrinsic regularization of the implicit surface itself. When the level set equation is derived from a variational approach, an energy functional is used which can accommodate a single or double-potential well term. Including this term maintains regularity of the implicit surface, without having to resort to reinitialization schemes as in standard level set algorithms. Furthermore, the variational approach to deriving the level set evolution equation allows for a conceptually straight-forward framework for adding additional terms to guide the surface evolution. In this work, I begin by analyzing the shortcomings of explicit parametrization approaches as well as the standard level set formulation. Next, I derive the DRLSE evolution equation using a variational approach. After that, I explain the use of a Frobenius norm term as particularly suited to salt body delineation. Last, I show the results of applying the Distance Regularized Level Set Evolution (DRLSE) algorithm to a salt model and salt data set, followed by a discussion of future work.

## LEVEL SET FUNDAMENTALS

Explicit methods of parametrizing curve boundaries expose us to a number of numerical problems. These include node distribution, evolving curves at sharp corners and cusps, as well the challenge of preventing evolving curve fronts from overlapping. By adding an extra dimension to the optimization problem, we can eliminate these issues, which is the main attraction behind using level set methods. In the case of segmenting a 2D object, level

set methodology expands the optimization space from a 2D (explicitly defined) curve to a 3D surface (which implicitly defines a curve). We evolve a surface  $\phi$  such that the curve represented by the contour at  $\phi = 0$  (the zero level set) is the boundary of the object that we wish to segment (Osher and Fedkiw, 2003). We choose force functions to evolve this surface in order to meet this end. These force functions are designed to conform the surface such that it delineates the segmentation object at  $\phi = 0$ . Equation 1 shows us the standard form of the level set equation, with  $F$  representing a directional force applied to the implicit surface  $\phi$ .

$$\frac{\partial \phi}{\partial t} = |F \nabla \phi| \quad (1)$$

Since this equation lacks an intrinsic means to maintain regularity, irregular or sharp features can form on the surface as it evolves, causing further evolution to be unstable. One approach to remediating this is to periodically reinitialize to a regular surface (Adalsteinsson and Sethian, 1995). When the implicit surface begins to become unstable (develops very sharp or flat shapes), the function is reinitialized by solving

$$\frac{\partial \phi}{\partial t} = \text{sign}(\phi_o)(1 - |\nabla \phi|) . \quad (2)$$

However, this approach is not always very accurate, especially when the implicit surface is not smooth or if there is a strong difference between the signed distance function and the surface being reinitialized. Furthermore, the question of how often reinitialization should be applied makes such approaches ad-hoc at best, and often expensive to implement well. For these reasons, the simplicity and robustness of deriving the evolution equation from a variational approach is a very attractive option.

## DISTANCE REGULARIZED LEVEL SET EVOLUTION

The DRLSE method is derived using a variational method. In our case, we intend to find the stationary value (minimum) of an energy functional. A standard approach to minimizing an energy functional is to find the steady state solution to the gradient flow equation

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi} . \quad (3)$$

From this, we substitute  $E$  with an energy functional that we minimize, as described in the next section. Because  $E$  is an energy functional derived using calculus of variation, the partial derivative shown in equation 3 becomes a Gâteaux derivative. The Gâteaux derivative is a generalization of directional derivative, and is used to formalize the functional derivative used in calculus of variations.

## External Energy

The energy functional that I use in this work can be represented very simply as the sum of a regularization term coupled with an external energy term

$$E(\phi) = R_p(\phi) + E_{ext}(\phi) . \quad (4)$$

The external energy term can be defined to include any term that utilizes information pertinent to the data of interest. We begin with a simple external energy term that is based on an active contour model, utilizing edge-based information, and composed of two terms that work in conjunction

$$E_{ext}(\phi) = \lambda L_g(\phi) + \alpha A_g(\phi) . \quad (5)$$

$L_g$  is an “edge” term acting to direct the curve evolution towards areas of strong contrast, while  $A_g$  is an “area” term that acts to accelerate the curve evolution by providing a ballooning or shrinking force (depending on how  $\alpha$  is chosen). The “edge” term computes a line integral of the function  $g$  along the zero level set contour,

$$L_g(\phi) = \int_{\Omega} g \delta_{\epsilon}(\phi) |\nabla \phi| dx , \quad (6)$$

while the “area” term (equation 7) computes a weighted area of the region inside the zero level set contour

$$A_g(\phi) = \int_{\Omega} g H(-\phi) dx . \quad (7)$$

Within the external energy functional terms (equations 6 and 7), I use a function  $g$  as the actual edge indicator

$$g = \frac{1}{1 + |\nabla(G_{\sigma} * I)|^2} . \quad (8)$$

In this case  $I$  is the amplitude of a pixel in the image we are segmenting, and  $G$  is a Gaussian smoothing function. By taking the gradient of this convolution,  $g$  tends to have smaller values at edge boundaries, acting as a very simple edge indicator function.

To detect the zero level set curve in our energy functional, it is necessary to construct a modified Dirac delta function that smoothly indicates the boundary

$$\delta_{\epsilon}(x) = \begin{cases} \frac{1}{2\epsilon}(1 + \cos(\frac{\pi x}{\epsilon})), & |x| \leq \epsilon \\ 0, & |x| > \epsilon \end{cases} . \quad (9)$$

We use  $\epsilon$  to set the window over which our smoothed Dirac delta function is applied.

For the area term as well as the Frobenius norm term (described later), we integrate  $g$  over the area enclosed by the zero level set. As with the Dirac delta function, I use a modified Heaviside function to smoothly indicate the boundary of the zero level set within a window defined by  $\epsilon$

$$H_{\epsilon}(x) = \begin{cases} \frac{1}{2}(1 + \frac{x}{\epsilon} + \frac{1}{\pi} \sin(\frac{\pi x}{\epsilon})), & |x| \leq \epsilon \\ 1, & x > \epsilon \\ 0, & x < -\epsilon \end{cases} . \quad (10)$$

## Regularization Term

To maintain numerical stability, we need to define the regularization term  $R_p$  such that it maintains the property of a function that is mathematically regular. For example, the signed distance function is very commonly used to initialize our level set algorithm because of its simplicity, but more importantly because it represents a regular implicit surface. Within the energy functional (equation 4) I include a regularization term in order to maintain the signed distance property of our surface as we evolve it

$$R_p(\phi) = \int_{\Omega} p(|\nabla\phi|)dx . \quad (11)$$

The signed distance function has the property of having a constant gradient (equal to one). For this reason, I construct a potential-well function

$$p_1(|\nabla\phi|) = \frac{1}{2}(|\nabla\phi| - 1)^2 , \quad (12)$$

such that  $R_p$  is minimized when  $|\nabla\phi| = 1$  (Li, 2010).

However, while using equation 12 allows us to maintain the signed distance property (and consequently regularity of the implicit surface), this formulation is not necessarily stable when we take the Gâteaux derivative of  $R_p$  and incorporate it into the gradient flow equation. Taking the Gâteaux derivative of equation 11, we get

$$\frac{\partial R_p}{\partial \phi} = -\text{div}(d_p(|\nabla\phi|)\nabla\phi) , \quad (13)$$

where  $\text{div}$  is the divergence operator,  $d_p$  is a function defined by

$$d_p(|\nabla\phi|) = \frac{p'(|\nabla\phi|)}{|\nabla\phi|} , \quad (14)$$

and  $p'$  is the first derivative of function  $p$ . With our choice of single-well potential function, equation 14 is simplified to  $d_p(|\nabla\phi|) = 1 - \frac{1}{|\nabla\phi|}$ . As a result, in the case where  $|\nabla\phi| = 0$ , our evolution becomes unstable since  $d_p = -\infty$ . For this reason, in order to avoid instability in areas where  $|\nabla\phi|$  gets close to zero, I modify the potential function to have two wells; one at  $|\nabla\phi| = 1$  as before, and a well at  $|\nabla\phi| = 0$ . This motivation leads us to define a double well equation

$$p_2(|\nabla\phi|) = \begin{cases} \frac{1}{(2\pi)^2}(1 - \cos(2\pi|\nabla\phi|)), & \text{if } |\nabla\phi| \leq 1 \\ \frac{1}{2}(|\nabla\phi| - 1)^2, & \text{if } |\nabla\phi| \geq 1 \end{cases} . \quad (15)$$

When the regularization and external energy terms are included together in equation 4, we derive the following complete energy functional

$$E(\phi) = \mu \int_{\Omega} p_2(|\nabla\phi|)dx + \lambda \int_{\Omega} g\delta_{\epsilon}(\phi) |\nabla\phi| dx + \alpha \int_{\Omega} gH_{\epsilon}(-\phi)dx . \quad (16)$$

The constants  $\mu$ ,  $\alpha$  and  $\lambda$  act as weights for each of the terms in our energy functional.

Combining the energy functional term above with the gradient flow equation (equation 3) that our derivation is based on, we get an equation that represents the evolution of our implicit surface, and can be directly implemented with a finite difference scheme

$$\frac{\partial \phi}{\partial t} = \mu \operatorname{div}(d_p(|\nabla \phi|) \nabla \phi) + \lambda \delta_\epsilon(\phi) \operatorname{div} \left( g \frac{\nabla \phi}{|\nabla \phi|} \right) + \alpha g \delta_\epsilon(\phi) . \quad (17)$$

### Use of the Frobenius Norm for Salt Delineation

One attribute of salt that can be useful in segmentation is the lack of stratification that is typically observed. A salt body as viewed in seismic data can usually be characterized as having a chaotic image gradient in the interior. One way that we can make use of this attribute in our DRLSE algorithm is by incorporating an appropriate term in our energy functional to quantify the chaos of the image gradient. By calculating the structural tensor in a neighborhood around each pixel

$$S(p) = \left\{ \begin{array}{cc} \sum_i I_{x_i} I_{x_i} & \sum_i I_{x_i} I_{y_i} \\ \sum_i I_{y_i} I_{x_i} & \sum_i I_{y_i} I_{y_i} \end{array} \right\} , \quad (18)$$

and then applying a variation of the Frobenius norm on said tensor, we are able to generate a map of values that represent the strength of stratification (Haukas, 2013).

In the representation shown in equation 18,  $I_{x_i}$  and  $I_{y_i}$  are the  $x$  and  $y$  gradient values of the image at index  $i$ , where  $i$  represents a pixel that exists in the neighborhood surrounding pixel  $p$ . For some pixel  $p$ , a structure tensor is generated by summing all  $m$  of the gradient multiplier elements, where  $m$  is equal to the number of terms in the neighborhood stencil used. In this case the neighborhood is a square image of size  $m = n \times n$  centered around pixel  $p$ .

At this point, we can either calculate the eigenvalues  $E(S(p))$  and perform the proper Frobenius norm

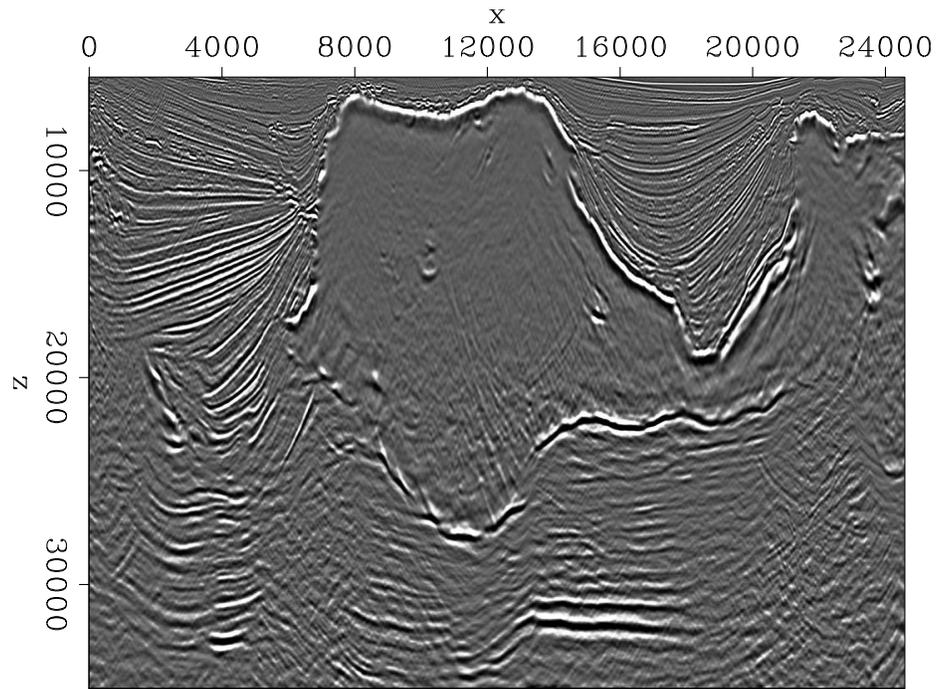
$$\|E(S(p))\|_{\text{Frob}} = \sqrt{\sum_{i=1}^2 \sum_{j=1}^2 |E(S(p))_{ij}|^2} , \quad (19)$$

or we can avoid the eigenvalue decomposition (and its added cost) and sum the square of the diagonal elements of  $S(p)$  as an approximate measure. I choose the latter approach for this work.

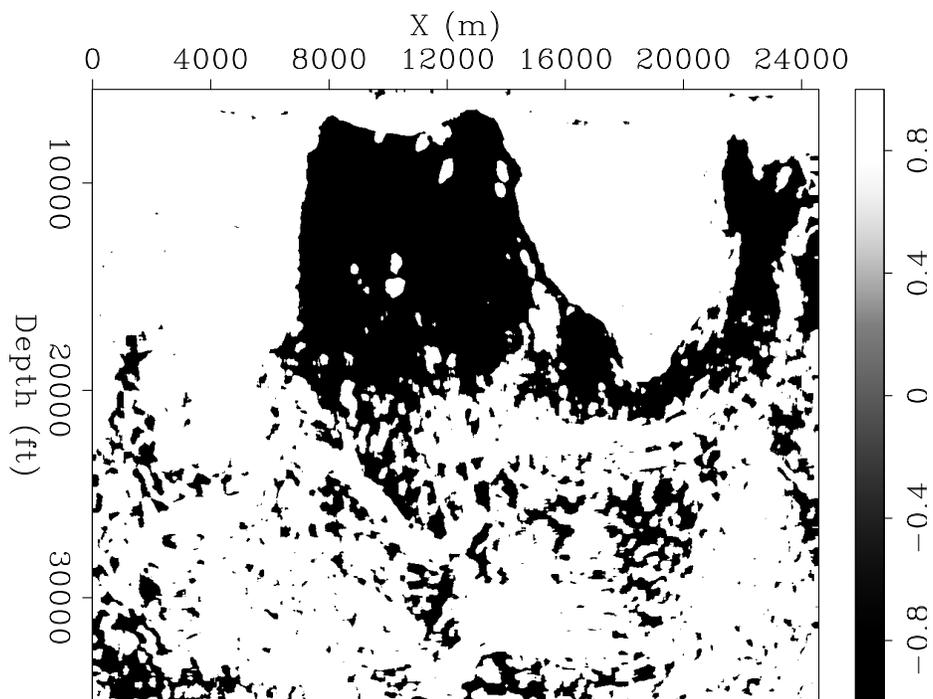
When we compare the stratification in the input data (Figure 1(a)) with the map of the Frobenius norm values (Figure 1(b)), we can see that the term does prove useful in identifying zones of salt.

## EXAMPLES AND DISCUSSION

The DRLSE algorithm converges quite well to the Sigsbee salt model (Figure 2(a)). This example demonstrates the effectiveness of the edge finding terms that our simple energy

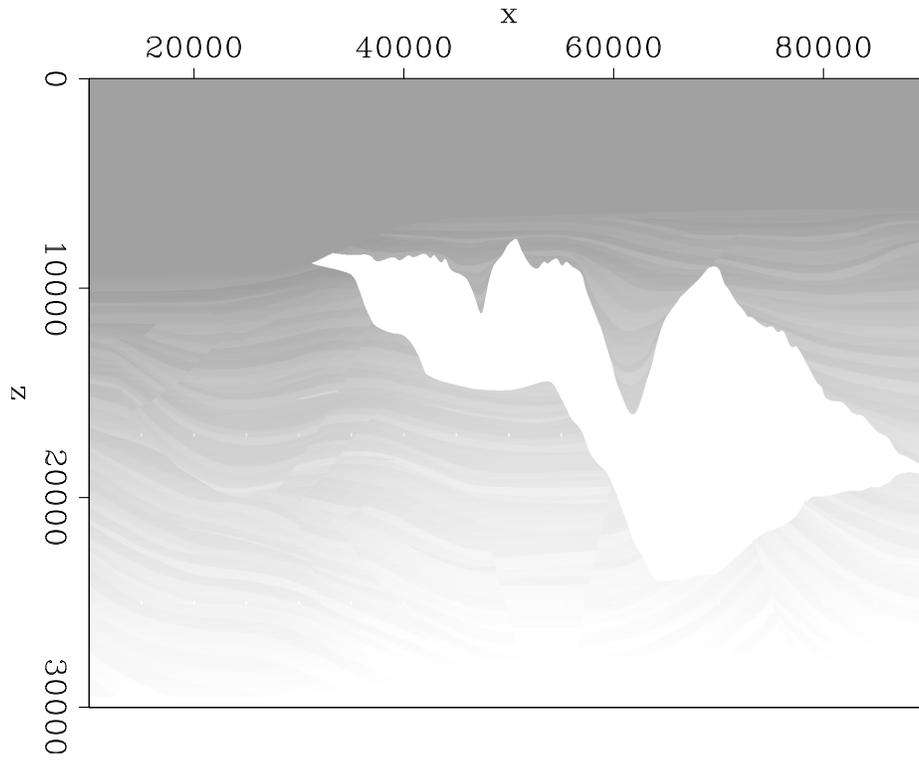


(a)

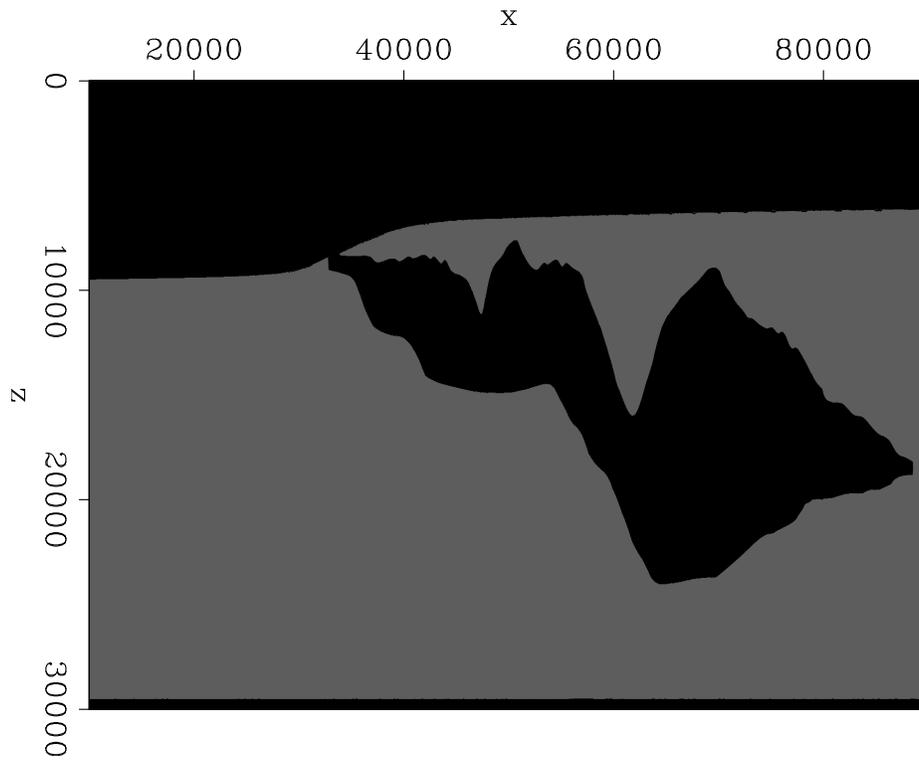


(b)

Figure 1: (a) Initial salt body data input (b) Map of Frobenius norm values, rescaled to binary representation. [ER] `taylor/. input-data,frob-map`



(a)



(b)

Figure 2: (a) Initial salt body model input (Sigsbee Model) (b) Level set output (without Frobenius norm). [ER] `taylor/. input-model,final-model`

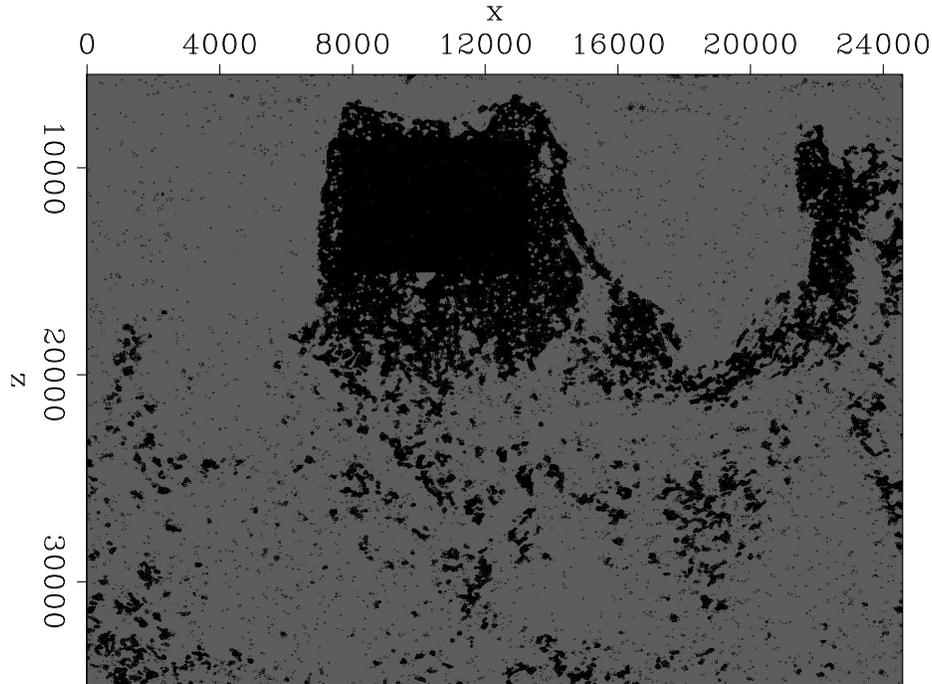


Figure 3: Level set output (with Frobenius norm term included). [ER] taylor/. final-data

functional includes. What can be noted is that the segmentation is unable to differentiate between the salt body (that was “selected” using the initial surface) and the water body layer above it. This energy functional as currently formulated doesn’t utilize the pixels within the initialized area to help determine the evolution of the level set. This is certainly an aspect that could be improved upon through incorporating an appropriate term in equation 4.

One obvious need in the results from this formulation is the lack of smoothness in the zero level set produced. Figure 3 demonstrates this lack of smoothness in the form of a “sprinkled” scattering of small segmented areas. Currently my research includes developing an effective smoothing term to add to the energy functional, namely by minimizing the curvature of the zero level set (Caselles et al., 1997). A term based on this attribute would mitigate the occurrence of “sprinkled” segmentation because these “sprinkled” regions contribute heavily to the total curvature, but are only a small part of the area within the zero level set.

A level set method, in particular the DRLSE method, is well suited to being incorporated into the work-flow of FWI. The gradient of the objective function can be used as a force to evolve the implicit surface (Lewis et al., 2012). When compared to other segmentation methods such as combinatorial graph cuts, the level set approach is more appropriate for FWI adaptation since it doesn’t require an input of hard constraints. The initialized surface that the algorithm uses as input can come from the initial FWI velocity model, or from a human input. However, these initial zero level sets are not hard constraints as in combinatorial graph cuts, and the accuracy of the input relative to the true segmentation is not as critical to a successful segmentation. Future extension of this work aims to incorporate the

DRLSE algorithm into a FWI work-flow for semi-automatic salt body segmentation.

## CONCLUSION

The level set method provides distinct advantages over explicit parametrization approaches when applied to seismic imaging segmentation problems. When it is derived with variational methods as a regularized algorithm, the problems associated with maintaining numerical stability in the implicit surface are ameliorated. Furthermore, this derivation provides us with a framework that allows us to further define our level set evolution by adding additional terms to our energy functional, as demonstrated by including a Frobenius norm term for salt delineation. Even with the very simple energy functional used in this work, the ability of the DRLSE algorithm to delineate salt bodies is evident in our results. Inclusion of additional external information, as well as smoothing terms, would further refine the salt body segmentation that this method can achieve.

## ACKNOWLEDGMENTS

I would like to thank Adam Halpert for his insight and suggestions, Biondo Biondi for his guidance on this project, and Robert Clapp for his technical expertise. I also would like to thank the sponsors of the Stanford Exploration Project for their financial support. Special thanks to SMARTJV and its partners for the Sigsbee model dataset, and to WesternGeco for the offshore salt dataset.

## REFERENCES

- Adalsteinsson, D. and J. A. Sethian, 1995, A fast level set method for propagating interfaces: *Journal of Computational Physics*, **118**, 269 – 277.
- Caselles, V., R. Kimmel, and G. Sapiro, 1997, Geodesic active contours: *Int. J. Comput. Vision*, **22**, 61–79.
- Haukas, Ravndal, F. B. S., 2013, Automated salt body extraction from seismic data using the level set method: *First Break*, **31**, 35–42.
- Kass, M., A. Witkin, and D. Terzopoulos, 1988, Snakes: Active contour models: *INTERNATIONAL JOURNAL OF COMPUTER VISION*, **1**, 321–331.
- Lee, C. P., 2005, Robust image segmentation using active contours: level set approaches.
- Lewis, W., B. Starr, and D. Vigh, 2012, A level set approach to salt geometry inversion in full-waveform inversion, *in* SEG Technical Program Expanded Abstracts 2012, 1–5, Society of Exploration Geophysicists.
- Li, Xu, G. F., 2010, Distance regularized level set evolution and its application to image segmentation: *IEEE Transactions on Image Processing*, **19**, 3243–3254.
- Mortensen, E. N. and W. A. Barrett, 1998, Interactive segmentation with intelligent scissors: *Graphical Models and Image Processing*, 349–384.
- Osher, S. and R. Fedkiw, 2003, Level set methods and dynamic implicit surfaces, volume **153**: Springer.



## SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-0463	ali
Barak, Ohad	723-9282	ohad
Biondi, Biondo	723-1319	biondo
Chang, Jason	724-2974	jason
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
Dahlke, Taylor	723-3187	taylor
de Ridder, Sjoerd	723-1250	sjoerd
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Le, Huy	723-3187	huyle
Leader, Chris	723-0463	chrisl
Levin, Stewart	726-1959	stew
Li, Elita	723-9282	myfusan
Maharramov, Musa	723-1250	musa
Shen, Xukai	723-0463	xukai
Shen, Yi	723-6006	yishen
Wong, Mandy	723-9282	mandyman
Zhang, Yang	723-6006	yang

**Dept fax number:** (650) 725-7344

## E-MAIL

Our Internet address is "*sep.stanford.edu*"; i.e., send Jon electronic mail with the address "*jon@sep.stanford.edu*".

## WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.

**STEERING COMMITTEE MEMBERS, 2012-2013**

<b>Name</b>	<b>Company</b>	<b>Tel #</b>	<b>E-Mail</b>
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Francois Audebert (Co-chair, 1st year)	TOTAL	–	francois.audebert@total.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Thomas Dickens	ExxonMobil	(713)431-6011	tom.a.dickens@exxonmobil.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Alejandro Valenciano (Co-chair, 2nd year)	PGS	–	alejandro.valenciano@pgs.com

## Research Personnel

**Ali Almomin** graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics and is currently working on seismic tomography. He is a member of SEG, EAGE, and SPE.



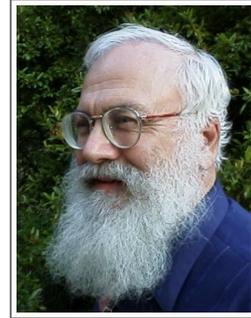
**Biondo L. Biondi** is professor of Geophysics at Stanford University. Biondo graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. He is co-director of the Stanford Exploration Project and of the Stanford Center for Computational Earth and Environmental Science. In 2004 the Society of Exploration Geophysicists (SEG) has honored Biondo with the Reginald Fessenden Award. Biondo recently published a book, 3-D Seismic Imaging, that is the first text book to introduce the theory of seismic imaging from the 3-D perspective. The book is published by SEG in the Investigations in Geophysics series. During 2007 gave a one-day short course in 28 cities around the world as the SEG/EAGE Distinguished Short Course Instructor (DISC). He is a member of AGU, EAGE, SEG and SIAM.



**Jason Chang** received his B.A. in geophysics from the University of California, Berkeley, in 2010. He joined SEP in autumn of 2011 and is currently working toward a Ph.D. in geophysics. He is a student member of SEG and AGU.



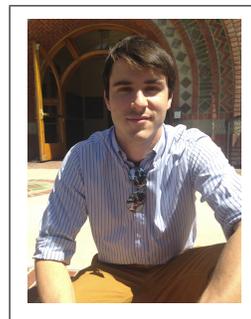
**Jon F. Claerbout** (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Emeritus 2008. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields*. Honorary member and SEG Fessenden Award “in recognition of his outstanding and original pioneering work in seismic wave analysis.” Founded the Stanford Exploration Project (SEP) in 1973. Elected Fellow of the American Geophysical Union. Authored three published books and five internet books. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG’s highest award. Honorary Member of the European Assn. of Geoscientists & Engineers (EAGE). EAGE’s highest recognition, the Erasmus Award.



**Robert Clapp** received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.



**Taylor Dahlke** is a second year student with SEP. He received his B.S. in civil engineering from the University of California, Berkeley in 2012, and joined SEP in July 2012. Currently, he is working towards a Ph.D. in geophysics with his research focused on applying levelset methodologies to perform salt body image segmentation. Taylor is a student member of SPE and Vice President of the Stanford SEG Student Chapter.



**Marine Denolle** received her Master in Geophysics from the Institut de Physique du Globe de Paris, France, in 2008. She started her PhD in the Department of Geophysics, School of Earth Sciences, Stanford in fall 2008 and will start her postdoctoral studies as a Green Scholar at Scripps, UC San Diego in winter 2014



**Antoine Guitton** received a M.Sc. in geophysics from Université de Strasbourg in 1996. He also received a M.Sc. and Ph.D in geophysics from Stanford University in 1998 and 2005. He is a recipient of the EAGE Arie van Weelden Award in 2004 and SEG J. Clarence Karcher Award in 2007. He received the Best Paper Presented by a Student at the Annual Meeting Award in 1999 for his work on robust norms. Antoine is currently working for Geoimaging Solutions Inc. as a senior research geophysicist working on imaging and velocity estimation. He is also a consulting professor in geophysics at Stanford University. He is a member of the EAGE, SEG and AGU.



**Adam Halpert** graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.



**Huy Le** graduated from the University of Oklahoma in 2012 with a B.S. in Geophysics. He has recently completed the Master program in Computational Geoscience, a joint program between the School of Earth Sciences and the Institute of Computational and Mathematical Engineering, and become a PhD student in Geophysics. His research interests include anisotropy, finite difference method, and full waveform inversion.



**Chris Leader** graduated from Oxford Univeristy in 2008 with a BA in Physics (with concentration on Astrophysics and Condensed Matter physics) and then from Imperial College London in 2009 with an MSc in Petroleum Geophysics (Distinction). He is currently a first year student in the Stanford Exploration Project on the PhD program working on Fourier methods of regularisation. Work experience involves 3D seismic processing for a Rio Tinto acquired dataset over summer 2009. He is a member of SEG, EAGE, PESGB and IOP.



**Stewart A. Levin** was acting director of the Stanford Exploration Project during Jon Claerbout's 1993-4 sabbatical year. After a distinguished career in industry at Mobil and Halliburton, he has returned to Stanford as a consulting professor in the Department of Geophysics.



**Yunyue (Elita) Li** graduated from China University of Petroleum, Beijing in July 2008 with a B.S. in Information and Computational Science. She joined SEP in the fall of 2008, and is currently working toward a Ph.D. in Geophysics. She is a student member of the SEG.



**Musa Maharramov** received his diploma c.l. in Mathematics from Baku State University in 1993. In 1995, he received a Ph.D. in Mathematics (Differential Equations) from the Azerbaijan Institute of Mathematics for his thesis “Asymptotic Solutions of Quasilinear Parabolic Equations”. In 1995-2001, he worked as an IT& Comms consultant, later IT manager, at Fluor Corporation in Baku. In 2001, he joined BP Azerbaijan, and in 2006 he started his work at BP High Performance Computing as a Computational Scientist. In that role he worked with the BP Imaging Team on mathematical, algorithmic and computational aspects of seismic migration and inversion. Musa joined SEP in the fall of 2011 and is pursuing his second Ph.D., in Geophysics. He is currently working on regularization of seismic inversion problems through the application of Geological/Geomechanical constraints and advanced optimisation techniques. Musa is a member of the SEG, EAGE and SIAM.



**Eileen Martin** graduated from the University of Texas at Austin in 2012 with a B.S. in Mathematics and Computational Physics. She is a Department of Energy Computational Science Graduate Fellow, and is working towards her Ph.D. in the Institute for Computational and Mathematical Engineering. She joined SEP in the summer of 2013, and is a member of AGU, SIAM and SEG.



**Gary Mavko** is professor of Geophysics at Stanford University. Gary received his B.S. in engineering physics from Cornell University, and his M.S. (1974) and a Ph.D. (1977) in geophysics from Stanford. Gary specializes in rock physics and is the co-author of "The Rock Physics Handbook". He is a member of AGU, EAGE, SEG, AAPG, and SPE.



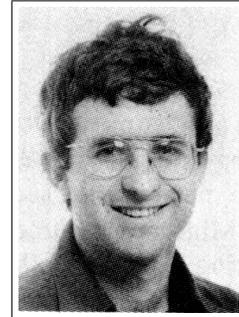
**Dave Nichols** received a B.A. in physics from Cambridge University in 1982; a M.Sc. in geophysics from Imperial College, London in 1983; and a Ph.D. in geophysics from Stanford University in 1994. From 1983 to 1987, he was employed by Western Geophysical in London and joined SEP in 1987. During the summer of 1991, he worked for Chevron Oil Field Research Company in La Habra, California. He currently works for Schlumberger in the Schlumberger Reservoir Completions Center in Rosharon, Texas.



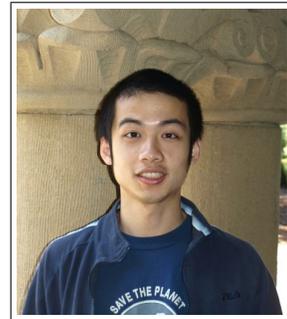
**Sjoerd de Ridder** hails from the Netherlands and received his BSc (2004) in Earth Sciences from Utrecht University. After an exchange year at Colorado School of Mines he attended Delft University of Technology. In Delft he worked on seismoelectric interferometry and obtained an MSc (2007) in Applied Earth Sciences with specialization in applied geophysics. In the fall of 2007 he started a PhD program at Stanford University to work on seismic-interferometry of low-frequency ambient seismic noise in SEP. He received the Jon F. Claerbout - Chevron Fellowship in 2009-2010. During his PhD program he spent summers at China University of Petroleum in Beijing (2010), BP America (2011), Chevron (2012) and BP Norway (2013). In his free time he likes to run, bike and hike while traveling and discussing philosophy, politics and history. In recent years he cycled over 3000 miles exploring China. Having finished his PhD in March of 2014 he started as an Associate Research Professor at the University of Science and Technology of China (USTC) in Hefei, Anhui.



**Shuki Ronen** is currently (2013) the chief geophysicist of Seabed Geosolutions, and a consulting faculty at SEP. The photo dates back to when he was a student at Stanford, 1985.



**Xukai Shen** graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



**Mandy Wong** graduated in 2004 with a B.Sc. in Physics and Mathematics from the University of British Columbia (UBC) in Vancouver, Canada. In 2006, she obtained a M.Sc. degree in Condensed Matter Theory at UBC. Afterward, Mandy worked for a geophysical consulting company, SJ Geophysics, based in Vancouver, Canada. Mandy joined SEP in 2008, and is working towards a Ph.D. in Geophysics. Her main research interest is imaging with multiples.



## SEP ARTICLES PUBLISHED OR IN PRESS

- Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis. SEG Technical Program Expanded Abstracts, **30**, 3887–3891.
- Almomin, A., and Biondi, B., 2012, Tomographic full waveform inversion: Practical and computationally feasible approach. SEG Technical Program Expanded Abstracts, 1–5.
- Almomin, A., and Biondi, B., 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations (Submitted).
- Ayeni, G., Tang, Y., and Biondi, B., 2010, Efficient seismic monitoring of hydrocarbon reservoirs using multiple shooting vessels: In Proceedings of the 2010 Offshore Technology Conference.
- Ayeni, G., and Biondi, B., 2010, Continuous reservoir monitoring with asynchronous simultaneous-source seismic data: 72nd Conference & Technical Exhibition, EAGE, Extended Abstracts.
- Ayeni G., and Biondi, B., 2010, Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets: Geophysics, **75**, no. 3, R61–R75.
- Ayeni, G., and Biondi, B., 2011, Wave-equation inversion of time-lapse seismic data sets: SEG Technical Program Expanded Abstracts (Submitted).
- Ayeni, G., Almomin, A., and Nichols, D., 2011, On the separation of simultaneous-source data by inversion: SEG Technical Program Expanded Abstracts (Submitted).
- Guerra, C., and Biondi, B., 2011, Fast 3D velocity updates using the pre-stack exploding reflector model: SEG Technical Program Expanded Abstracts **29**, 4375–4379.
- Guerra, C., and Biondi, B., 2010, Fast 3D velocity updates using the pre-stack exploding reflector model: Geophysics (Accepted).
- Biondi, B., 2010, Velocity estimation by image focusing analysis: Geophysics, **75**, pp. 49–60.
- Biondi, B., 2010, Velocity estimation by image focusing analysis: SEG Technical Program Expanded Abstracts **29**, 4066–4070.
- Chang, J., and de Ridder, S., 2012, Correlation energy between surface and borehole stations at Valhall: SEG Technical Program Expanded Abstracts **31** (Submitted).
- de Ridder, S., N. Crook, S.S. Haines, and S.T. Ide, 2010, Seismic investigation of underground coal fires: A feasibility study at the Southern Ute Nation Coal Fire site. Symposium on the Application of Geophysics to Engineering and Environmental Problems **23**, 630–638.
- de Ridder, S., and B. Biondi, 2010, Low-frequency passive seismic interferometry for land data: SEG Technical Program Expanded Abstracts **29**, 4041–4046.
- de Ridder, S., and J. Dellinger, 2011, Imaging of near-surface structures at Valhall using ambient-seismic-noise Eikonal tomography: EAGE Third Passive Seismic Workshop, Athens, Greece, PSP13.
- de Ridder, S., and J. Dellinger, 2011, Ambient seismic noise eikonal tomography for near-surface imaging at Valhall. The Leading Edge, **30**, no. 5, pp. 936–942.
- de Ridder, S., 2011, Ambient seismic noise tomography at Valhall. SEG Technical Program Expanded Abstracts **30**, 1597–1601.
- de Ridder, S., 2011, Ambient seismic noise tomography for exploration seismology at Valhall. Presented at 2011 AGU Fall Meeting, San Francisco, CA, 5-9 December 2011.
- Grobbe, N., F.C. Schoemaker, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Electrokinetic fields and waves: Theory, experiments, and numerical modeling. Geophysical Research Abstracts, Vol. **14**, EGU2012–10636.
- de Ridder, S., and B. Biondi, 2012, Continuous passive seismic monitoring of CCS projects

- by correlating seismic noise: A feasibility study. 74th Conference & Technical Exhibition, EAGE, Extended Abstracts (Accepted).
- de Ridder, S., and B. Biondi, 2012, Reservoir monitoring by passive seismic interferometry. SEG/SPE/AAPG Joint Workshop, La Jolla, CA, 24-29 June 2012 (Accepted).
- de Ridder, S., 2012, Ambient seismic noise correlations for reservoir monitoring. SEG Technical Program Expanded Abstracts **31** (Submitted).
- de Ridder, S., and B. Biondi, 2012, Continuous reservoir monitoring by ambient seismic noise tomography. SEG Technical Program Expanded Abstracts **31** (Submitted).
- Schoemaker, F.C., N. Grobbe, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Experimental validation of the electrokinetic theory and development of seismoelectric interferometry by cross-correlation: International Journal of Geophysics (Accepted).
- Dellinger, J., J. Yu, and S. de Ridder, 2010, Virtual-source interferometry of 4C OBC data at Valhall without a low-cut recording filter: SEG/EAGE Summer Research Workshop, Snowbird, Utah, 15-20 August 2010.
- Farghal, M.S., and Levin, S.A., 2012, Hunting for microseismic reflections using multiplets: SEG Technical Program Expanded Abstracts **31** (Submitted).
- Guitton, A., Ayeni, G., and Esteban, D.A., 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, *77*, No. 2, R117–R127.
- Guitton, A., and Esteban, D.A., 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: Geophysical Prospecting (In Press).
- Guitton, A., 2012, Blocky regularization schemes for full waveform inversion: Geophysical Prospecting (In Press).
- Halpert, A., Model-building with image segmentation and fast image updates: SEG Technical Program Expanded Abstracts **30**, 4035–4039.
- Halpert, A., Clapp, R.G., and B. Biondo, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts **29**, 1276–1280.
- Halpert, A., Clapp, R.G., and B. Biondo, 2011, Interpreter guidance for automated seismic image segmentation: 73rd Conference & Exhibition, EAGE, Extended Abstracts.
- Leader, C., Shen, X., and R. Clapp, 2011, Memory efficient reverse time migration in 3D: Proceedings of Nvidia GPU Technology Conference 2011/2012 (Accepted).
- Leader, C., and R. Clapp, 2012, Least squares reverse time migration on GPUs - balancing IO and computation: 74th Conference & Technical Exhibition, EAGE, Extended Abstracts (Accepted).
- Leader, C., and A. Almomin, 2012, How incoherent can we be? Phase encoded linearised inversion with random boundaries: SEG Technical Program Expanded Abstract **31** (Submitted).
- Li, Y., Y. Zhang, and J. Claerbout, 2010, Geophysical applications of a novel and robust L1 solver: SEG Technical Program Expanded Abstracts **29**, 3519–3523.
- Li, Y., and B. Biondi, 2011, Migration velocity analysis for anisotropic models: SEG Technical Program Expanded Abstracts **30**, 201–206.
- Li, Y., Nicoles, D., Osypov, K., and Bachrach, R., 2011, Anisotropic tomography with rock physics constraints: 73rd Conference & Technical Exhibition, EAGE, Extended Abstracts.
- Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: Geophysics **77**, 1–9.
- Li, Y., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: SEG Technical Program Expanded Abstract **31** (Submitted).
- Maharramov, M., 2011, Random boundary condition for low-frequency wave propagation:

- SEG Technical Program Expanded Abstracts **30**, 2962.
- Maharramov, M., and B. Nolte, 2011, Efficient one-way wave-equation migration in tilted transversally isotropic media: 73rd Conference & Technical Exhibition, EAGE, Extended Abstracts.
- Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Technical Program Expanded Abstracts **29**, 1975–1979.
- Shen, X., and R. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts (Submitted).
- Tang, Y., and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEG Technical Program Expanded Abstracts **29**, 4280–4285.
- Tang, Y., and B. Biondi, 2011, Target-oriented wavefield tomography using synthesized Born data: Geophysics (Accepted).
- Tang, Y., and B. Biondi, 2011, Subsalt velocity analysis by target-oriented wavefield tomography: A 3-D field-data example: SEG Technical Program Expanded Abstracts (Submitted).
- Wong, M., Biondi, B.L., and Ronen, S., 2010, Joint inversion of up- and down-going signal for ocean bottom data: SEG Technical Program Expanded Abstracts **29**, 2752–2756.
- Wong, M., and Ronen, S., and Biondi, B.L., 2011, Least-squares reverse-time migration/inversion for ocean bottom data: A case study. SEG Technical Program Expanded Abstracts **30**, 2369–2373.
- Wong, M., and Ronen, S., and Biondi, B.L., 2012, Joint imaging with streamer and ocean bottom data. SEG Technical Program Expanded Abstracts, 1–5.
- Wong, M., and Ronen, S., and Biondi, B.L., 2012, Imaging with multiples using linearized full-wave inversion. SEG Technical Program Expanded Abstracts, 1–5.
- Wong, M., 2013, Handling salt reflection in Least-squares RTM SEG Technical Program Expanded Abstracts, (submitted).
- Zhang, Y., Claerbout, J., Guitton, A., 2011, A new bidirectional deconvolution that overcomes the minimum phase assumption. 73rd Conference & Technical Exhibition, EAGE, Extended Abstracts.

