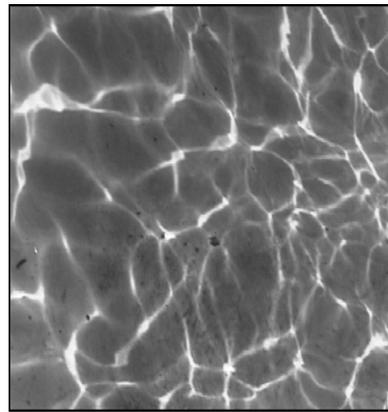
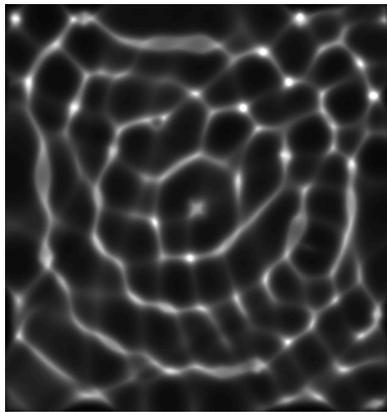


STANFORD EXPLORATION PROJECT

*Abdullah Al Theyab, Gboyega Ayeni, James Berryman, Biondo Biondi, David Chen,
Jon Claerbout, Robert Clapp, William Curry, Claudio Guerra, Antoine Guitton,
Adam Halpert, Simon Luo, Mohammad Maysami, Sjoerd de Ridder, Guojian Shan,
Xukai Shen, Jeff Shragge, Yaxun Tang, and Alejandro Valenciano*

Report Number 134, April 2008



Copyright © 2008

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 report¹ 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 library². We assume you have a UNIX workstation with Fortran, Fortran90, 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, 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 Fortran90 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 online³.

¹<http://sepwww.stanford.edu/private/docs/sep134>

²<http://sepwww.stanford.edu/public/docs/sepdata/lib/toc.html>

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

SEP134 — TABLE OF CONTENTS

Imaging Hessian

<i>Yaxun Tang</i> , Wave-equation Hessian by phase encoding	1
<i>Claudio Guerra and Alejandro Valenciano</i> , Reconciling processing and inversion: Multiple attenuation prior to wave-equation inversion	25
<i>Gboyega Ayeni and Biondo Biondi</i> , Time-lapse wave-equation inversion .	39

Velocity Analysis

<i>Biondo Biondi</i> , Automatic wave-equation migration velocity analysis . . .	65
<i>Claudio Guerra and Biondo Biondi</i> , Prestack exploding reflector modeling: The crosstalk problem	79
<i>Abdullah Al Theyab</i> , Ray tracing modeling and inversion of light intensity under a water surface	93
<i>Sjoerd de Ridder</i> , Stable simulations of illumination patterns caused by focusing of sunlight by water waves	105

Imaging with non-standard coordinates and sources

<i>Jeff Shragge</i> , Angle-domain common-image gathers in generalized coordinates	111
<i>Jeff Shragge and Guojian Shan</i> , 3D shot-profile migration in ellipsoidal coordinates	123
<i>Jeff Shragge</i> , Toward 3D conical-wave migration in tilted elliptic cylindrical coordinates	135
<i>Sjoerd de Ridder</i> , Spectral analysis of the non-proliferation experiment . .	147

Computational Interpretation

<i>Adam Halpert and Robert G. Clapp</i> , Image segmentation for velocity model construction and updating	159
<i>Robert G. Clapp</i> , Lloyd and Viterbi for QC and auto-picking	171
<i>Robert G. Clapp and David M. Chen and Simon Luo</i> , Hypercube viewer	179

Interpolation

<i>William Curry and Guojian Shan</i> , Interpolation of near offsets using multiples and prediction-error filters	193
--	-----

<i>Jon Claerbout and Antoine Guitton</i> , An algorithm for interpolation using Ronen's pyramid.....	221
<i>Xukai Shen</i> , Data interpolation in pyramid domain.....	225

Amplitudes and anisotropy

<i>James G. Berryman</i> , Biot-Gassmann analysis of partial and patchy saturated reservoirs for both reflected and transmitted seismic waves.....	231
<i>James G. Berryman</i> , Maximum entropy spectral analysis.....	245
<i>James G. Berryman</i> , Seismic anisotropy for polar media and an extended	267
<i>Mohammad Maysami</i> , Seismic reflector characterization by a multiscale detection-estimation method.....	275
SEP phone directory.....	285
Research personnel.....	287
('SEP article published or in press, 2008',).....	293

Wave-equation Hessian by phase encoding

Yaxun Tang

ABSTRACT

I present a method for computing wave-equation Hessian operators, also known as resolution functions or point-spread functions, under the Born approximation. The proposed method modifies the original explicit Hessian formula, enabling efficient computation of the operator. A particular advantage of this method is that it reduces or eliminates on-disk storage of Green's functions. The modifications, however, also introduce undesired cross-talk artifacts. I examine two different phase-encoding schemes, namely, plane-wave phase encoding and random phase encoding, to suppress the cross-talk. I apply the Hessian operator obtained by using random phase encoding to the Sigsbee2A synthetic data set, where a better subsalt image with higher resolution is obtained.

INTRODUCTION

Migration is an important tool for imaging subsurface structures using reflection seismic data. The classic imaging principle for shot-based migration states that reflectors are located where the forward-propagated source wavefield correlates with the backward-propagated receiver wavefield (Claerbout, 1971). However, this imaging principle is only the adjoint of the forward Born modeling operator (Lailly, 1983), which provides reliable structural information of the subsurface, but distorts the amplitude of the reflectors because of the non-unitary nature of the Born modeling operator. To improve relative amplitude behavior, the imaging problem can be formulated as an inverse problem based on the minimization of a least-squares functional. The inverse problem can be formulated either in the data space (Lailly, 1983; Tarantola, 1984; Nemeth et al., 1999; Clapp, 2005) or in the model space (Beylkin, 1985; Chavent and Plessix, 1999; Plessix and Mulder, 2004; Valenciano et al., 2006; Yu et al., 2006). The data-space approach can be solved iteratively using the gradient-based method (Nemeth et al., 1999; Clapp, 2005) without explicit construction of the Hessian, the matrix of the second derivatives of the error functional with respect to the model parameters. The iterative solving, however, is relatively costly and converges very slowly.

On the other hand, the model-space approach requires explicitly constructing the Hessian and applying its pseudo-inverse to the migrated image. The full Hessian of the least-squares functional is too big and expensive to be computed in practical applications; hence Chavent and Plessix (1999); Plessix and Mulder (2004) approximate it by a diagonal matrix. In the case of high-frequency asymptotics, and with an infinite aperture, the Hessian is diagonal in most cases (Beylkin, 1985). For a finite range of frequencies and limited acquisition geometry, however, the Hessian is no longer diagonal and not even diagonally dominated (Chavent and Plessix, 1999; Pratt et al., 1998; Plessix and Mulder, 2004; Valenciano et al., 2006). It has been shown by Valenciano et al. (2006) that, in areas of poor illumination,

e.g., subsalt regions, the Hessian's main diagonal energy is smeared along its off-diagonals. Therefore, the migrated image pre-multiplied by a diagonal matrix cannot perfectly recover the amplitude information, especially in poorly illuminated areas. That's why Valenciano et al. (2006) suggest computing a limited number of the Hessian off-diagonals to compensate for poor illumination and improve the inversion result. However, computing the Hessian off-diagonals, even for a limited number, is very expensive by direct implementation of the explicit Hessian formula. A huge number of Green's functions (easily several hundred terabytes for a typical 3-D survey) must be pre-computed and stored and then retrieved from the disk to generate the Hessian. Such operations not only require high-volume storage, but also high-speed I/O and network. Though computer speed continues to improve rapidly, computing such a huge matrix is still a challenge.

To make the Hessian computation more affordable, I describe a method based on the phase-encoding technique. In this method, the original explicit Hessian formula is slightly modified to enable efficient computation of the Hessian operator. The proposed method makes the Hessian computation similar to the shot-profile migration but with slightly modified imaging conditions. The new method eliminates the need to pre-compute and store Green's functions, but it also introduces cross-talk artifacts. I examine two phase-encoding schemes, plane-wave phase encoding and random phase encoding to attenuate the cross-talk.

This paper is organized as follows. First, I briefly review the theory of formulating the inverse problem in the model space. Then I discuss how the explicit Hessian can be efficiently computed using phase encoding. Finally, I apply the phase-encoded Hessian to the Sigsbee2A model, where a better subsalt image with higher resolution is obtained.

LEAST-SQUARES HESSIAN

Forward modeling using conventional two-way acoustic wave equation is a non-linear process that generates not only primaries, but also multiples. In imaging, however, multiples are often treated as noise and we mainly use primary information to image the subsurface. Primaries can be modeled using the Born approximation of the non-linear two-way wave equation by assuming a smooth background slowness field and small slowness perturbations. By making such assumptions, we can synthesize primaries by using the Born modeling equation as follows:

$$d(\mathbf{r}_k, \mathbf{s}_i, \omega) = \omega^2 \sum_{l=1}^L f_s(\omega) G_i(\mathbf{s}_i, \mathbf{x}_l, \omega) G_i(\mathbf{x}_l, \mathbf{r}_k, \omega) m(\mathbf{x}_l), \quad (1)$$

where $d(\mathbf{r}_k, \mathbf{s}_i, \omega)$ is the modeled seismic data for receiver location \mathbf{r}_k ($k = 1, \dots, M$) and shot location \mathbf{s}_i ($i = 1, \dots, N$) for a single frequency ω , $f_s(\omega)$ is the source signature. $m(\mathbf{x}_l)$ ($l = 1, \dots, L$) is the reflectivity at \mathbf{x}_l in the subsurface, $G_i(\mathbf{s}_i, \mathbf{x}_l, \omega)$ is the monochromatic Green's function from the source \mathbf{s}_i to the image point \mathbf{x}_l for the i th shot, and $G_i(\mathbf{x}_l, \mathbf{r}_k, \omega)$ is the monochromatic Green's function from the image point \mathbf{x}_l to the receiver \mathbf{r}_k for the i th shot.

Given the linear modeling equation 1, the adjoint equation is obtained by simply transposing and taking the complex conjugate of the modeling operator:

$$m_{mig}(\mathbf{x}_l) = \sum_{\omega} \omega^2 \sum_{i=1}^N f'_s(\omega) G'_i(\mathbf{s}_i, \mathbf{x}_l, \omega) \sum_{k=1}^M G'_i(\mathbf{x}_l, \mathbf{r}_k, \omega) d(\mathbf{r}_k, \mathbf{s}_i, \omega), \quad (2)$$

where $'$ stands for the complex conjugate. Equation 2 is the standard wave-equation shot-profile migration, which has been widely used in the industry. Migration positions the reflectors correctly in the subsurface. The amplitude information, however, is not well preserved: since the Born modeling operator (equation 1) is non-unitary, its adjoint, i.e., the migration operator (equation 2) is not the true inverse of the modeling operator. To better recover the amplitudes, we need to go beyond migration by formulating the imaging problem as an inverse problem, which enables us to find a generalized inverse of the modeling operator.

In using inverse theory, we first formulate an objective function in the least-square sense as follows:

$$F(\mathbf{m}) = \frac{1}{2} \sum_{\omega} \sum_{k=1}^M \sum_{i=1}^N |d(\mathbf{r}_k, \mathbf{s}_i, \omega) - d_{obs}(\mathbf{r}_k, \mathbf{s}_i, \omega)|^2, \quad (3)$$

where $d(\mathbf{r}_k, \mathbf{s}_i, \omega)$ represents the modeled data by using equation 1, and $d_{obs}(\mathbf{r}_k, \mathbf{s}_i, \omega)$ represents the observed data. Now the goal is to find a model $\mathbf{m} = (m(\mathbf{x}_1), \dots, m(\mathbf{x}_L))$ that best fits the recorded data \mathbf{d}_{obs} in the least-square sense. Equation 3 can also be written in a more compact form as follows:

$$F(\mathbf{m}) = \|\mathbf{d} - \mathbf{d}_{obs}\|_2 = \|\mathbf{L}\mathbf{m} - \mathbf{d}_{obs}\|_2, \quad (4)$$

where \mathbf{L} is the forward modeling operator defined in equation 1, and $\|\cdot\|_2$ stands for the L_2 norm. Equation 4 can be solved directly using the gradient-based methods (Nemeth et al., 1999; Clapp, 2005). The main obstacle to its wide use is its cost; three wavefield propagations are required within each iteration, making it impractical for large-scale problems. Another drawback is its relatively slow convergence rate, since only the gradient information of the objective function F is used in the optimization process.

Instead of solving equation 4 directly in the data space, we can reformulate it and solve it in the model space. Because \mathbf{L} is a linear operator, F is a quadratic function. Its minimum is obtained when \mathbf{m} satisfies:

$$\mathbf{m} = \mathbf{H}^{-1}\mathbf{L}'\mathbf{d}_{obs}, \quad (5)$$

where $\mathbf{H} = \mathbf{L}'\mathbf{L}$ is the Hessian operator. The Hessian provides the curvature information of the objective function F . We can rewrite equation 5 by recognizing that $\mathbf{L}'\mathbf{d}_{obs}$ is the migrated image \mathbf{m}_{mig} (equation 2)

$$\mathbf{H}\mathbf{m} = \mathbf{m}_{mig}. \quad (6)$$

Now we can formulate our new objective function as follows:

$$J(\mathbf{m}) = \|\mathbf{H}\mathbf{m} - \mathbf{m}_{mig}\|_2. \quad (7)$$

Equation 7 says that once we explicitly compute the Hessian operator \mathbf{H} , we can use the linear conjugate-gradient or any other method to minimize the new objective function J defined in equation 7.

Equation 7 is mathematically equivalent to equation 3, with the only difference being that the former is formulated in the model space, while the latter in the data space. The main advantage of model-space approach is that we can invert the model parameters in a target-oriented fashion (Valenciano et al., 2006), which substantially reduces the size of

the problem. For example, we can choose to invert only areas of particular interest, such as subsalt regions, where reservoirs are often located. In fact, the model-space approach divides the original problem defined by equation 3 into two stages: computing the explicit Hessian and minimizing the objective function J . Since minimizing the objective function J is trivial once the Hessian is obtained, various regularization schemes can be easily tried to improve the inversion result without any additional cost. This is another important advantage of formulating inverse problem in the model space.

The Hessian can be obtained by taking the second derivative of the objective function F with respect to the model parameters, the formula is given as follows (see Appendix A for derivation):

$$H(\mathbf{x}_p, \mathbf{x}_q) = \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G'_i(\mathbf{s}_i, \mathbf{x}_q, \omega) \sum_{k=1}^M G_i(\mathbf{r}_k, \mathbf{x}_p, \omega) G'_i(\mathbf{r}_k, \mathbf{x}_q, \omega) \quad (8)$$

Hereafter, I call $H(\mathbf{x}_p, \mathbf{x}_q)$ in equation 8 the exact Hessian, since it is derived strictly from the least-squares functional F . It is cumbersome and very expensive to compute the exact Hessian by directly implementing the above equation, because large storage of Green's function is needed. In the subsequent sections, I introduce an alternative method based on phase encoding for computing the Hessian operator, which reduces the computational requirement. As I will demonstrate, by using this approach, we do not need to save any Green's functions, and the cost for computing the Hessian is also significantly reduced.

PHASE-ENCODED HESSIAN

Encoding of the receiver-side Green's functions

Suppose we have the following function:

$$\tilde{H}(\mathbf{x}_p, \mathbf{x}_q) = \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G'_i(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \sum_{u=1}^M G_i(\mathbf{r}_u, \mathbf{x}_p, \omega) \alpha_u(\omega) \sum_{v=1}^M G'_i(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha'_v(\omega), \quad (9)$$

where we introduce an extra summation $\sum_{v=1}^M$ for the receiver-side Green's functions; $\alpha_u(\omega)$ and $\alpha'_v(\omega)$ are some weighting functions to be specified later. Equation 9 can be directly implemented by storing the Green's functions, just as with the direct implementation of equation 8. However, equation 9 offers more flexibility and can be very efficiently implemented without explicitly saving the Green's functions.

With the extra summation, the term $\sum_{u=1}^M G_i(\mathbf{r}_u, \mathbf{x}_p, \omega) \alpha_u(\omega)$ can now be seen as the extrapolated wavefield at the image point \mathbf{x}_p , using the composite source as the source function ($f_c(\mathbf{r}_u, \omega) = \alpha_u(\omega)$, $u = 1, \dots, M$). The same thing holds for the other summation term $\sum_{v=1}^M G'_i(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha'_v(\omega)$, except that it is the complex conjugate of the extrapolated wavefield at the image point \mathbf{x}_q . To make it clearer, we define the receiver wavefield $R_i(\mathbf{x}, \omega)$

corresponding to the receiver composite source for the i th shot gather as

$$R_i(\mathbf{x}, \omega) = \sum_{u=1}^M G_i(\mathbf{r}_u, \mathbf{x}, \omega) \alpha_u(\omega), \quad (10)$$

and the source wavefield corresponding to a point source located at \mathbf{s}_i on the surface as

$$S_i(\mathbf{x}, \omega) = f_s(\omega) G_i(\mathbf{s}_i, \mathbf{x}, \omega). \quad (11)$$

Substituting equations 10 and 11 into equation 9 leads to

$$\tilde{H}(\mathbf{x}_p, \mathbf{x}_q) = \sum_{\omega} \omega^4 \sum_{i=1}^N S_i(\mathbf{x}_p, \omega) S_i'(\mathbf{x}_q, \omega) R_i(\mathbf{x}_p, \omega) R_i'(\mathbf{x}_q, \omega), \quad (12)$$

which means $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q)$ can be computed by cross-correlating the source and receiver wavefields with their shifted complex conjugates (\mathbf{x}_q is the neighborhood point around \mathbf{x}_p). It is similar to the wave-equation shot-profile migration process, except the imaging condition is slightly modified. In other words, we do not have to save the Green's functions. Instead, we can replace each shot gather with composite sources, then extrapolate it into the subsurface and use the imaging condition defined in equation 12 to generate $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q)$. This process is efficient because multiple Green's functions are computed at the same time during the wavefield extrapolation.

It can be shown that after some simple algebraic manipulation, equation 9 can be rewritten as follows:

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G_i'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{k=1}^M G_i(\mathbf{r}_k, \mathbf{x}_p, \omega) G_i'(\mathbf{r}_k, \mathbf{x}_q, \omega) |\alpha_k(\omega)|^2 + \\ &\quad \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G_i'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{u=1}^M \sum_{v=1(u \neq v)}^M G_i(\mathbf{r}_u, \mathbf{x}_p, \omega) G_i'(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha_u(\omega) \alpha_v'(\omega). \end{aligned} \quad (13)$$

If we let the weighting function $\alpha_k(\omega)$ satisfy $|\alpha_k(\omega)| = 1$, the first term in equation 13 becomes the exact Hessian $H(\mathbf{x}_p, \mathbf{x}_q)$; however, the second term is the undesired cross-talk from the cross-correlations among different receiver-side Green's functions.

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= H(\mathbf{x}_p, \mathbf{x}_q) + \\ &\quad \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G_i'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{u=1}^M \sum_{v=1(u \neq v)}^M G_i(\mathbf{r}_u, \mathbf{x}_p, \omega) G_i'(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha_u(\omega) \alpha_v'(\omega) \end{aligned} \quad (14)$$

Simultaneous encoding of the source-side and receiver-side Green's functions

We can further encode the source-side Green's function by synthesizing composite sources from the source locations. For simplicity, we assume ocean bottom cable (OBC) or land acquisition geometry, where all the shots share the same receiver array. Therefore we have the following relations:

$$G_i(\mathbf{s}_i, \mathbf{x}, \omega) = G(\mathbf{s}_i, \mathbf{x}, \omega) \quad (15)$$

$$G_i(\mathbf{r}_k, \mathbf{x}, \omega) = G(\mathbf{r}_k, \mathbf{x}, \omega). \quad (16)$$

Then the following function can be constructed:

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) \beta_i(\omega) \sum_{j=1}^N G'(\mathbf{s}_j, \mathbf{x}_q, \omega) \beta'_j(\omega) \times \\ &\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) \alpha_u(\omega) \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha'_v(\omega), \end{aligned} \quad (17)$$

where we introduce two extra summations: $\sum_{j=1}^N$ for the source-side Green's function and $\sum_{v=1}^M$ for the receiver-side Green's function. Let us once again define the composite source wavefield $S(\mathbf{x}, \omega)$ and composite receiver wavefield $R(\mathbf{x}, \omega)$ as follows:

$$S(\mathbf{x}, \omega) = f_s(\omega) \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}, \omega) \beta_i(\omega), \quad (18)$$

and

$$R(\mathbf{x}, \omega) = \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}, \omega) \alpha_u(\omega). \quad (19)$$

Substituting equations 18 and 19 into equation 17 leads to

$$\tilde{H}(\mathbf{x}_p, \mathbf{x}_q) = \sum_{\omega} \omega^4 S(\mathbf{x}_p, \omega) S'(\mathbf{x}_q, \omega) R(\mathbf{x}_p, \omega) R'(\mathbf{x}_q, \omega). \quad (20)$$

Equation 20 tells us that to compute the simultaneously encoded Hessian $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q)$ for OBC or land acquisition geometry, only two wavefield propagations are required: one for the composite source wavefield defined by equation 18, and the other for the composite receiver wavefield defined by equation 19. We may rewrite equation 17 as

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) |\beta_i(\omega)|^2 \times \\ &\quad \sum_{k=1}^M G(\mathbf{r}_k, \mathbf{x}_p, \omega) G'(\mathbf{r}_k, \mathbf{x}_q, \omega) |\alpha_k(\omega)|^2 + \\ &\quad \sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N \sum_{j=1, i \neq j}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_j, \mathbf{x}_q, \omega) \beta_i(\omega) \beta'_j(\omega) \times \\ &\quad \sum_{u=1}^M \sum_{v=1, u \neq v}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha_u(\omega) \alpha'_v(\omega). \end{aligned} \quad (21)$$

If we can choose proper weighting functions $\alpha_k(\omega)$ and $\beta_i(\omega)$ such that $|\alpha_k(\omega)| = 1$ and $|\beta_i(\omega)| = 1$, then the first term in equation 21 becomes the exact Hessian $H(\mathbf{x}_p, \mathbf{x}_q)$, and the second term becomes the undesired cross-talk:

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= H(\mathbf{x}_p, \mathbf{x}_q) + \\ &\sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N \sum_{j=1(i \neq j)}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_j, \mathbf{x}_q, \omega) \beta_i(\omega) \beta_j'(\omega) \times \\ &\sum_{u=1}^M \sum_{v=1(u \neq v)}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \alpha_u(\omega) \alpha_v'(\omega). \end{aligned} \quad (22)$$

From equations 14 and 22, we face a similar situation encountered in phase-encoding migration (Romero et al., 2000); i.e., our exact Hessian is contaminated by cross-talk artifacts, so we seek to define weighting functions $\alpha_k(\omega)$ and $\beta_i(\omega)$ that attenuate the cross-talk as much as possible. In the next two sections, I examine two different phase-encoding schemes to attenuate that cross-talk, namely, plane-wave phase encoding and random phase encoding.

PLANE-WAVE PHASE ENCODING

Let us first encode only the receiver-side Green's functions. Suppose we choose the weighting functions to be

$$\alpha_k(\omega) = \alpha_k(p_r, \omega) = A_r(\omega) e^{i\omega p_r (\mathbf{r}_k - \mathbf{r}_0)}, \quad (23)$$

which is the known plane-wave phase-encoding function (Whitmore, 1995; Liu et al., 2006), where $A_r(\omega)$ is a real function depending upon the angular frequency ω , $i = \sqrt{-1}$; the vector \mathbf{r}_0 is some reference location, and p_r is the ray-parameter at the surface. Here we limit our discussion to 2-D, so p_r is a scalar, but it is very straightforward to extend the analysis to 3-D. As proved by Liu et al. (2006), summing an infinite number of plane waves ranging from $p_r = -\infty$ to $+\infty$ completely attenuates the cross-talk in the plane-wave migration, and the final migration result is exactly equivalent to the standard shot-profile migration result. The same property holds here in the scenario of Hessian computation, as proved in Appendix B: stacking over p_r and choosing $A_r(\omega)$ to satisfy $|\omega|^{-1} A_r^2(\omega) = 1$, the approximate Hessian $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r)$ converges to the exact Hessian $H(\mathbf{x}_p, \mathbf{x}_q)$:

$$\sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r) = H(\mathbf{x}_p, \mathbf{x}_q). \quad (24)$$

Similarly, in the case of simultaneous encoding, we can choose the weighting functions $\alpha_k(\omega)$ and $\beta_i(\omega)$ to be

$$\alpha_k(\omega) = \alpha_k(p_r, \omega) = A_r(\omega) e^{i\omega p_r (\mathbf{r}_k - \mathbf{r}_0)} \quad (25)$$

$$\beta_i(\omega) = \beta_i(p_s, \omega) = A_s(\omega) e^{i\omega p_s (\mathbf{s}_i - \mathbf{s}_0)}, \quad (26)$$

where p_r and p_s are the receiver-side and source-side plane-wave ray parameters. Once again, the analysis is limited to 2-D, so p_r and p_s are scalars. The vectors \mathbf{r}_0 and \mathbf{s}_0

are some reference locations for receiver and source plane waves respectively. $A_r(\omega)$ and $A_s(\omega)$ are real functions. Similar to the analysis in Appendix B, Appendix C proves that by stacking over p_s and p_r and choosing $A_s(\omega)$ and $A_r(\omega)$ to satisfy $|\omega|^{-1}A_s^2(\omega) = 1$ and $|\omega|^{-1}A_r^2(\omega) = 1$, the final stacking result is equivalent to the exact Hessian:

$$\sum_{p_s=-\infty}^{+\infty} \sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_s, p_r) = H(\mathbf{x}_p, \mathbf{x}_q). \quad (27)$$

RANDOM PHASE ENCODING

Instead of using the plane-wave encoding phase function, we can use random phases to disperse unwanted cross-terms. By doing so, we do not need to stack over p_r or p_s to attenuate the cross-talk, so the cost is significantly reduced. To encode only the receiver-side Green's function, we can choose the weighting function in equation 13 to be

$$\begin{aligned} \alpha_u(\omega) &= e^{\iota\gamma_u(\omega)} \\ \alpha_v(\omega) &= e^{\iota\gamma_v(\omega)}. \end{aligned}$$

Since $|e^{\iota\gamma_u(\omega)}| = 1$, then equation 13 becomes

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= H(\mathbf{x}_p, \mathbf{x}_q) + \\ &\sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G_i(\mathbf{s}_i, \mathbf{x}_p, \omega) G'_i(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\sum_{u=1}^M \sum_{v=1(u \neq v)}^M G_i(\mathbf{r}_u, \mathbf{x}_p, \omega) G'_i(\mathbf{r}_v, \mathbf{x}_q, \omega) e^{\iota(\gamma_u(\omega) - \gamma_v(\omega))}. \end{aligned} \quad (28)$$

Suppose we choose the phase function $\gamma_r(\omega) = \gamma_u(\omega) - \gamma_v(\omega)$ to be a sequence of random numbers between 0 and 2π . When we sum over ω to generate the final result, the phases will not agree, and consequently, the cross-talk will be attenuated. To maximize the phase differences from each frequency, a uniformly distributed random sequence could be used.

Similarly, for the case of simultaneous encoding, we choose the weighting functions in equation 21 to be

$$\begin{aligned} \alpha_u(\omega) &= e^{\iota\gamma_u(\omega)} \\ \alpha_v(\omega) &= e^{\iota\gamma_v(\omega)} \\ \beta_i(\omega) &= e^{\iota\gamma_i(\omega)} \\ \beta_j(\omega) &= e^{\iota\gamma_j(\omega)}. \end{aligned}$$

Then equation 21 becomes

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q) &= H(\mathbf{x}_p, \mathbf{x}_q) + \\ &\sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N \sum_{j=1(i \neq j)}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_j, \mathbf{x}_q, \omega) e^{\iota(\gamma_i(\omega) - \gamma_j(\omega))} \times \\ &\sum_{u=1}^M \sum_{v=1(u \neq v)}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) G'(\mathbf{r}_v, \mathbf{x}_q, \omega) e^{\iota(\gamma_u(\omega) - \gamma_v(\omega))}. \end{aligned} \quad (29)$$

By choosing phase functions $\gamma_s(\omega) = \gamma_i(\omega) - \gamma_j(\omega)$ and $\gamma_r(\omega) = \gamma_u(\omega) - \gamma_v(\omega)$ to be uniformly distributed random sequences, we maximize the attenuation of the cross-talk in equation 29 when stacking over frequencies.

COST COMPARISON

In this section, I compare the cost for different methods to have an idea of the savings generated by using the phase-encoding method. As discussed before, Hessian computation mainly contains two parts: wavefield propagation (Green's functions) and cross-correlation among different Green's functions. Since the cross-correlation parts are similar for methods with or without phase encoding, I will compare the cost for the first part, i.e, wavefield propagation for Green's functions.

Let us assume a 3-D seismic survey that has N_s shots in total, covering a surface area that can be divided into $N_x \times N_y$ cells. Table 1 illustrates the number of wavefield propagations and the size of Green's functions for an image cube with the size $N'_x \times N'_y \times N'_z$ for different methods. N_ω is the number of frequencies of the wavefield, $N_{p_{sx}}$ and $N_{p_{sy}}$ are the source ray parameters for the in-line and cross-line directions, and $N_{p_{rx}}$ and $N_{p_{ry}}$ are the receiver ray parameters for the in-line and cross-line directions respectively. For simultaneous encoding methods, OBC or land recording geometry has been assumed.

Table 1: Comparison of the costs for different methods

Method	Wavefield propagations	Size of Green's functions
direct computation	$N_x N_y N_\omega$	$N_x N_y N'_x N'_y N'_z N_\omega$
plane-wave, receiver-side	$(1 + N_{p_{rx}} N_{p_{ry}}) N_s N_\omega$	0
plane-wave, simultaneous	$(N_{p_{sx}} N_{p_{sy}} + N_{p_{rx}} N_{p_{ry}}) N_\omega$	0
random, receiver-side	$2 N_s N_\omega$	0
random, simultaneous	$2 N_\omega$	0

Considering that $N_{p_{sx}}, N_{p_{sy}}, N_{p_{rx}}$ and $N_{p_{ry}}$ are usually much smaller than N_x and N_y in practice, the phase-encoding methods generally need far fewer wavefield propagations. The most efficient method would be the simultaneous random encoding method, which requires only $2N_\omega$ wavefield propagations; simultaneous plane-wave encoding is also relatively efficient, considering its ability to get rid of the cross-talk. If the acquisition geometry is not OBC or land geometry, it would be efficient to use the receiver-side random encoding method; its cost is comparable to that of the shot-profile migration. Besides needing fewer wavefield propagations, phase-encoding methods do not require to save any Green's functions on the hard disk, which is good news for practical applications, since the size of the Green's functions ($N_x N_y N'_x N'_y N'_z N_\omega$) by direct computation can easily reach a ridiculously large number.

NUMERICAL EXAMPLES

In this section, I show several numerical examples for two different velocity models. The first is a simple constant-velocity model, which I use to verify the proposed algorithms for Hessian computation; the second is the more complicated Sigsbee2A velocity model.

Because of the complex salt body and limited acquisition geometry, there are shadow zones under the salt where conventional wave-equation migration algorithms often fail to produce reliable images. Therefore, it is useful to show how inversion by using the phase-encoded Hessian can improve the final image.

Verification of the algorithm: A constant velocity model

I verify the proposed phase-encoding algorithms on a constant velocity model ($v = 2000\text{m/s}$) for different acquisition geometries. Figure 1 shows the diagonal of the Hessian (when $\mathbf{x}_p = \mathbf{x}_q$). The diagonal of the Hessian is computed using different methods, with an acquisition geometry containing only one shot located at -600m and two receivers located at 600m and 1200m on the surface. The frequency band of the seismic experiment is from 5 Hz to 35 Hz , and all the frequencies are used to generate the following results. Figure 1(a) is the exact diagonal of Hessian $H(\mathbf{x}_p, \mathbf{x}_p)$ uncontaminated by any cross-talk artifacts. Figure 1(b) is obtained by directly implementing equation 9. No weighting functions are applied ($\alpha_k(\omega) = 1$), so the result is contaminated by strong cross-talk, which can be easily identified as the vertical stripes on the right side of the image. Figure 1(c) is obtained by using plane-wave phase encoding after stacking 61 receiver-side-encoded plane waves. It successfully removes the cross-talk, and Figure 1(c) looks very similar to Figure 1(a). Figure 1(d) shows the result obtained by using the random phase encoding method. The cross-talk is greatly reduced by the random phase functions, but is not completely attenuated. As I show later, the result of random phase encoding can be improved by stacking more receivers or shots.

Figure 2 shows the Hessian with off-diagonals (with size 21×21) at the image point $x = 680\text{m}$, $z = 800\text{m}$. The acquisition geometry is the same as that in Figure 1. The horizontal and vertical axes in Figure 2 show the horizontal and vertical offsets away from that image point. Figure 2(a) is the exact Hessian; Figure 2(b) shows the Hessian with strong cross-talk; Figures 2(c) and (d) show the results obtained with plane-wave phase encoding and random phase encoding respectively. Both phase-encoding schemes successfully remove the cross-talk shown in Figure 2(b).

In the next example, I slightly change the acquisition geometry used in the previous example from two receivers to 401 receivers. The receivers range from -2000m to 2000m , with a spacing of 10m . Once again, Figure 3 shows the diagonal of the Hessian operator, while Figure 4 illustrates the off-diagonals at the image point $x = 680\text{m}$, $z = 800\text{m}$. The results shown in Figure 3(c) and Figure 4(c) are obtained by using receiver-side plane-wave phase encoding, which perfectly removes the cross-talk. Figure 3(d) and Figure 4(d) show the results of random phase encoding. The cross-talk in the bottom left and right corners is also well removed. Also note that some random background noise has been introduced (Figure 3(d)).

Inversion with the phase-encoded Hessian: Sigsbee2A

To demonstrate the power of inversion, I apply the model-space inversion approach to the Sigsbee2A model. The explicit Hessian operator is computed using the random phase-encoding method with the frequency band from 5 Hz to 35 Hz , equivalent to that of the

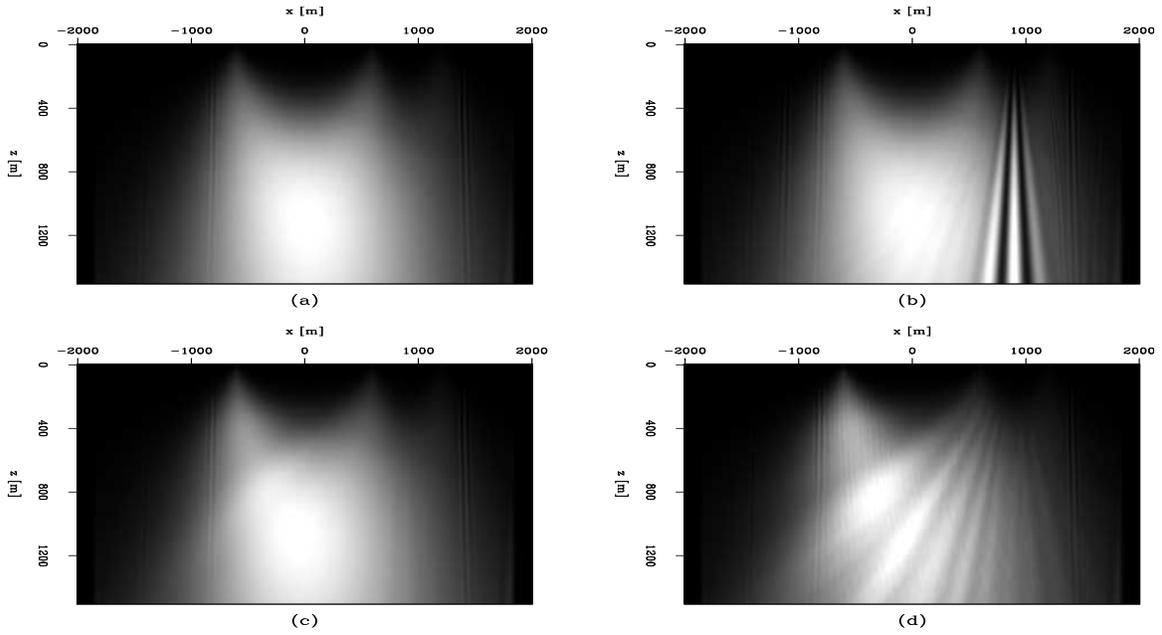


Figure 1: Diagonal of the Hessian for a constant-velocity model with only one shot and two receivers. (a) shows the exact diagonal of the Hessian obtained by directly implementing equation 8; (b) is obtained by using equation 9, and it is contaminated by strong cross-talk, which can be easily identified as the vertical stripes in the result; (c) is obtained by using plane-wave phase-encoding with 61 receiver-side-encoded plane waves stacked together to generate the result; the cross-talk is successfully removed; (d) is obtained by using the random phase-encoding; the cross-talk is dispersed, and the result contains less cross-talk.

[ER] [yaxun1/. const-hess-comparison](#)

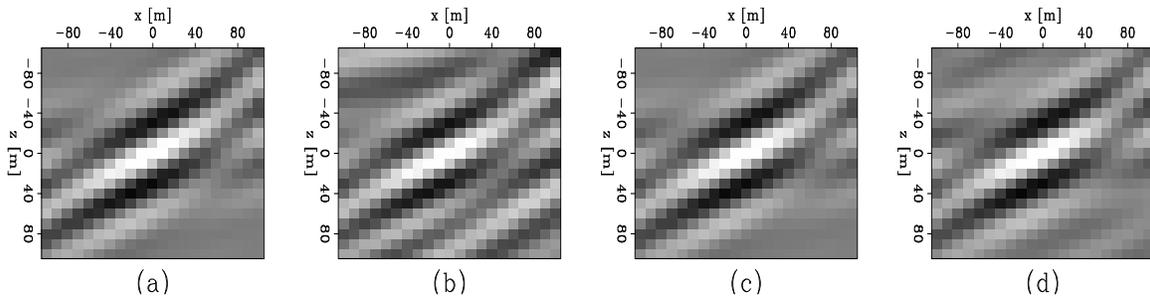


Figure 2: The Hessian operator for image point $x = 680\text{m}$, $z = 800\text{m}$ with the same acquisition geometry as in Figure 1. The size of the Hessian operator is 21×21 . Panel (a) shows the exact Hessian operator, free of cross-talk; (b) is obtained by using equation 9, and is contaminated by strong cross-talk, which can be easily identified by comparing with (a); (c) is obtained by using plane-wave phase-encoding with 61 receiver-side-encoded plane waves stacked together to generate the result; the cross-talk is successfully removed; (d) is obtained by using random phase-encoding, which also successfully removes the cross-talk.

[ER] [yaxun1/. const-hess-offd-comparison](#)

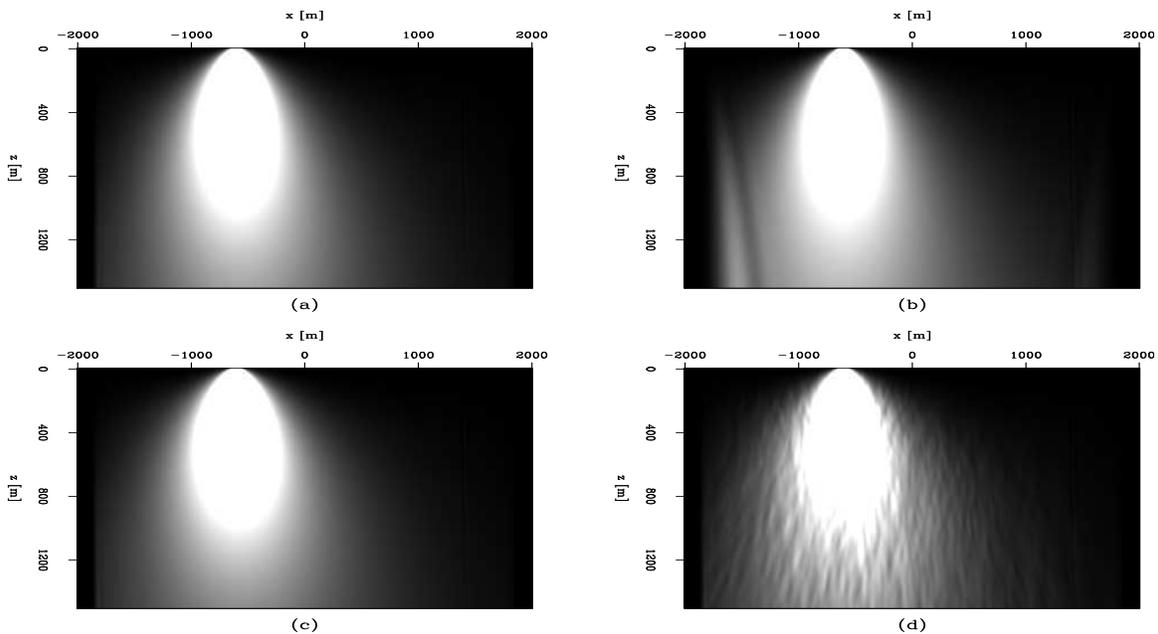


Figure 3: Diagonal of the Hessian for a constant velocity model with only one shot but 401 receivers. Panel (a) shows the exact diagonal of the Hessian obtained by directly implementing equation 8; (b) is obtained by using equation 9, and is contaminated by strong cross-talk, which can be easily identified by comparing with (a); (c) is obtained by using plane-wave phase-encoding with 61 receiver-side-encoded plane waves stacked together; the cross-talk is successfully removed; (d) is obtained by using the random phase-encoding method; the cross-talk in the corners is also successfully removed, but random background noise is introduced. [ER] [yaxun1/. const-hess-more-comparison](#)

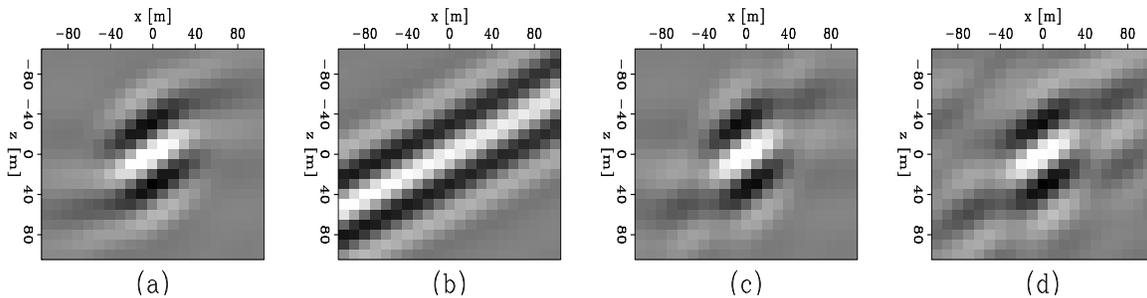


Figure 4: The Hessian operators for image point at $x = 680\text{m}$ and $z = 800\text{m}$, with the same acquisition geometry as in Figure 3. The size of the Hessian operator is 21×21 . Panel (a) shows the Hessian operator obtained by using the exact equation 8; it is free of cross-talks; (b) is obtained by using equation 9, and is contaminated by strong cross-talk, which can be easily identified by comparing with (a), note the cross-talk significantly influences the illumination pattern for that image point; (c) is obtained by using the plane-wave phase-encoding method with 61 receiver-side-encoded plane waves stacked together to generate the result; the cross-talk is successfully removed; (d) is obtained by using the random phase-encoding method, which also successfully removes the cross-talk. [ER] [yaxun1/. const-hess-more-offd-comparison](#)

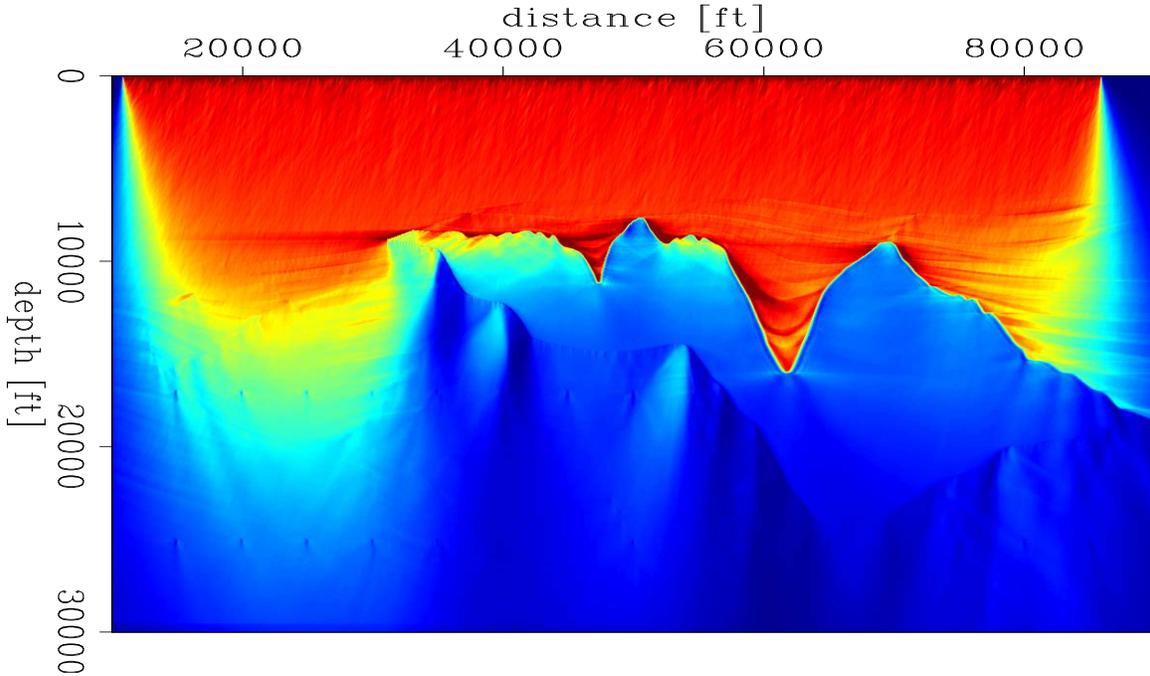


Figure 5: Diagonal of Hessian for the Sigsbee2A model. The Hessian is obtained by using random phase encoding, which takes the limited receiver array into consideration. The result shows the Hessian for the entire frequency band (5 Hz - 35 Hz). [CR] `yaxun1/. sigsb2a-hess-diag-random`

migrated image \mathbf{m}_{mig} . I apply Two different strategies: the first is to compute only the diagonal of the Hessian operator and use it to normalize the migrated image; the other is to compute a limited number of Hessian off-diagonals, and then use the linear conjugate gradient method to invert the reflectivity \mathbf{m} .

Figure 5 shows the diagonal of the Hessian using random phase encoding. Note the uneven illumination caused by the complex salt body and limited acquisition geometry. For comparison, Figure 6 shows the source intensity (SI), computed using the following equation:

$$\begin{aligned}
 H(\mathbf{x}_p, \mathbf{x}_p) \approx H_{SI}(\mathbf{x}_p, \mathbf{x}_p) &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_p, \omega) \\
 &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 |G(\mathbf{s}_i, \mathbf{x}_p, \omega)|^2.
 \end{aligned} \tag{30}$$

Comparing equation 30 with equation 8, $H_{SI}(\mathbf{x}_p, \mathbf{x}_p)$ is a crude approximation to the exact diagonal of Hessian, because it assumes the constant receiver-side Green's functions and ignores the effects of the limited receiver arrays. It over-estimates the total energy that enters the earth and returns to be recorded by the receivers. That is why Figure 6 shows a better, but inaccurate, illumination coverage below the salt.

Figure 7 shows the conventional shot-profile migrated image, where the shadow zones below the salt are easily identified; Figure 8 shows the result of normalizing the migrated

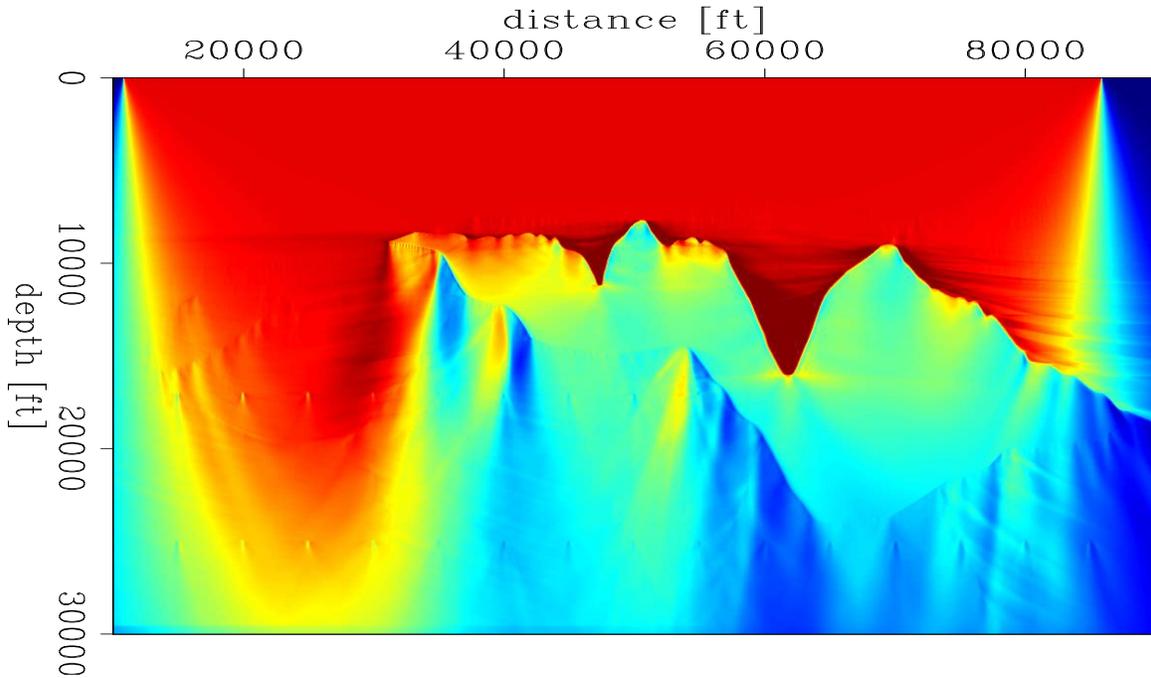


Figure 6: Source intensity for the Sigsbee2A model obtained by using equation 30, which assumes the receiver-side Green's functions to be constant; it ignores the effects of limited receiver array and over-estimates the total energy that enters the earth and returns to be recorded by the receivers. [CR] `yaxun1/. sigsb2a-hess-diag-souint`

image with the diagonal of the Hessian shown in Figure 5. For comparison, Figure 9 shows the result of normalizing the migrated image with the source intensity shown in Figure 6. Figure 8 shows slightly more balanced amplitude across the section than Figure 9, especially in areas below the salt. This is because Figure 5 takes the limited receiver arrays into consideration, and hence better predicts the illumination pattern in the subsurface.

Figure 10 shows the Hessian operator obtained by random phase encoding for a particular region of interest under the salt: (a) shows the diagonal of the Hessian; (b) is obtained by convolving the Hessian operator (with a size 21×21) with a collection of point scatterers in the model space. It demonstrates the varying shapes and non-stationarities of the Hessian operators across the model space. Note that in well-illuminated areas, the Hessian operator is well focused, while in poorly illuminated areas, the Hessian operator has a preferred dipping orientation, which means these image points are illuminated at only a few dip angles.

Figure 11 shows the comparison between migration and inversion: (a) is the migrated image, and (b) is the inverted image obtained using the randomly phase-encoded Hessian operator. The result is obtained after 20 iterations of the linear conjugate-gradient method. Figure 12 plots the residuals as a function of iteration number, the residual converges after about 12 iterations. In the inversion result, the vertical resolution is greatly enhanced; the shadow zones that in the migrated image are now filled in with structures; the sediments and the large dipping fault extend closer to the salt body. However, the inversion result also shows increased noise, which might be caused by the null space or the random noise

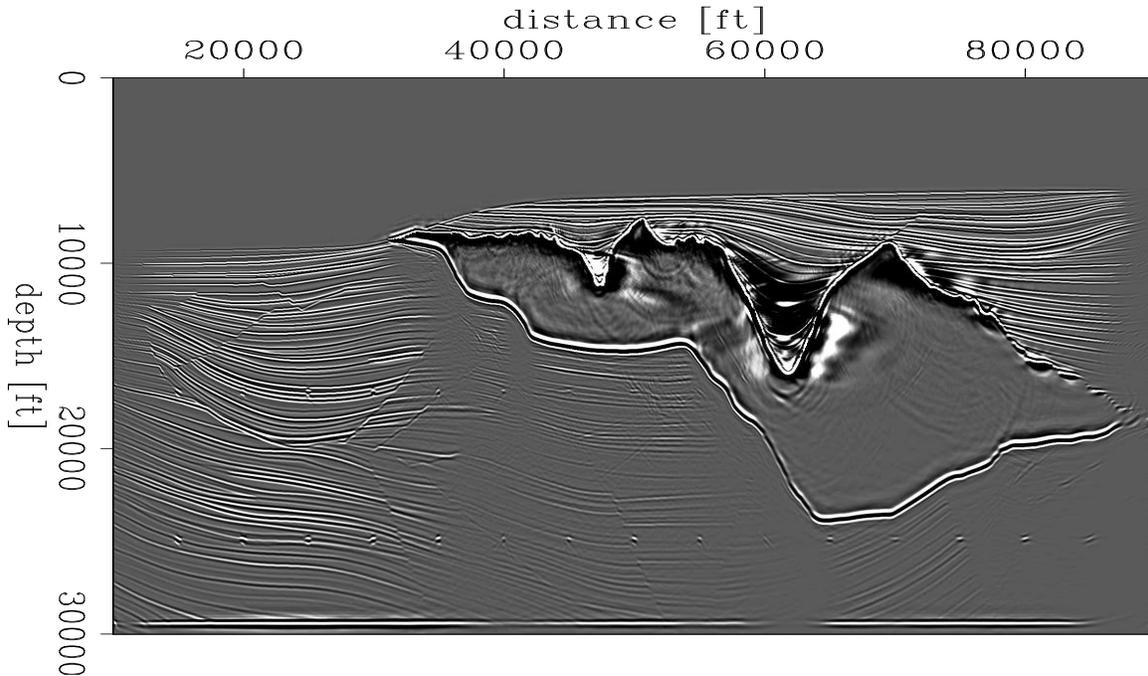


Figure 7: Conventional shot-profile migration result of Sigsbee2A model. Note the shadow zones under the salt. [CR] [yaxun1/. sigsb2a-imag-ffd](#)

introduced in the randomly phase-encoded Hessian, or both. To suppress the noise, a proper regularization term can be introduced in the inversion process.

CONCLUSIONS

I introduce a method based on phase encoding that allows efficient computation of the explicit Hessian operator. The proposed algorithm closely resembles shot-profile migration, except that a slightly different imaging condition is used, so that no Green's functions need to be pre-computed and saved on disk; hence, Hessians with larger model spaces and more frequencies can be obtained. However, this method also generates undesired cross-talk. I examine plane-wave phase encoding and random phase encoding to attenuate the cross-talk. Numerical examples demonstrate that these two phase-encoding schemes work well in suppressing the cross-talk. The inversion examples of using randomly phase-encoded Hessian on the Sigsbee2A model also show promising results.

ACKNOWLEDGMENT

I would like to thank Alejandro Valenciano for helpful discussions on the inversion part.

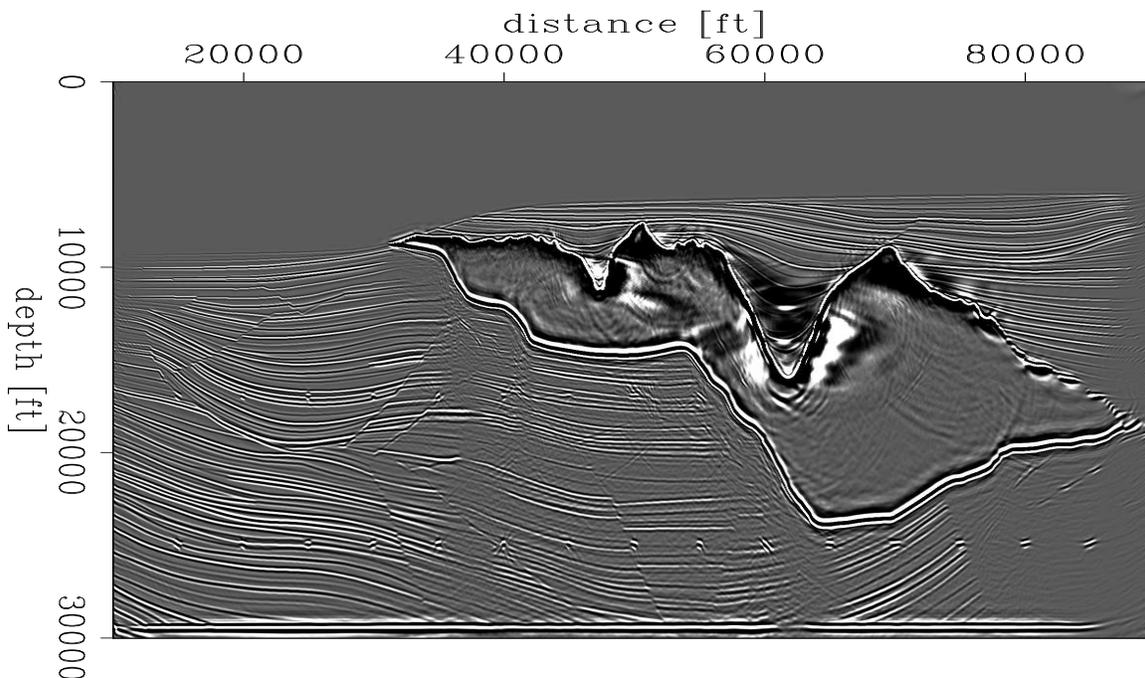


Figure 8: The migrated image (Figure 7) is normalized by the diagonal of the Hessian shown in Figure 5. [CR] `yaxun1/. sigsb2a-imag-ffd-decon-diag-random`

REFERENCES

- Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform: *J. Math. Phys.*, **26**, 99–108.
- Chavent, G. and R.-E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: *Geophysics*, **64**, 508–515.
- Claerbout, J. F., 1971, Towards a unified theory of reflector mapping: *Geophysics*, **36**, 467–481.
- Clapp, M. L., 2005, Imaging under salt: Illumination compensation by regularized inversion, *in* Ph.D. Thesis, Stanford University.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration: *Proc. Conf. on Inverse Scattering, Theory and Applications, Expanded Abstracts*, Philadelphia, SIAM.
- Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: *Geophysics*, **71**, S129–S139.
- Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, 208–221.
- Plessix, R.-E. and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: *Geophys. J. Int.*, **157**, 975–987.
- Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: *Geophys. J. Int.*, **133**, 341–362.
- 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.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation:

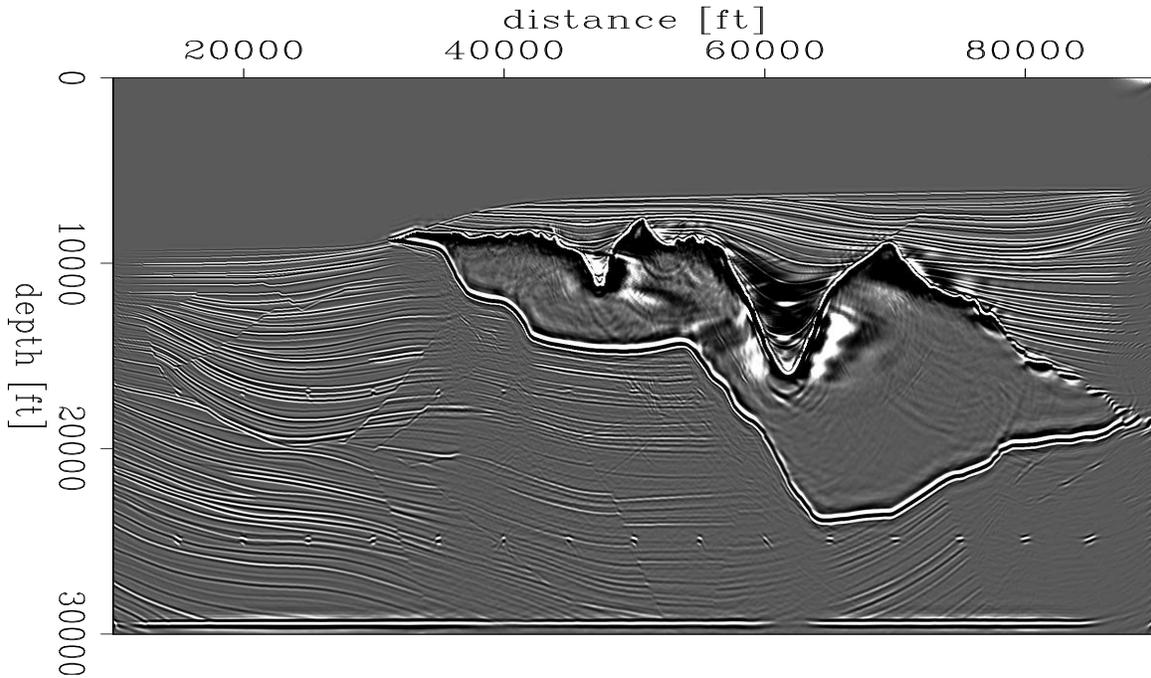


Figure 9: The migrated image (Figure 7) is normalized by the source intensity shown in Figure 6. [CR] `yaxun1/. sigsb2a-imag-ffd-decon-diag-souint`

Geophysics, **49**, 1259–1266.

Valenciano, A. A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: *Geophysics*, **71**, A35–A38.

Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismograms, *in* Ph.D. thesis, University of Tulsa.

Yu, J., J. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: *Geophysics*, **71**, S53–S62.

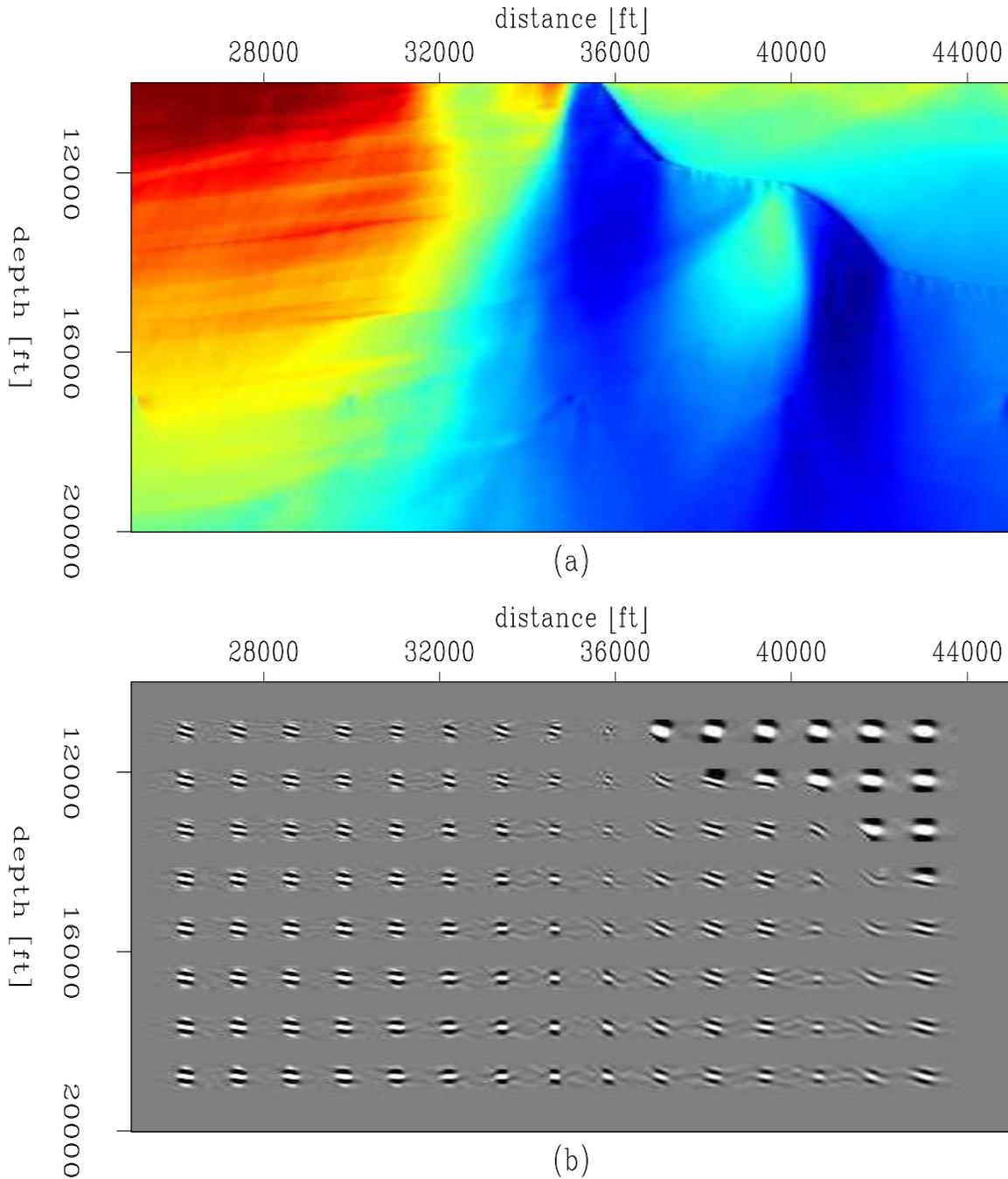


Figure 10: Randomly phase-encoded Hessian operators for Sigsbee2A. Panel (a) shows the diagonal of the randomly phase-encoded Hessian for a particular region of interest under the salt, note the uneven illumination due the complex salt body and limited acquisition geometry; (b) shows the result by convolving the Hessian operator (with a size 21×21) with a collection of point scatters, it demonstrates the varying shapes of the operators. [CR]

`yaxun1/. sigsb2a-hess-win`

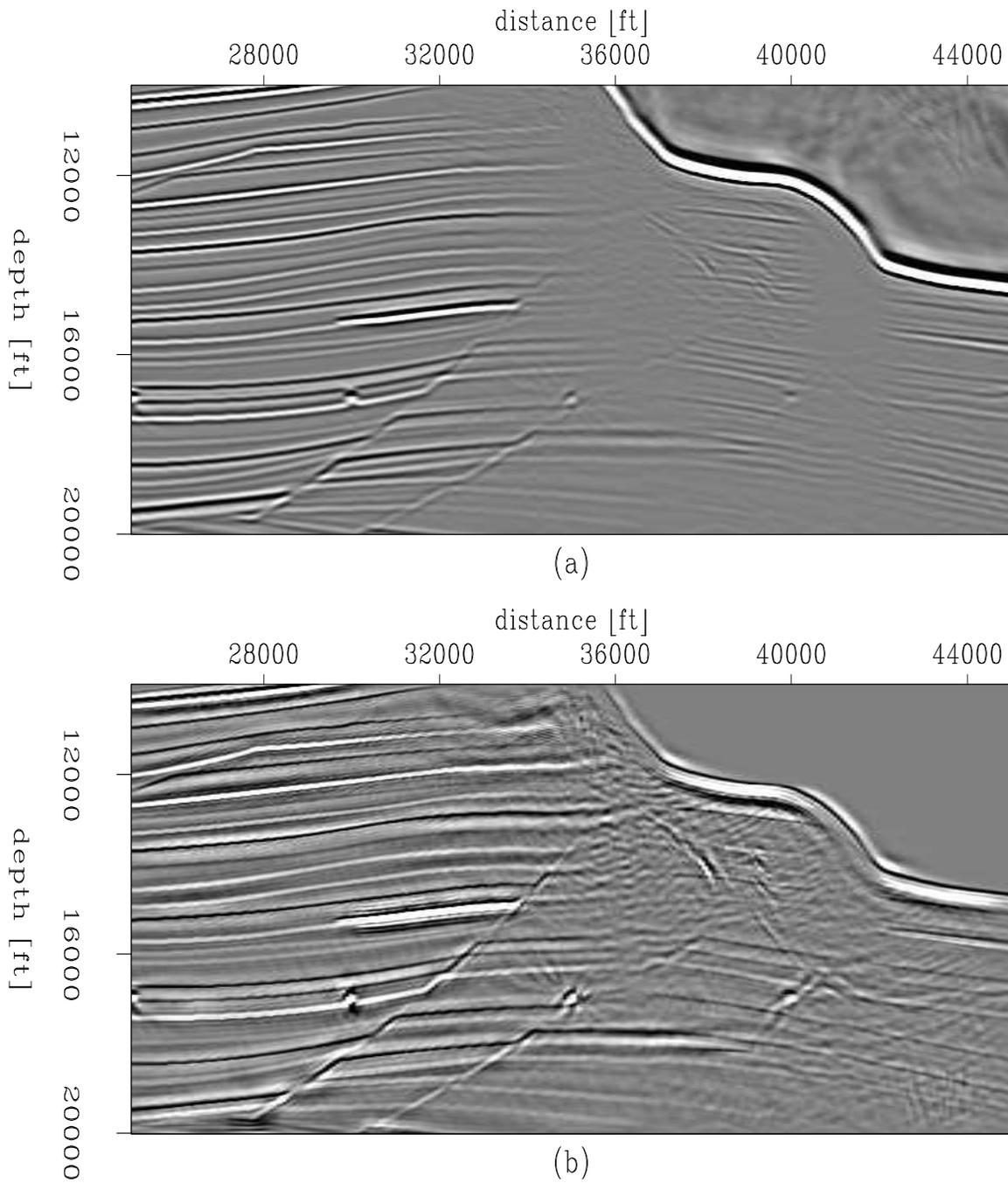
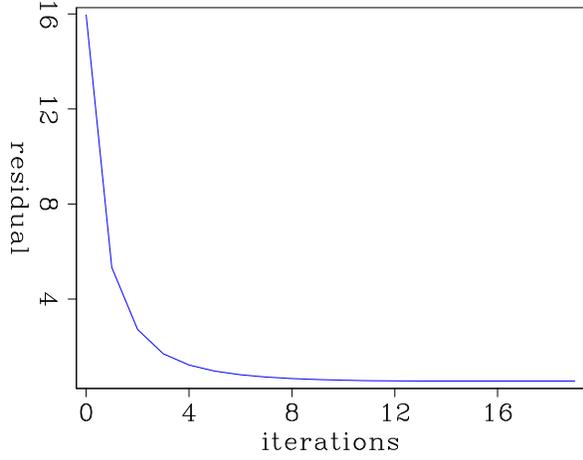


Figure 11: Comparison between migration and inversion. Panel (a) shows the conventional shot-profile migration result; (b) shows the inversion result using the 21×21 Hessian operators obtained by using the random phase-encoding method. [CR]

yaxun1/. sigsb2a-invt-win

Figure 12: Residual vs. number of iterations for the Sigsbee2A model; the inversion converges after about 12 iterations. [CR]

`yaxun1/. sigsb2a-invt-fmov-win`



APPENDIX A

This appendix derives the Hessian operator of a linear least-squares functional. Let us start with the objective function

$$F(\mathbf{m}) = \frac{1}{2} \sum_{\omega} \sum_{k=1}^M \sum_{i=1}^N |d(\mathbf{r}_k, \mathbf{s}_i, \omega) - d_{obs}(\mathbf{r}_k, \mathbf{s}_i, \omega)|^2, \quad (\text{A-1})$$

where the forward modeling operator is defined by

$$d(\mathbf{r}_k, \mathbf{s}_i, \omega) = \omega^2 \sum_{l=1}^L f_s(\omega) G(\mathbf{s}_i, \mathbf{x}_l, \omega) G(\mathbf{x}_l, \mathbf{r}_k, \omega) m(\mathbf{x}_l). \quad (\text{A-2})$$

The first-order partial derivative of the above objective function with respect to the model parameter (the gradient) is:

$$\frac{\partial F(\mathbf{m})}{\partial m(\mathbf{x}_p)} = \sum_{\omega} \omega^2 \sum_{i=1}^N \sum_{k=1}^M (d'(\mathbf{r}_k, \mathbf{s}_i, \omega) - d'_{obs}(\mathbf{r}_k, \mathbf{s}_i, \omega)) f_s(\omega) G(\mathbf{s}_i, \mathbf{x}_p, \omega) G(\mathbf{x}_p, \mathbf{r}_k, \omega). \quad (\text{A-3})$$

The Hessian is then obtained by taking the second-order partial derivatives of the objective function with respect to the model parameters:

$$\begin{aligned} H(\mathbf{x}_p, \mathbf{x}_q) &= \frac{\partial^2 F(\mathbf{m})}{\partial m(\mathbf{x}_p) \partial m(\mathbf{x}_q)} \\ &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \sum_{k=1}^M G(\mathbf{x}_p, \mathbf{r}_k, \omega) G'(\mathbf{x}_q, \mathbf{r}_k, \omega). \end{aligned} \quad (\text{A-4})$$

We can further assume that the reciprocity holds, such that:

$$G(\mathbf{x}, \mathbf{y}, \omega) = G(\mathbf{y}, \mathbf{x}, \omega). \quad (\text{A-5})$$

Therefore, the Hessian can be rewritten as follows:

$$H(\mathbf{x}_p, \mathbf{x}_q) = \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \sum_{k=1}^M G(\mathbf{r}_k, \mathbf{x}_p, \omega) G'(\mathbf{r}_k, \mathbf{x}_q, \omega). \quad (\text{A-6})$$

Equation A-6 is equivalent to equation 8.

APPENDIX B

This appendix demonstrates that the receiver-side plane-wave-encoded Hessian $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r)$ converges to the exact Hessian $H(\mathbf{x}_p, \mathbf{x}_q)$ by stacking over receiver ray parameters. By using the weighting functions defined in equation 23, the approximate Hessian for a single p_r takes the form:

$$\begin{aligned} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r) &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 A_r^2(\omega) G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) e^{i\omega p_r(\mathbf{r}_u - \mathbf{r}_0)} \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) e^{-i\omega p_r(\mathbf{r}_v - \mathbf{r}_0)}. \quad (\text{B-1}) \end{aligned}$$

If we stack over p_r from $-\infty$ to $+\infty$, we have

$$\begin{aligned} \sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r) &= \sum_{p_r=-\infty}^{+\infty} \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 A_r^2(\omega) G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) e^{i\omega p_r(\mathbf{r}_u - \mathbf{r}_0)} \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) e^{-i\omega p_r(\mathbf{r}_v - \mathbf{r}_0)} \\ &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 A_r^2(\omega) G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \sum_{p_r=-\infty}^{+\infty} e^{-i\omega p_r(\mathbf{r}_v - \mathbf{r}_u)}. \quad (\text{B-2}) \end{aligned}$$

Note that

$$\sum_{p_r=-\infty}^{+\infty} e^{-i\omega p_r(\mathbf{r}_v - \mathbf{r}_u)} \approx \frac{1}{|\omega|} \delta(\mathbf{r}_v - \mathbf{r}_u). \quad (\text{B-3})$$

If we choose the real function $A_r(\omega)$ such that it satisfies

$$\frac{A_r^2(\omega)}{|\omega|} = 1, \quad (\text{B-4})$$

then we get

$$\begin{aligned}
\sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_r) &= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 \frac{A_r^2(\omega)}{|\omega|} G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\
&\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \delta(\mathbf{r}_v - \mathbf{r}_u) \\
&= \sum_{\omega} \omega^4 \sum_{i=1}^N |f_s(\omega)|^2 G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\
&\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) G'(\mathbf{r}_u, \mathbf{x}_q, \omega) \\
&= H(\mathbf{x}_p, \mathbf{x}_q). \tag{B-5}
\end{aligned}$$

APPENDIX C

This appendix demonstrates that the simultaneously plane-wave-encoded Hessian $\tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_s, p_r)$ converges to the exact Hessian $H(\mathbf{x}_p, \mathbf{x}_q)$ by stacking over both source and receiver ray parameters. The simultaneously plane-wave-encoded Hessian takes the form

$$\begin{aligned}
\tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_s, p_r) &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 A_s^2(\omega) A_r^2(\omega) \times \\
&\quad \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) e^{\iota\omega p_s(\mathbf{s}_i - \mathbf{s}_0)} \sum_{j=1}^N G'(\mathbf{s}_j, \mathbf{x}_q, \omega) e^{-\iota\omega p_s(\mathbf{s}_j - \mathbf{s}_0)} \times \\
&\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) e^{\iota\omega p_r(\mathbf{r}_u - \mathbf{r}_0)} \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) e^{-\iota\omega p_r(\mathbf{r}_v - \mathbf{r}_0)}. \tag{C-1}
\end{aligned}$$

By stacking over p_s and p_r , we get

$$\begin{aligned}
\sum_{p_s=-\infty}^{+\infty} \sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_s, p_r) &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 A_s^2(\omega) A_r^2(\omega) \times \\
&\quad \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) \sum_{j=1}^N G'(\mathbf{s}_j, \mathbf{x}_q, \omega) \sum_{p_s=-\infty}^{+\infty} e^{-\iota\omega p_s(\mathbf{s}_j - \mathbf{s}_i)} \times \\
&\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \sum_{p_r=-\infty}^{+\infty} e^{-\iota\omega p_r(\mathbf{r}_v - \mathbf{r}_u)}. \tag{C-2}
\end{aligned}$$

Once again note that

$$\sum_{p_r=-\infty}^{+\infty} e^{-i\omega p_s(\mathbf{s}_j - \mathbf{s}_i)} \approx \frac{1}{|\omega|} \delta(\mathbf{s}_j - \mathbf{s}_i) \quad (\text{C-3})$$

$$\sum_{p_r=-\infty}^{+\infty} e^{-i\omega p_r(\mathbf{r}_v - \mathbf{r}_u)} \approx \frac{1}{|\omega|} \delta(\mathbf{r}_v - \mathbf{r}_u). \quad (\text{C-4})$$

If we choose real functions $A_s(\omega)$ and $A_r(\omega)$ such that they satisfy

$$\frac{A_s^2(\omega)}{|\omega|} = 1 \quad (\text{C-5})$$

$$\frac{A_r^2(\omega)}{|\omega|} = 1, \quad (\text{C-6})$$

then we get

$$\begin{aligned} \sum_{p_s=-\infty}^{+\infty} \sum_{p_r=-\infty}^{+\infty} \tilde{H}(\mathbf{x}_p, \mathbf{x}_q, p_s, p_r) &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 \frac{A_s^2(\omega)}{|\omega|} \frac{A_r^2(\omega)}{|\omega|} \times \\ &\quad \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) \sum_{j=1}^N G'(\mathbf{s}_j, \mathbf{x}_q, \omega) \delta(\mathbf{s}_j - \mathbf{s}_i) \times \\ &\quad \sum_{u=1}^M G(\mathbf{r}_u, \mathbf{x}_p, \omega) \sum_{v=1}^M G'(\mathbf{r}_v, \mathbf{x}_q, \omega) \delta(\mathbf{r}_v - \mathbf{r}_u) \times \\ &= \sum_{\omega} \omega^4 |f_s(\omega)|^2 \sum_{i=1}^N G(\mathbf{s}_i, \mathbf{x}_p, \omega) G'(\mathbf{s}_i, \mathbf{x}_q, \omega) \times \\ &\quad \sum_{k=1}^M G(\mathbf{r}_k, \mathbf{x}_p, \omega) G'(\mathbf{r}_k, \mathbf{x}_q, \omega) \\ &= H(\mathbf{x}_p, \mathbf{x}_q). \end{aligned} \quad (\text{C-7})$$

Reconciling processing and inversion: Multiple attenuation prior to wave-equation inversion

Claudio Guerra and Alejandro Valenciano

ABSTRACT

Seismic inversion is very sensitive to the presence of noise. In an inversion scheme, noise is any event in the data not predicted by the forward modeling because of either inaccurate physics or inadequate model parameterization. Therefore, noise exists in the residual space and, if coherent, slows convergence to an acceptable result, or in the worst case, dominates the whole process, inhibiting the efficacy of inversion. The ideal solution is to incorporate the modeling of the noise into the forward modeling. A more practical approach is to make the data agree with the physical assumptions of the inversion scheme. In this context, noise attenuation is a pre-processing step before inversion. Here, we illustrate the problem by applying one-way wave-equation inversion to a portion of the well known Sigsbee2b data. In the present example, noise takes the form of multiples, not modeled by the one-way wave-equation. After characterizing the noise in the migrated data, we use a dip filter to estimate it and a non-stationary adaptive filter technique to subtract it from the migrated data.

INTRODUCTION

Migration with the full wave equation is, in principle, capable of creating an image of the subsurface by reversing the propagation of all wave modes (Wapenaar et al., 1987). It is a computationally demanding process, which is why the wave equation is approximated by the one-way wave equation. Migration with the one-way approximation is formulated such that, ideally, the only seismic events present in the input data are primary reflections. When this is not the case, because the pre-processing was unable to completely eliminate noise (multiple reflections, for instance), the final image presents migrated events that do not correspond to the geology. In other words, these events are not modeled by the physical theory on which the migration method is based. They are frequently called migration noise or migration artifacts.

As is widely known, inversion problems are very sensitive to noise. In seismic, several authors have shown that coherent noise degrades the performance of the velocity inversion and yields erroneous results (Chauris and Noble, 2001; Shin and Min, 2006; Li and Symes, 2007). Therefore it is usually necessary to include some kind of a priori information about the model as an additional regularization equation in the inversion or to perform noise suppression prior to inversion.

Noise also affects least-squares inversion of migrated images into reflectivity. In offshore data, multiples are the main coherent noise. Valenciano (2008) addresses the problem of migrated multiple reflections in inversion by applying a pre-processing step to attenuate them. Here, using the Sigsbee2b dataset (Paffenholz et al., 2002), we characterize the

multiples in the subsurface-offset domain, detail the pre-processing step and show the impact it has on inversion.

LINEAR LEAST-SQUARES INVERSION

Linear least-squares inversion provides a theoretical approach to compensate for experimental deficiencies (e.g., limited acquisition geometry) and complexities of the overburden, while maintaining consistency with the acquired data. For seismic imaging, it can be summarized as follows. Given a linear modeling operator \mathbf{L} , compute synthetic data \mathbf{d} using $\mathbf{d} = \mathbf{L}\mathbf{m}$, where \mathbf{m} is a reflectivity model.

The quadratic cost function,

$$S(\mathbf{m}) = \|\mathbf{d} - \mathbf{d}_{obs}\|^2 = \|\mathbf{L}\mathbf{m} - \mathbf{d}_{obs}\|^2, \quad (1)$$

is formed, where \mathbf{d}_{obs} denotes the recorded data.

The reflectivity model $\hat{\mathbf{m}}$ that minimizes $S(\mathbf{m})$ is given by

$$\hat{\mathbf{m}} = (\mathbf{L}'\mathbf{L})^{-1}\mathbf{L}'\mathbf{d}_{obs} = \mathbf{H}^{-1}\mathbf{m}_{mig}, \quad (2)$$

where \mathbf{L}' (the migration operator) is the adjoint of the linear modeling operator \mathbf{L} , \mathbf{m}_{mig} is the migrated image, $\mathbf{m}_{mig} = \mathbf{L}'\mathbf{d}_{obs}$, and $\mathbf{H} = \mathbf{L}'\mathbf{L}$ is the Hessian of $S(\mathbf{m})$.

Since the model space can be large, computing the inverse of the Hessian matrix is a big challenge for most geophysical imaging problems. For this reason, it is often more feasible to compute the inverse image as the solution of the linear system of equations,

$$\mathbf{H}\hat{\mathbf{m}} = \mathbf{m}_{mig}, \quad (3)$$

by using an iterative inversion algorithm. In this approach to the inverse problem, only one matrix-vector multiplication of the Hessian matrix with the model vector is necessary per iteration, in contrast with other methods (Clapp, 2005) that require one migration and one modeling every iteration. Still, explicit computation of the Hessian for the entire model space is too expensive in practice. Valenciano (2008) discusses how exploiting the structure of the Hessian matrix and the localization in the model space makes this problem tractable.

By using a priori information about the covariance of the model (model regularization), one can add regularization to solve the otherwise ill-posed inversion problem. A more customary regularization for the inversion in the poststack image domain is to add a damping factor that penalizes an increase of the values of the model. This regularization makes no use of any physical knowledge we might have about the seismic reflectors. It is implemented by adding a small value to the diagonal of $\mathbf{H}(\mathbf{x}, \mathbf{x}')$ in equation 3:

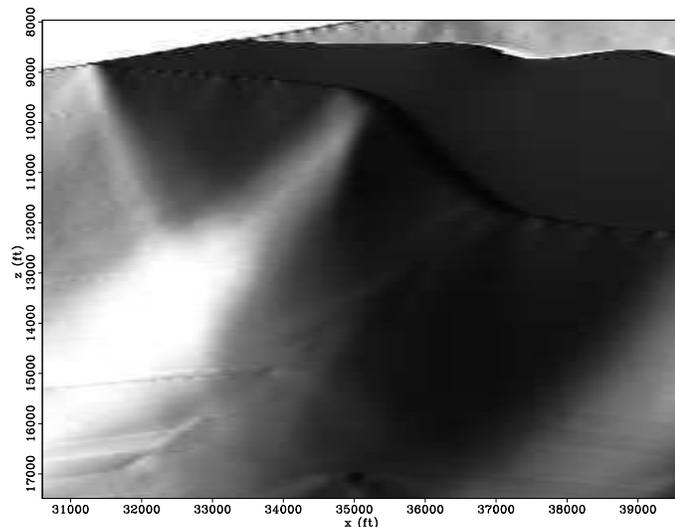
$$(\mathbf{H} + \varepsilon\mathbf{I})\hat{\mathbf{m}} - \mathbf{m}_{mig} = \mathbf{r} \approx 0, \quad (4)$$

where \mathbf{I} is the identity operator, \mathbf{r} is the residuals vector, and ε is a scalar parameter that governs the strength of the regularization.

NOISE CHARACTERIZATION

The occurrence of shadow zones below the salt body is strongly conditioned by the concavity of its base. In the Sigsbee2b model, the comparison of the diagonal of the Hessian matrix in Figure 1 with the reflectivity model in Figure 2 shows that where the base of the salt is concave down, the transmitted energy is focused and high illumination values are produced; where the base of the salt is concave up, the transmitted energy is spread along to divergent paths, producing low illumination zones. Unfortunately, for reflected energy and also for multiple energy, this geometry acts in the opposite way, concentrating energy where the base of the salt is concave up and dispersing energy where it is concave down. In the shadow zones it is very likely that the strongest remaining amplitudes correspond to multiple reflections, while in the well-illuminated zones these events produce little or no spurious interference with the primary energy (Figure 3). Consequently, during one-way wave-equation inversion, multiple energy can dominate the residuals, surpassing primaries in the shadow zones.

Figure 1: Illumination pattern given by the diagonal of the Hessian matrix of the Sigsbee2B model. Dark gray represents low illumination and light gray represents high illumination. [ER] `claudio1/. diag-Sis`



One solution to this problem is to attenuate the multiples (events that, in inversion, do not fit the model) from the migration. This requires that the multiples be separated from the primaries or modeled in such a way that during subtraction the primaries are minimally affected.

Traditional data-space demultiple schemes dealing either with periodicity or differential moveout become ineffective in sub-salt settings for two main reasons: a) complex ray paths make periodicity assumptions fail, and b) salt focusing effects concentrates sub-salt reflections in the near offset range, making Radon-type demultiple schemes ineffective. Sava and Guitton (2005) and Alvarez et al. (2007) demonstrated that primaries and multiples have different behavior in image space (subsurface-offset or reflection-angle), and that multiples can be adaptively subtracted from the migration in the image domain without significantly affecting the primaries.

The Sigsbee2B model (Paffenholz et al., 2002) has a high-reflectivity water-bottom. In contrast with Sigsbee2a, this characteristic generates strong peg-leg multiples, which bounce between the salt edge and the water-bottom. In Figure 4 we use part of the zero-offset section of Sigsbee2b data (non-free-surface data) to highlight these multiples and the

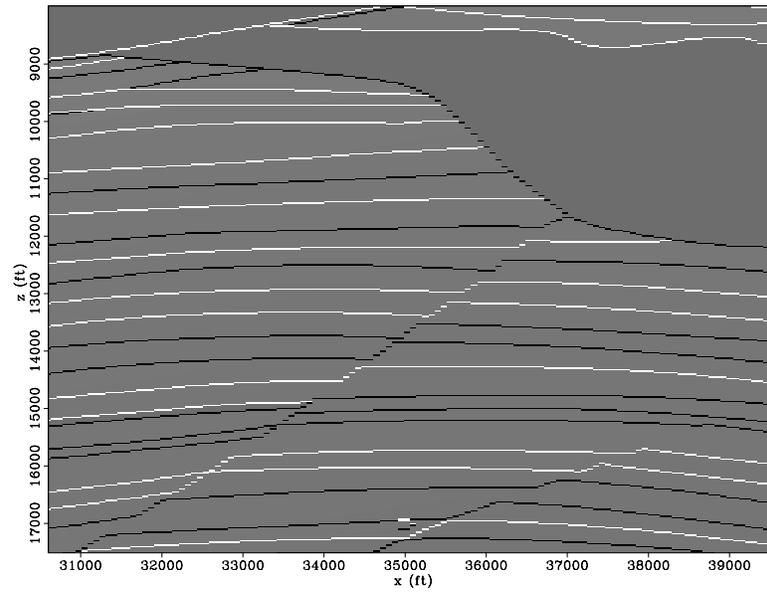


Figure 2: Reflectivity model of Sigsbee2B model. [ER] `claudio1/. refle-new`

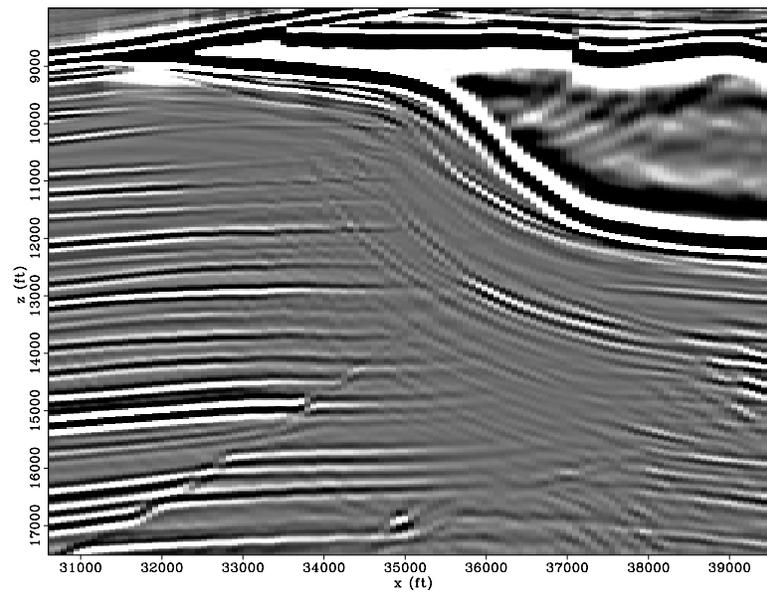
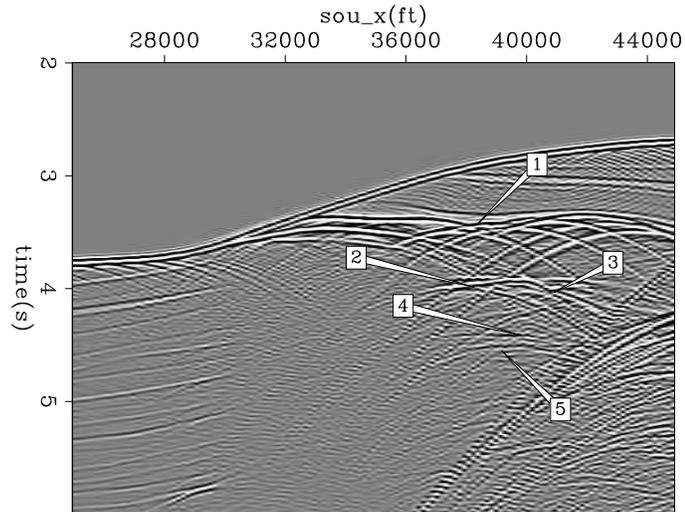


Figure 3: Migration of Sigsbee2b data - zero-subsurface-offset section. [ER] `claudio1/. mig-stack`

salt limits in the time domain. The diffractions labeled 1 correspond to the undulating top of the salt; that labeled 2 is the reflection from its base; 1st-order peg-leg diffracted multiples, concurrent with the base of the salt, are labeled 3; the 1st-order peg-leg multiple from the base of the salt is identified with label 4; and internal multiples generated inside the salt body are labeled 5.

Figure 4: Sigsbee2b zero-offset time section. The labels correspond to 1) top of the salt; 2) base of the salt; 3) 1st-order peg-leg diffracted multiples; 4) internal multiples originated at the salt limits; and, 5) 1st-order peg-leg multiple from the base of the salt. [ER] claudio1/. ZOtime



After prestack depth-migration, multiple reflections are distorted, because they are downward propagated beyond their actual reflection depth, using the incorrect velocities of the time-concurrent primaries. If their reflection traveltimes are coincident with sub-salt reflections, their propagation directions are severely deviated; therefore, the shape they assume shows a strong imprint of the velocity model.

In Figure 5 the migration result of Sigsbee2b is shown in a magnified view. The frame on the left is the zero-subsurface-offset section, and the one on the right is a subsurface offset gather – ODCIG – taken at CMP position 38825 ft. The 1st-order peg-leg diffracted multiples, labeled 1, show the effect of over-migration, because they are migrated with salt velocity; labels 2 and 4 stand for migrated peg-leg multiples from the base of the salt, and immediately above them, labeled 3 and 5, are the migrated internal multiples. Some reflectors show focusing close to zero-offset. Notice the strong correlation between the shape of migrated multiples and the base of the salt caused by the deviation of the propagation direction mentioned above. For comparison, Figure 6 shows the migration of Sigsbee2a. In this case, multiples related to the water-bottom (labels 1, 2 and 4) are much weaker or absent, because the water-bottom in this model is characterized by a “soft” interface.

Multiples are more evident in regions of low illumination. In Sigsbee data, these regions are associated with a concave-up base of the salt, in which downgoing energy is defocused and upgoing energy is focused. Much of the remaining energy in these regions corresponds to multiples. Figure 7 exemplifies this problem. The ODCIG on the right frame was selected from a low illumination area. The horizontal to upward-curved events are migrated multiples, which contaminate the entire sub-salt section. Because of the low illumination, some reflectors appear as dipping events below 14000 ft. Notice that, in the zero-subsurface-offset section, because of the imprint of the salt velocity, all the multiple modes show different dips than the reflectors.

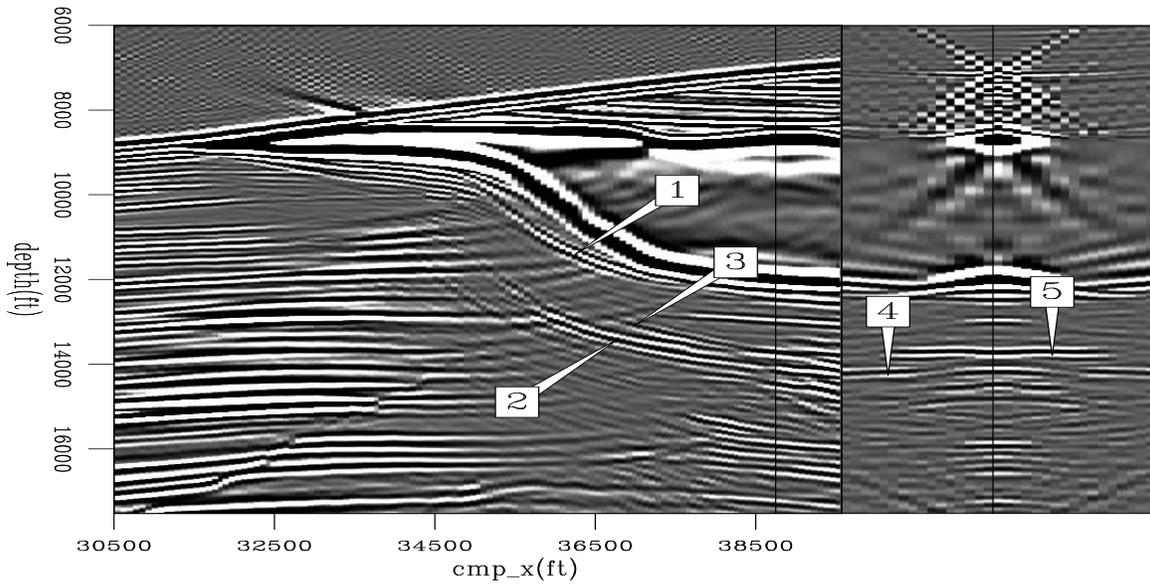


Figure 5: Sigsbee2b shot-profile migration using the cross-correlation imaging condition. The labels correspond to migrated events from 1) 1st-order peg-leg diffracted multiples; 2) 1st-order peg-leg multiple from the base of the salt; 3) internal multiples originated at the salt limits; 4) 1st-order peg-leg multiple from the base of the salt; and, 5) internal multiples originated at the salt limits. The front face corresponds to the image at zero subsurface-offset, and the side face corresponds to the subsurface-offset gather at $x = 38600$ ft. [ER] claudio1/. sig2b01

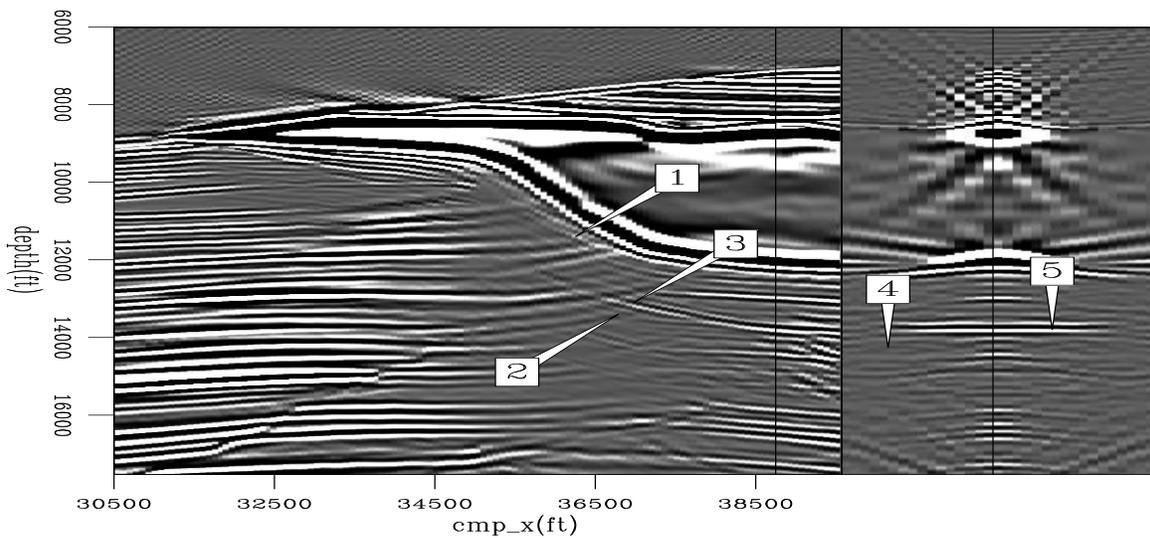


Figure 6: Sigsbee2a shot-profile migration using the cross-correlation imaging condition. The labels correspond to migrated events from 1) 1st-order peg-leg diffracted multiples; 2) 1st-order peg-leg multiple from the base of the salt; 3) internal multiples originated at the salt limits; 4) 1st-order peg-leg multiple from the base of the salt; and, 5) internal multiples originated at the salt limits. The front face corresponds to the image at zero subsurface-offset, and the side face corresponds to the subsurface-offset gather at $x = 38600$ ft. Compare with Figure 5. [ER] claudio1/. sig2a01

The different dip behavior of primaries and multiples can be used as a criterion to separate them using a dip filter in the $k_x - k_h$ wavenumber domain in a pre-processing step. After discrimination, and prior to subtraction from the original data, amplitude and phase of the estimated multiples must be adjusted. We perform this correction by using non-stationary filters, according to the strategy of Alvarez and Guitton (2007). They advocate that the best adjustment between the estimated multiples and the multiples in the data is achieved by simultaneously matching the estimation of primaries and multiples to the data containing both. The strategy uses small overlapping patches of the input data to compute local filters in a least-squares inverse problem.

The next section presents the results of pre-processing and its impact on the inversion output.

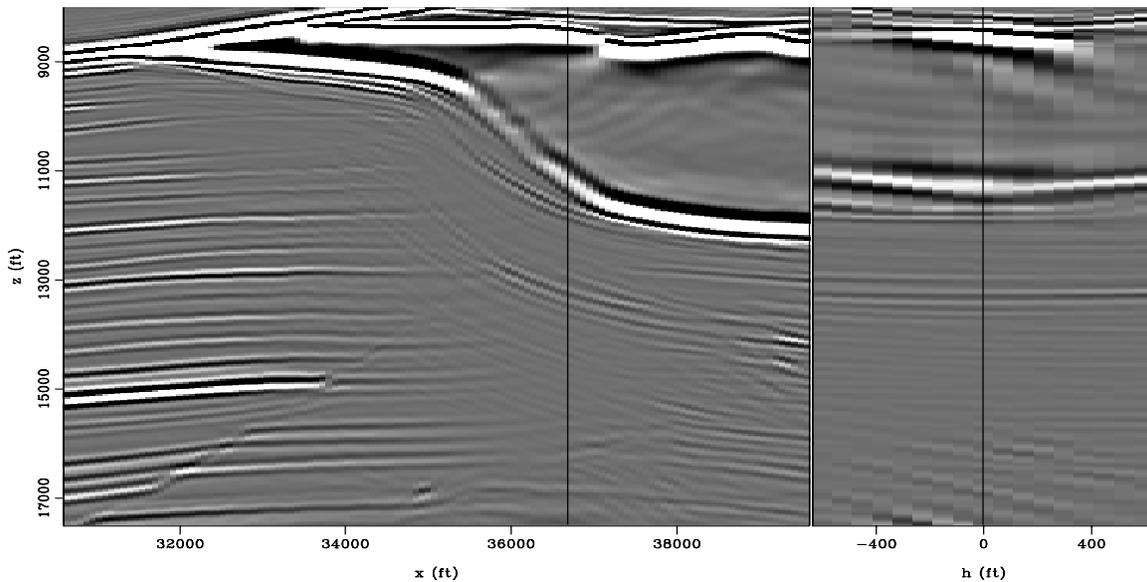


Figure 7: Sigsbee2b shot-profile migration (subsurface-offset) using the cross-correlation imaging condition. The front face corresponds to the zero subsurface-offset image, and the side face corresponds to the subsurface-offset gather at $x = 36800$ ft. Poorly illuminated areas are dominated by migrated multiples. [ER] `claudio1/. migoff1`

RESULTS

The Hessian matrix for the Sigsbee2B model (diagonal shown in Figure 1) and the migration (Figure 3) were computed using one-way wave-equation propagators. As result, the multiples were neither included in the modeling nor propagated with the correct velocity/direction during migration. For this reason, the multiples will be treated as noise for the inversion. In the following, we will show how they prevent the inversion fail from recovering the primary signal in the shadow zones.

Figure 7 shows the migration result in the subsurface-offset domain. The front face corresponds to the image at zero subsurface offset, and the side face corresponds to the subsurface-offset gather at a position inside the shadow zone ($x = 36800$ ft). Since the primaries are illuminated at few reflection angles and from a dominant slanted wave-

propagation direction, their signature in the subsurface-offset domain is a slanted line. Because of the multiple bounces on dipping interfaces (15° water-bottom and 30° salt flank), multiple energy is spread horizontally in the ODCIG, resembling a zero-reflection-angle event. Figure 8 shows the migration result transformed to the reflection-angle domain. The front face corresponds to the image at the 0° reflection angle, and the side face corresponds to the reflection-angle gather at a position inside the shadow zone ($x = 36800$ ft). Notice how multiples are concentrated around the 0° reflection angle, and primaries are mapped at higher reflection angles. This behavior of the multiples and primaries can also be found in field datasets (Valenciano, 2008).

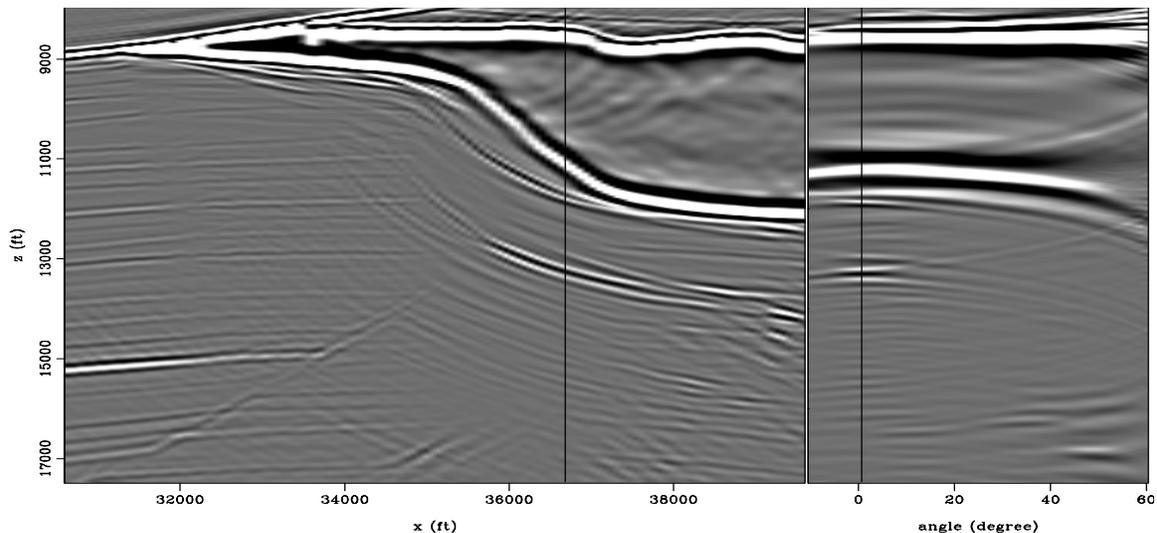


Figure 8: Sigsbee2b shot-profile migration (reflection angle) using the cross-correlation imaging condition. The front face corresponds to the image at the 0° reflection angle, and the side face corresponds to the reflection-angle gather at $x = 36800$ ft. [ER]

claudio1/. migang1

We used a three-stage pre-processing strategy to attenuate the multiple energy before inversion. It includes discriminating the multiple energy in the $k_x - k_h$ plane (to generate a model of the multiples); performing amplitude and phase correction of the discriminated multiples in a least-square sense; and subtracting them from the original data.

The discrimination of the multiple energy relies on the differences in dip patterns as described previously: multiples show up in the $k_x - k_h$ plane at low k_h and high k_x , and primaries at high k_h and low k_x . Therefore, we build a model of the primaries by submitting the migrated data to a $k_x - k_h$ filter at every depth step. Figure 9 shows a depth slice at $z = 14150$ ft of the migration in the $k_x - k_h$ plane, and Figure 10 shows the result of filtering. The model for the multiples (Figure 11) is obtained after subtraction of the filtered result (Figure 10) from the migration (Figure 9).

Since the separation of primaries and multiples is not perfect, some of the multiples are present in the model of primaries, and some primary energy leaks to the model of multiples. Because of this cross-talk between the estimates of primaries and multiples, it is desirable to control the amount of attenuation. Therefore, the impact on the primaries of the multiple-attenuation process can be ameliorated if the amplitude and phase of the multiples model

Figure 9: Depth slice at $z = 14150$ ft of the migration in the $k_x - k_h$ plane.

[ER] `claudio1/. not-filtered`

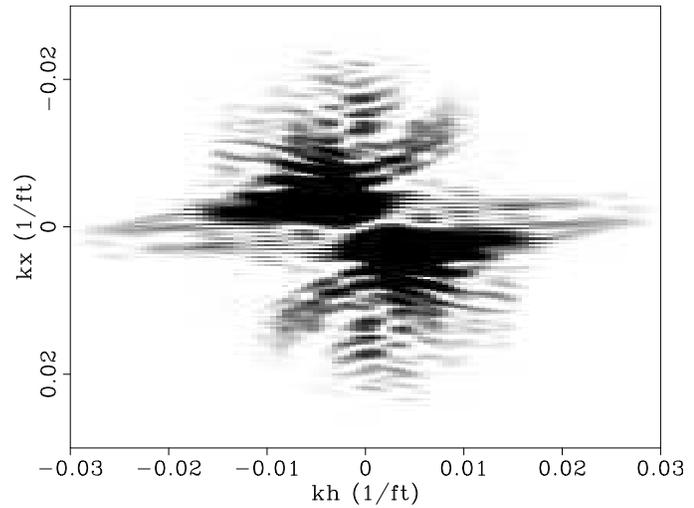


Figure 10: Primaries after band-passing the migration (Figure 9) in the $k_x - k_h$ plane.

[ER] `claudio1/. filtered`

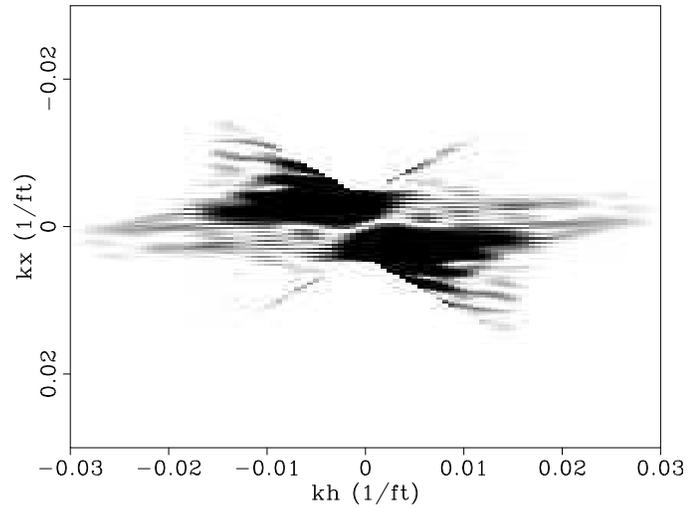
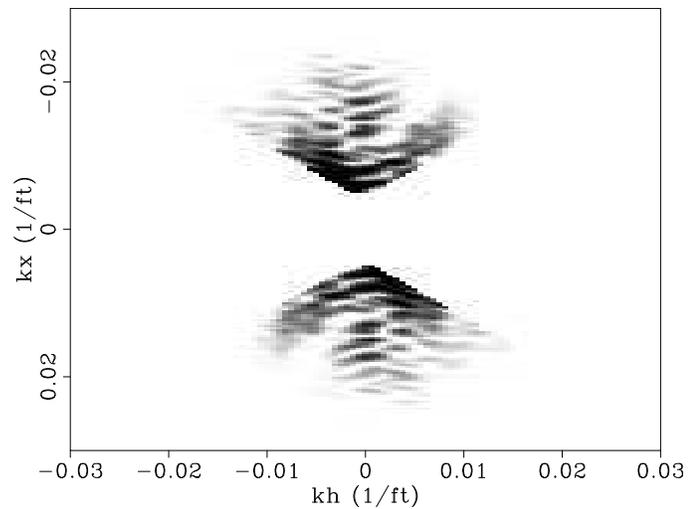


Figure 11: The model for the multiples in the $k_x - k_h$ plane, obtained after subtraction of the filtered result (Figure 10) from the migration (Figure 9).

[ER] `claudio1/. diff-filtered`



are adjusted, minimizing cross-talk and the differences from the multiples present in the original data in a least-squares sense. To perform this adjustment, we use simultaneous adaptive matching of primaries and multiples, formulated by Alvarez and Guitton (2007).

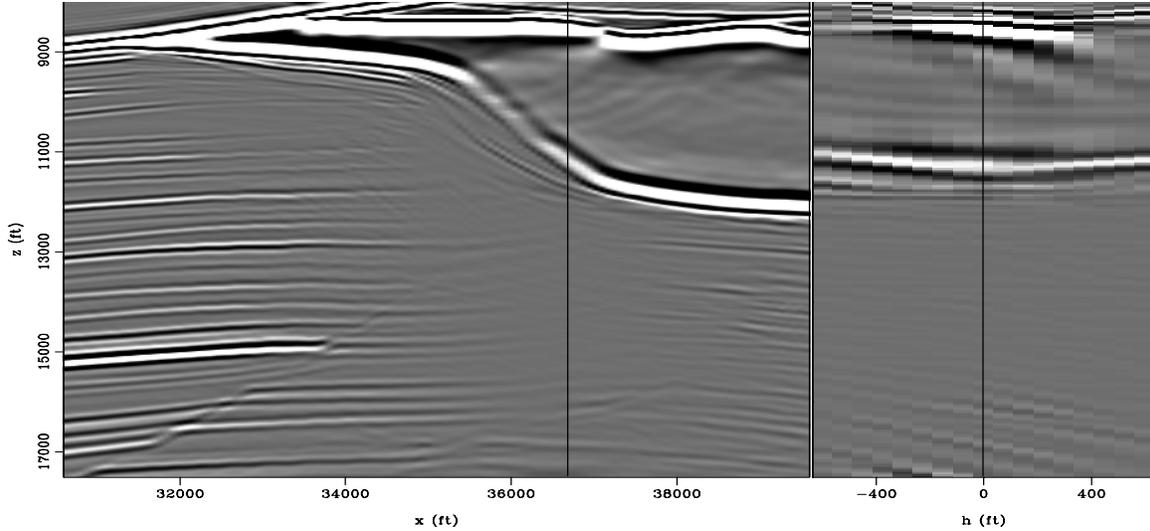


Figure 12: Sigsbee2b shot-profile migration (subsurface-offset) using the cross-correlation imaging condition, after multiple attenuation. The front face corresponds to the image at zero subsurface-offset, and the side face corresponds to the subsurface-offset gather at $x = 36800$ ft. Compare with Figure 7. [ER] `claudio1/. migoff1-filt`

After pre-processing, the multiples have been largely attenuated, making the migration image more suitable for inversion. Figure 12 shows the filtered migration in the subsurface-offset domain. As in Figure 7, the front face corresponds to the image at zero subsurface offset, and the side face corresponds to the subsurface-offset gather at a position inside the shadow zone ($x = 36800$ ft). The angle domain corroborates this hypothesis (Figure 13). The front face corresponds to the image at the 0° reflection angle, and the side face corresponds to the reflection-angle gather at a position inside the shadow zone ($x = 36800$ ft).

Inversion was computed using equation 3 for two different right-hand-side vectors. One is the migrated image at zero subsurface offset without filtering (front panel, Figure 7), and the other is the migrated image at zero subsurface offset after filtering (front panel, Figure 12).

The results of the inversion are shown in Figures 14 and 15 (unfiltered and filtered input), and should be compared with the migrations shown in Figures 7 and 12, respectively. The inversion with the unfiltered migration is more unstable. This should require the use of a high value of the regularization parameter, which reduces the effectiveness of the inversion outside the shadow zones.

The inversion of the filtered migration shows a better convergence behavior. Again, no regularization was applied. Notice the well-collapsed diffractor at $(35000, 17000)$, in contrast to the unfiltered version (Figure 14). For the latter, conjugate-gradient iterations focused on reducing the multiple energy in the space of the residuals, decreasing the overall efficacy of inversion.

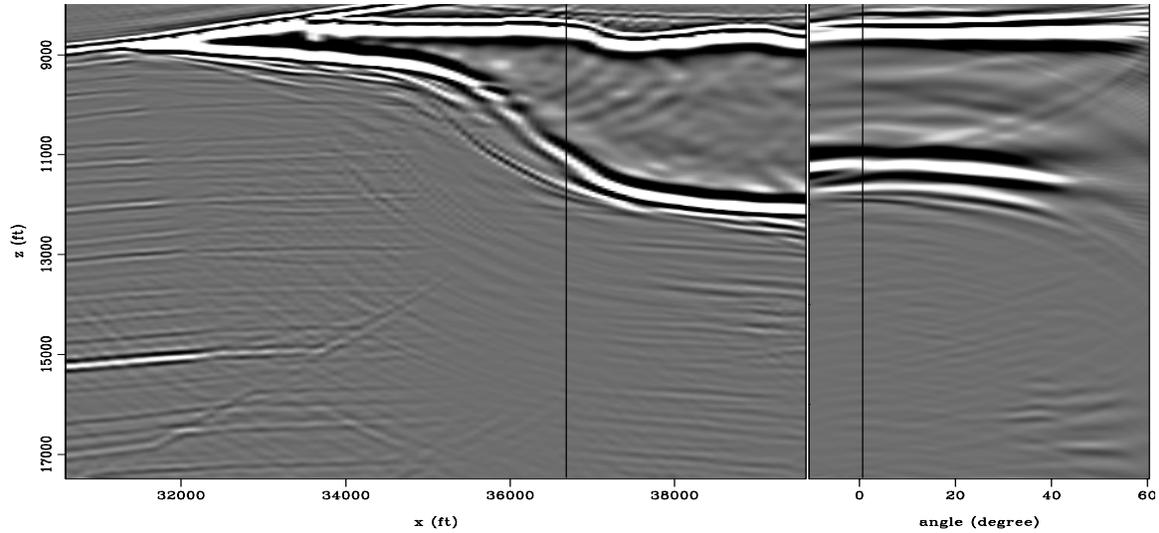


Figure 13: Sigsbee2b shot-profile migration (reflection angle) using the cross-correlation imaging condition, after multiple attenuation. The front face corresponds to the image at the 0° reflection angle, and the side face corresponds to the reflection-angle gather at $x = 36800$ ft. Compare with Figure 8. [ER] `claudio1/. migang1-filt`

Figure 14: Inversion in the poststack image domain with the unfiltered migration as the input (equation 3) after seven iterations of a conjugate-gradient iterative solver and no regularization. [ER] `claudio1/. invz`

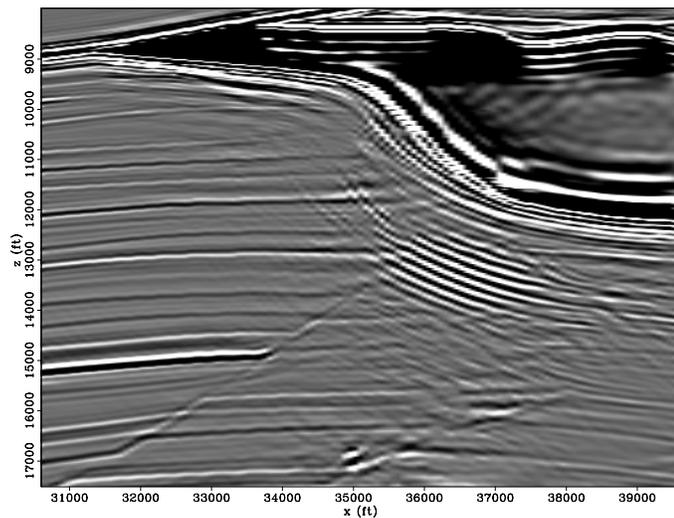
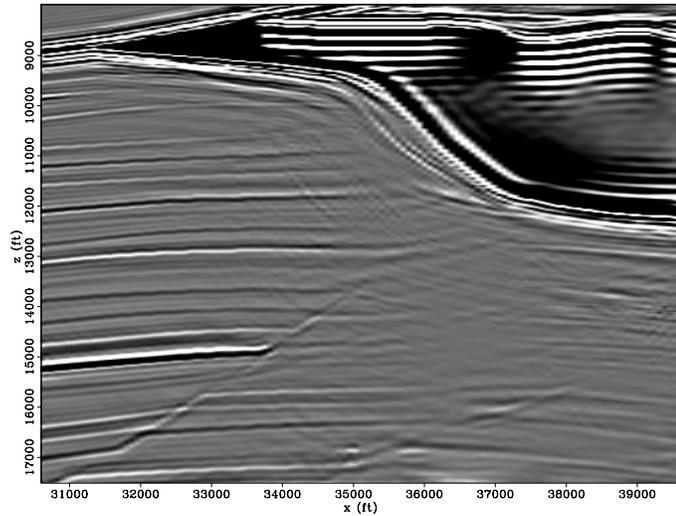


Figure 15: Inversion in the poststack image domain with a filtered migration as the input to equation 3. [ER]

claudio1/. invz-filt



Figures 16 and 17 show the residuals of the inversion with the unfiltered migration and with the filtered migration, respectively, after seven iterations of the conjugate-gradient solver. Except for the residuals in the salt body, the highest residual amplitudes of the unfiltered inversion correspond to the multiples in the shadow zone. Notice that some amount of the diffractor energy (35000,17000) is still present in the residuals. On the other hand, the residuals of the inversion with the filtered data do not show the diffractor energy. Additionally, the residuals of the fault close to the diffractor are smaller than that of the unfiltered version. Unfortunately the multiples were not completely removed by filtering, so again the inversion procedure increases their amplitudes.

Overall, the inversion results after pre-processing have more balanced amplitudes, allowing the continuation of the reflector inside the shadow zones with much-improved kinematics. The reflectors also gain more vertical and horizontal resolution, particularly seen at the two faults present in the reflectivity model (Figure 2) and at the diffractor mentioned above.

Figure 16: Residuals of the inversion in the poststack image domain with the unfiltered migration as the input after seven iterations. [ER]

claudio1/. resinvz

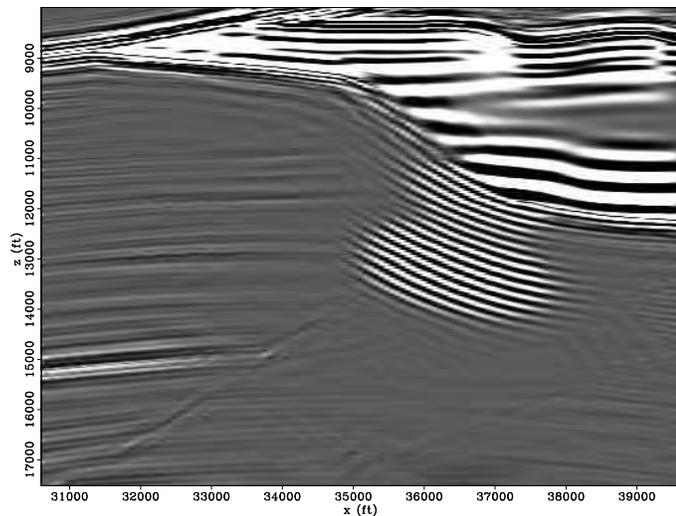
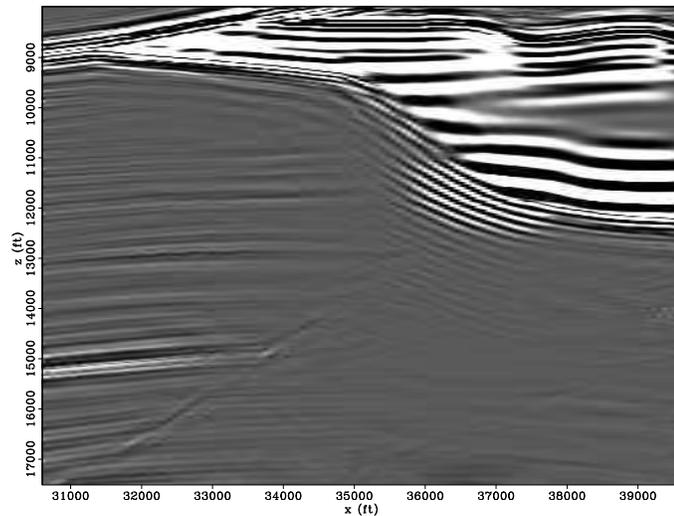


Figure 17: Residuals of the inversion in the poststack image domain with a filtered migration as the input after seven iterations. [ER]

claudio1/. resinvz-filt



CONCLUSIONS

We showed how different modes of multiple reflections present in Sigsbee2b data behave after migration. The multiple reflections we face are not surface-related; therefore, they can still persist after multiple attenuation during processing. As the model used to compute this dataset is based on a realistic geological setting, the same behavior is likely to happen in real data. We explored the separability of primaries from multiples after migration by applying simple $k_x - k_h$ filters. Once discriminated, primaries and multiples were submitted to a simultaneous, adaptive, non-stationary filtering to adjust amplitude and phase and decrease the cross-talk between them. The adapted multiples were then subtracted from the migrated data.

The inversion results show that, if the physics of the wave-propagation is not adequate to account for all the propagation modes in the migrated data, coherent noise – in our case multiples – will dominate the residual space, decreasing the efficiency of inversion. It is imperative, therefore, to make data satisfy the physical approximations by means of pre-processing that, in some sense reconciles processing and inversion.

REFERENCES

- Alvarez, G., B. Biondi, and A. Guitton, 2007, Attenuation of specular and diffracted 2D multiples in image space: *Geophysics*, **72**, V97–V109.
- Alvarez, G. and A. Guitton, 2007, Simultaneous adaptive matching of primaries and multiples with non-stationary filters: SEP-125, 61–77.
- Chauris, H. and M. Noble, 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets: *Geophysical Journal International*, **144**, 14–26.
- Clapp, M. L., 2005, Imaging Under Salt: Illumination Compensation by Regularized Inversion: PhD thesis, Stanford University.
- Li, J. and W. Symes, 2007, Interval velocity estimation via NMO-based differential semblance: *Geophysics*, **72**, U75–U88.
- Paffenholz, J., B. McLain, J. Zaskie, and P. Keliher, 2002, Subsalt multiple attenuation and

- imaging: Observations from the Sigsbee2b synthetic data set: 72nd Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2122–2125.
- Sava, P. and A. Guitton, 2005, Multiple attenuation in the image space: *Geophysics*, **70**, V10–V20.
- Shin, C. and D. Min, 2006, Waveform inversion using a logarithmic wavefield: *Geophysics*, **71**, R31–R42.
- Valenciano, A., 2008, Imaging by Wave-Equation Inversion: PhD thesis, Stanford University.
- Wapenaar, C. P. A., N. A. Kinering, and A. J. Berkhout, 1987, Principle of prestack migration based on the full elastic two-way wave equation: *Geophysics*, **52**, 151–173.

Time-lapse wave-equation inversion

Gboyega Ayeni and Biondo Biondi

ABSTRACT

A regularized least-squares inversion scheme is proposed as a method for correcting poor and uneven subsurface illumination under complex overburden and for attenuating image differences resulting from differences in acquisition geometries. This approach involves a joint inversion of migrated images from different vintages with filters that are explicitly derived from the Hessian of the least-squares solution to the linearized wave equation. By using such a formulation, both the imaging and monitoring challenges are solved as a single problem that yields a more accurate image of the subsurface and its time-evolution. Numerical tests show that this joint inversion technique yields more accurate time-lapse images than those obtained by differencing independently migrated or inverted images.

INTRODUCTION

Time-lapse (4D) seismic is based on the premise that changes in reservoir fluid content (associated with production) cause changes in the acoustic properties of rocks and that such changes are detectable in recorded seismic data. Recent developments in seismic acquisition and imaging, and improved understanding of the relationships between seismic data and rock/fluid properties have contributed to the successes in applications of time-lapse seismic to hydrocarbon reservoir monitoring. Although time-lapse imaging of conventional reservoirs is now a mature technology, there has been little success reported in regions of complex geology (e.g., subsalt reservoirs), or in areas where acquisition repeatability is difficult, expensive, or impossible (e.g., because of latter development facilities between surveys).

In recent years, emphasis has been placed on developing better acquisition systems (to ensure high repeatability of surveys and high fidelity datasets) and processing schemes to better discriminate between the effects of reservoir changes from those attributable and unwanted noise. However, time-lapse seismic applications in relatively complex frontier areas, (e.g., deep-water Gulf of Mexico), have been limited mainly because of the imaging difficulties below the complex overburdens and the associated uneven illumination patterns. Also, rapid changes in acquisition techniques throughout the life of most reservoirs usually lead to datasets with different geometries, often causing considerably high-amplitude noise in the time-lapse image.

The proposed joint inversion method utilizes a system of non-stationary deconvolution filters derived from an explicit computation of a target-oriented, least-squares Hessian of the linearized wave-equation (Valenciano et al., 2006). The inversion process attenuates distortions caused by uneven illumination and geometry differences. The joint inversion

scheme enables incorporation of prior knowledge of the reservoir location, extent and geometry into the inversion as a form of model styling (or regularization). An important advantage of the proposed scheme over separate inversion is that different regularization goals can be established for the baseline and time-lapse changes during inversion. Outputs from this scheme include an inverted baseline image and its evolution through time (i.e., time-lapse difference images). A fundamental assumption in the current formulation is that the background baseline velocity model is known and that it does not change significantly between surveys. As opposed to schemes that attempt to directly recover velocity variations between surveys (e.g., Ajo-Franklin et al. (2005)), we assume that changes in reservoir properties can be recovered from the resulting time-lapse reflectivity images.

In this paper, we briefly describe time-lapse seismic monitoring (TLSM) practice and challenges, we summarize the basic idea of linear inversion, and then describe a joint inversion formulation — based on the simultaneous deconvolution of two or more surveys — that attenuates unwanted image differences. Finally, we demonstrate that the joint inversion scheme could be used to correct for both illumination and geometry differences using two synthetic datasets (one from Born single-scattering and the other from two-way variable-density acoustic modeling).

BACKGROUND

The acoustic properties of reservoir rocks are strongly influenced by the fluids stored in their pore spaces. Batzle and Wang (1992) discuss some of the most important seismic properties of reservoir fluids and how they influence seismic properties. Lumley (1995) and Calvert (2005) give comprehensive review of the theory, processing and applications of seismic monitoring. Since its adoption as a reservoir monitoring tool, many successful applications of TLSM have been reported (Lefeuvre et al., 2003; Whitcombe et al., 2004; Zou et al., 2006).

Repeatability is a major consideration for successful application of time-lapse seismic to reservoir monitoring, especially in formations with weak seismic responses, where imaged reservoir changes are subtle. Non-repeatability can result from differences in survey acquisition geometry and binning, cable feathering, tides, source-wavelet bandwidth and phase variability, differential static time-shifts, ambient noise, residual multiple energy, and relative mispositioning of imaged reflection events (Rickett and Lumley, 2001; Johnston, 2005). Laws and Kragh (2000) and Eiken et al. (2003) discuss acquisition techniques that reduce some of these uncertainties.

Some common time-lapse seismic processing schemes (which in general involve among other processes, spatial re-alignment, match filtering, amplitude balancing, and warping) and caveats have been discussed by several authors (Eastwood et al., 1994; Ross and Altan, 1997; Rickett and Lumley, 2001; Calvert, 2005; Johnston, 2005). Although advances in such processing methods have improved the success rate of TLSM, most methods fail in complex geological environments. One dominant reason for such failure is the uneven illumination that cause unreliable seismic amplitudes, thus precluding their use in TLSM. For the example of subsalt reservoirs, large contrasts at salt boundaries, complexity of seismic travel paths, and the associated complex and uneven illumination patterns (Muerdter and Ratcliff, 2001) are major hindrances to application of TLSM. Some recent acquisition techniques (e.g., wide-, rich- and full-azimuth, ocean-bottom, and virtual source surveys)

improve imaging and illumination of reservoirs under complex overburdens (Korneev and Bakulin, 2006; Sava, 2006). Although, such acquisition methods, coupled with recent developments in wave-equation migration, have led more accurate images, they have seen limited applications in seismic monitoring in complex geology. The method discussed in this paper attempts to solve some of the monitoring challenges in these complex environments. The primary objective of the inversion scheme is to compensate for uneven subsurface illumination (and illumination differences when surveys have different geometries), while simultaneously optimizing the estimation of time-lapse changes between surveys.

Although some of the current best practices in TLSM improve the reliability of time-lapse responses and their interpretations, many limitations still exist. Most shortcomings are of lower importance in reservoirs with large seismic responses or low background noise — where such artifacts are submerged by the stronger time-lapse response. However, in many stiff reservoirs and those with complex or uneven illumination patterns (e.g., subsalt reservoirs), small inaccuracies can cause considerably large spurious 4D responses. We envisage that with the gradual increase in demand for more optimal reservoir management (hence the need for more accurate time-lapse amplitudes), and rapid changes in acquisition systems and techniques (hence the need to utilize surveys with potentially widely varying geometries for monitoring), new imaging and monitoring methods will be required to overcome the shortcomings in current time-lapse imaging practice. The need for such new methods is even more apparent with the gradual shift towards more complex exploration frontiers (e.g., deep-water Gulf of Mexico and deep-water West-Africa), where most current methods fail. With the general trend of increased speed of modern computers and the gradual drop in storage cost and computational memory, we believe a joint inversion scheme could be feasible for small- to medium-sized TLSM objectives in the immediate future.

Some authors have discussed joint inversion methods for time-lapse changes of certain reservoir properties. Sarkar et al. (2003) obtained more accurate estimates of changes in reservoir elastic parameters by a coupled impedance inversion of base and monitor datasets. Ajo-Franklin et al. (2005) attenuate some artifacts that persist in image differences from separately inverted datasets by solving a coupled tomographic inversion problem. Also, Zhang et al. (2005) show examples of simultaneous inversion for time-lapse changes using a convolutional model.

Although our primary objective is to invert for reflectivity amplitudes, it is possible to directly extend our formulation to changes in reservoir rock and fluid properties (e.g., pressure and saturation), since there are known relationships between such properties and reflectivity (or impedance) amplitudes. An advantage of this method over those of Sarkar et al. (2003) and Zhang et al. (2005) is that most geometry and illumination artifacts are directly attenuated by the deconvolution operator. Also, we avoid making a constant wavelet assumption throughout (or over large regions) of the model space as is typical in many time-lapse inversion schemes.

In the following sections we discuss the basic theory of wave-equation inversion (and some approximations) and then introduce the joint inversion scheme as an alternative to current standard time-lapse processing methods. Then, using synthetic examples, we demonstrate how the technique could be used to attenuate geometry and illumination differences between two or multiple surveys.

THEORY

Linear inversion

Given a linear modeling operator \mathbf{L} , the seismic data \mathbf{d} can be computed as

$$\mathbf{L}\mathbf{m} = \mathbf{d}, \quad (1)$$

where \mathbf{m} is the reflectivity model. The modeling operator, \mathbf{L} , in this study, is an approximation to the seismic acquisition process. Two different surveys — say a baseline and monitor — acquired at different times ($\mathbf{t} = \mathbf{0}$ and $\mathbf{t} = \mathbf{1}$ respectively) over the same earth model can be represented as follows:

$$\begin{aligned} \mathbf{L}_0\mathbf{m}_0 &= \mathbf{d}_0, \\ \mathbf{L}_1\mathbf{m}_1 &= \mathbf{d}_1, \end{aligned} \quad (2)$$

where \mathbf{m}_0 and \mathbf{m}_1 are respectively the reflectivity models at the times when the datasets \mathbf{d}_0 and \mathbf{d}_1 were acquired, and \mathbf{L}_0 and \mathbf{L}_1 are the modeling operators that defining the acquisition process for the two surveys (a baseline and monitor). We assume that there is not a significant change in the background velocity within such time-interval as to significantly affect wave propagation. The quadratic cost functions for equation 2 are given by

$$\begin{aligned} S(\mathbf{m}_0) &= \|\mathbf{L}_0\mathbf{m}_0 - \mathbf{d}_0\|^2, \\ S(\mathbf{m}_1) &= \|\mathbf{L}_1\mathbf{m}_1 - \mathbf{d}_1\|^2, \end{aligned} \quad (3)$$

and the least-squares solutions are

$$\begin{aligned} \hat{\mathbf{m}}_0 &= (\mathbf{L}_0^\dagger\mathbf{L}_0)^{-1}\mathbf{L}_0^\dagger\mathbf{d}_0 = (\mathbf{L}_0^\dagger\mathbf{L}_0)^{-1}\tilde{\mathbf{m}}_0 = \mathbf{H}_0^{-1}\tilde{\mathbf{m}}_0, \\ \hat{\mathbf{m}}_1 &= (\mathbf{L}_1^\dagger\mathbf{L}_1)^{-1}\mathbf{L}_1^\dagger\mathbf{d}_1 = (\mathbf{L}_1^\dagger\mathbf{L}_1)^{-1}\tilde{\mathbf{m}}_1 = \mathbf{H}_1^{-1}\tilde{\mathbf{m}}_1, \end{aligned} \quad (4)$$

where $\tilde{\mathbf{m}}_0$ and $\tilde{\mathbf{m}}_1$ are the migrated baseline and monitor images, $\hat{\mathbf{m}}_0$ and $\hat{\mathbf{m}}_1$ are the inverted images, \mathbf{L}_0^\dagger and \mathbf{L}_1^\dagger are the migration operators (adjoints to the modeling operators \mathbf{L}_0 and \mathbf{L}_1 respectively), and $\mathbf{H}_0 \equiv \mathbf{L}_0^\dagger\mathbf{L}_0$ and $\mathbf{H}_1 \equiv \mathbf{L}_1^\dagger\mathbf{L}_1$, are the Hessian matrices. Here, and in other parts of the paper, the symbol \dagger denotes transposed complex conjugate. Note that all discussions in this paper are based on (but not limited to) one-way wave-equation migration methods.

The Hessian matrices are the second derivatives of the cost functions (equation 9) with respect to all model points in the image. Further discussion on the wave-equation Hessian is given in the the next section and more detailed treatments of its derivation are found in Plessix and Mulder (2002), Mulder and Plessix (2004), and Valenciano et al. (2006). Because the Hessian matrices are generally not invertible for almost any practical scenario, equation 4 is solved iteratively as follows:

$$\begin{aligned} \mathbf{H}_0\hat{\mathbf{m}}_0 &= \tilde{\mathbf{m}}_0, \\ \mathbf{H}_1\hat{\mathbf{m}}_1 &= \tilde{\mathbf{m}}_1. \end{aligned} \quad (5)$$

An inverted time-lapse image, $\Delta\hat{\mathbf{m}}$, can be obtained as difference between the two images, $\hat{\mathbf{m}}_1$ and $\hat{\mathbf{m}}_0$, obtained from equation 5:

$$\Delta\hat{\mathbf{m}} = \hat{\mathbf{m}}_1 - \hat{\mathbf{m}}_0. \quad (6)$$

Computing the time-lapse image using equation 6 is regarded as *separate inversion* in other parts of this paper.

Joint Inversion

Another approach is to express the inversion as a single problem, in which both the baseline and time-lapse images are estimated simultaneously. We write the data modeling operation as follows:

$$\begin{aligned} \mathbf{L}_0 \mathbf{m}_0 &= \mathbf{d}_0, \\ \mathbf{L}_1 (\mathbf{m}_0 + \Delta \mathbf{m}) &= \mathbf{d}_1, \end{aligned} \quad (7)$$

where $\mathbf{m}_0 + \Delta \mathbf{m} = \mathbf{m}_1$. In matrix form, we can combine these expressions to write

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \Delta \mathbf{m} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}. \quad (8)$$

In principle, using an iterative solver, we can solve for the least-squares solution to equation 8 by minimizing the cost function

$$S(\mathbf{m}_0, \Delta \mathbf{m}) = \left\| \begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \Delta \mathbf{m} \end{bmatrix} - \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} \right\|^2. \quad (9)$$

The cost of this approach is proportional to the number of iterations times twice the cost of one set of migrations — since one iteration requires one modeling and one migration for the baseline and monitor datasets. Since several iterations would typically be required to reach convergence, and the inversion process would usually be repeated several times to fine-tune inversion parameters, the overall cost of this scheme may be considerably high. One advantage of the method discussed in the remaining part of this section is that modifications can be made to inversion parameters and the inversion repeated several times without the need for new migration or modeling. The least-squares solution to equation 8 is given by

$$\begin{bmatrix} \mathbf{L}_0^\dagger \mathbf{L}_0 + \mathbf{L}_1^\dagger \mathbf{L}_1 & \mathbf{L}_1^\dagger \mathbf{L}_1 \\ \mathbf{L}_1^\dagger \mathbf{L}_1 & \mathbf{L}_1^\dagger \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_0^\dagger & \mathbf{L}_1^\dagger \\ \mathbf{0} & \mathbf{L}_1^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \quad (10)$$

or

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \quad (11)$$

This can be re-casted as

$$\begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_1 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \quad (12)$$

Thus, the inverted baseline and time-lapse images ($\hat{\mathbf{m}}_0$ and $\Delta \hat{\mathbf{m}}$ respectively) can be obtained from equation 12. As earlier noted, the Hessian matrices \mathbf{H}_0 and \mathbf{H}_1 (and hence the joint deconvolution operator) are not invertible, equation 11 must be solved iteratively.

In most TLSM problems, the general geology and reservoir architecture of the study area are known — thus providing some information that can be used to determine appropriate regularization for the inversion. Such regularization incorporates prior knowledge of the reservoir geometry and location, and expectation of changes in different parts of the study area. For example, equation 12 can be extended to include such prior information, to give

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_1 \\ \epsilon_0 \boldsymbol{\Lambda}_0 \mathbf{R}_0 & \mathbf{0} \\ -\epsilon_1 \boldsymbol{\Lambda}_1 \mathbf{R}_0 & \epsilon_1 \boldsymbol{\Lambda}_1 \mathbf{R}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad (13)$$

where, \mathbf{R}_0 and \mathbf{R}_1 are the spatial regularization terms for the baseline and time-lapse images respectively, $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ are temporal regularization (or coupling) between the surveys, while ϵ_0 and ϵ_1 determine strength of the regularization on the baseline and time-lapse images. The spatial and coupling operators can be applied over one or several dimensions (e.g., stacked-image, subsurface offset, or subsurface opening-angles), depending on the problem size, computational domain and a priori information available. While the spatial regularization operator contains information on the structural geometry of the reservoir (or implied properties of correctly migrated gathers, e.g. horizontal angle gathers, or near-zero concentration of amplitudes in subsurface offset gathers), the temporal regularization ensures that the reservoir changes evolve relatively smoothly. The temporal regularization term is similar to that used in the so-called spatio-temporal tomographic inversion demonstrated by Ajo-Franklin et al. (2005).

We have extended this development to include additional surveys using the same procedure above (see Appendix A). For example, for three surveys (one baseline and two monitors), we can obtain the time-lapse images from

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_2 & \mathbf{H}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}}_1 \\ \Delta \hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_2 \end{bmatrix}, \quad (14)$$

and including a regularization operator, we have

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_2 & \mathbf{H}_2 \\ \hline \epsilon_0 \mathbf{\Lambda}_0 \mathbf{R}_0 & \mathbf{0} & \mathbf{0} \\ -\epsilon_1 \mathbf{\Lambda}_1 \mathbf{R}_0 & \epsilon_1 \mathbf{\Lambda}_1 \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & -\epsilon_2 \mathbf{\Lambda}_2 \mathbf{R}_0 & \epsilon_2 \mathbf{\Lambda}_2 \mathbf{R}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}}_1 \\ \Delta \hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_2 \\ \hline \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (15)$$

Conceptually, equation A-7 can be useful in cases where several relatively small or sparse surveys, with different geometries are conducted at relatively short intervals, as proposed for example in the seismic monitoring of CO_2 sequestration (Arogunmati and Harris, 2007). We note that because different surveys would have some coincident shot or receiver locations, the overall computational cost of the Hessian (see next section) for different surveys is reduced.

Target-oriented Hessian

The cost of computing the full Hessian matrices (\mathbf{H}_1 to \mathbf{H}_N) in equation A-7 is prohibitive and not practical for almost any reasonably sized survey. Several authors (Shin et al., 2001; Rickett, 2003; Guitton, 2004; Valenciano et al., 2006; Symes, 2008) discussed common approximations to the Hessian. Also, schemes such as the random-phase and plane-wave encoding of the source and receiver Green's functions (Tang, 2008) ensure highly efficient approximate Hessian computation. In this paper, we follow the target-oriented approach of Valenciano et al. (2006).

Under the Born approximation, the synthetic seismic data, $\mathbf{d}(\mathbf{s}, \mathbf{r}; \omega)$ at a given frequency, ω , recorded by receiver $\mathbf{r}(x_r, y_r, z_r)$, from a shot $\mathbf{s}(x_s, y_s, z_s)$ and scattering point

$\mathbf{x}(x, y, z)$, is given by

$$\mathbf{d}(\mathbf{s}, \mathbf{r}; \omega) = \omega^2 \sum_{\mathbf{x}} f(s) \mathbf{G}(\mathbf{x}, \mathbf{s}; \omega) \mathbf{G}(\mathbf{x}, \mathbf{r}; \omega) \mathbf{m}(\mathbf{x}), \quad (16)$$

where $f(s)$ is the source function, $\mathbf{G}(\mathbf{x}, \mathbf{s}; \omega)$ and $\mathbf{G}(\mathbf{x}, \mathbf{r}; \omega)$ are the Green's functions from the shot and receiver positions to the image point \mathbf{x} , with reflectivity $\mathbf{m}(\mathbf{x})$. Equation 16 describes the linear modeling process in equation 9.

The adjoint of the modeling operation in equation 16 is termed migration, which expressed in shot-profile formulation is

$$\tilde{\mathbf{m}}(\mathbf{x}) = \sum_w \omega^2 \sum_s f^\dagger(s) \mathbf{G}^\dagger(\mathbf{x}, \mathbf{s}; \omega) \sum_r \mathbf{G}^\dagger(\mathbf{x}, \mathbf{r}; \omega) \mathbf{d}(\mathbf{s}, \mathbf{r}; \omega), \quad (17)$$

where $\tilde{\mathbf{m}}(\mathbf{x})$ is the migrated image. The migration operator is equivalent the negative gradient of the least-squares cost function (Tarantola, 1987) defined in equation 9, while the Hessian is (Plessix and Mulder, 2002; Valenciano et al., 2006)

$$\mathbf{H}(\mathbf{x}, \mathbf{y}) = \sum_w \omega^4 \sum_s \left| f^\dagger(s) \right|^2 \mathbf{G}^\dagger(\mathbf{x}, \mathbf{s}; \omega) \mathbf{G}(\mathbf{y}, \mathbf{s}; \omega) \sum_r \mathbf{G}^\dagger(\mathbf{x}, \mathbf{r}; \omega) \mathbf{G}(\mathbf{y}, \mathbf{r}; \omega), \quad (18)$$

where $\mathbf{y}(x, y, z)$ corresponds to all model points.

In order to explicitly compute the Hessian using equation 18, Green's functions from the surface shot and receiver positions to all image points need to be computed. The post-stack Hessian matrix is of size $N_x \times N_y \times N_z$ by $N_x \times N_y \times N_z$, where N_i is the number of image points along the i -axis. Including secondary dimensions such as the subsurface offsets or the subsurface opening-angle, the cost of computing equation 18 increases considerably (in proportion to the number of components along such dimension). It is not practicable to compute or store such a matrix, even for small surveys.

Because reservoirs typically are limited in extent, the region of interest is usually smaller than the full image space. The target-oriented approach (Valenciano et al., 2006) involves an explicit computation of Green's and the Hessian for a specific area of interest and for only a limited number of off-diagonal elements, (i.e., points around any image point). This method also takes into account the sparsity and structure of the Hessian, as well as particular acquisition geometry to greatly reduce the computational cost of the Hessian.

The target-oriented approach reduces equation 18 to

$$\mathbf{H}(\mathbf{x}_T, \mathbf{x}_T + \mathbf{a}_x) = \sum_w \omega^4 \sum_s \left| f^\dagger(s) \right|^2 \mathbf{G}^\dagger(\mathbf{x}_T, \mathbf{s}; \omega) \mathbf{G}(\mathbf{x}_T + \mathbf{a}_x, \mathbf{s}; \omega) \sum_r \mathbf{G}^\dagger(\mathbf{x}_T, \mathbf{r}; \omega) \mathbf{G}(\mathbf{x}_T + \mathbf{a}_x, \mathbf{r}; \omega), \quad (19)$$

where \mathbf{a}_x is the offset from the target image-point \mathbf{x}_T defining the deconvolution filter size and hence the number of off-diagonal terms to be computed. The filter size \mathbf{a}_x can be determined heuristically or from an analysis of the amplitudes of filter coefficients away from the diagonal. As noted by Valenciano (2008), the frequency sampling required to prevent wrap-around artifacts for the local filter (or row of the Hessian) for a given image point is coarser than that used in that used in migration. This further reduces the overall computational cost of equation 19. Examples of the target-oriented Hessian for models in Figures 1 and 2 for two and three surveys are shown in Figures 3 and 4.

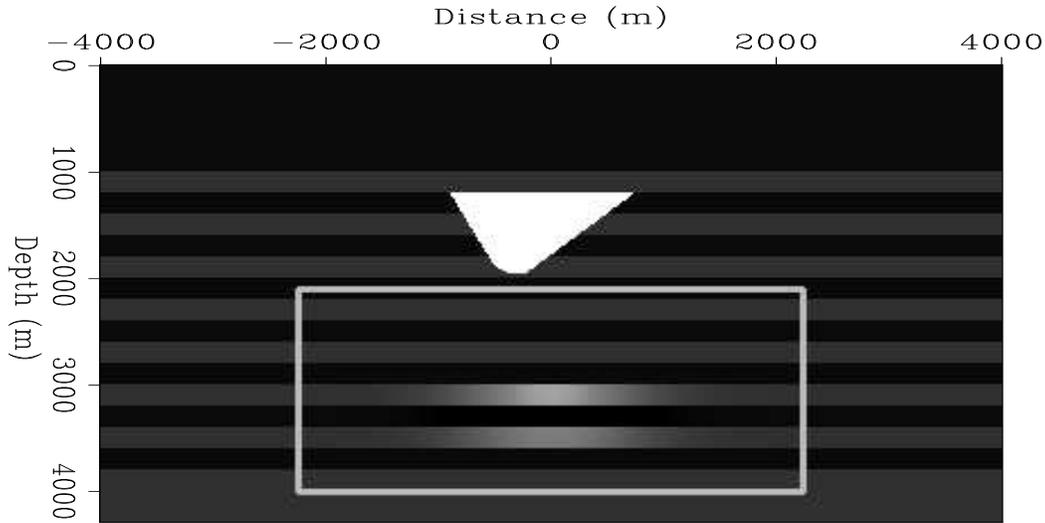


Figure 1: Full velocity model. The box indicates target area for which Figures 3 and 4 were computed, while the anomaly centered at distance 0m and depth 3000m represents the approximate location of reservoir change. The triangular block is a salt with velocity 4500m/s, while the surrounding sediments have alternating velocity of 2500m/s and 2700m/s. The densities range from 2.5g/cc to 3.0g/cc. [ER] gayeni1/. velsalt

In a single survey, each row of the Hessian is a convolution filter (or point-spread function) describing the effects of the limited-bandwidth seismic waveform, geometry and illumination on a reflectivity spike in the subsurface. In multiple surveys, each band belonging to individual sub-matrices contains similar information from a single or combination of surveys as shown in equation A-7. Given the block and sparse matrix structure, we can use fast mathematical formulations specifically designed for such systems. Also, note that the empty bins in Figures 3 and 4 are neither computed nor stored and that because of the matrix symmetry, only one-half of it needs to be computed. The structure of the problem gives a large leeway for parallelization over several domains in the Hessian computation and inversion. Finally, since we assume that there is not a significant variation in the background velocity between surveys, and since some shot and receiver locations would be re-occupied during the monitor survey(s), the Green's functions can be reused. In the following sections, we demonstrate this inversion approach and compare results with those obtained using two approximations to the Hessian.

NUMERICAL EXAMPLES

We simulated changes in the reservoir properties (resulting from production) at different times during the production cycle as an increasing Gaussian anomaly centered at $x = 0\text{m}$ and $z = 3000\text{m}$ for the 2D model shown in Figure 1. The modeling parameters for the two survey geometries considered are summarized in Table 1.

The baseline survey was modeled with geometry A, while two monitors were modeled with were modeled (for the same time-period) using geometries A and B. In all examples considered, the receiver spread is kept constant while the shots moves along. Also, we

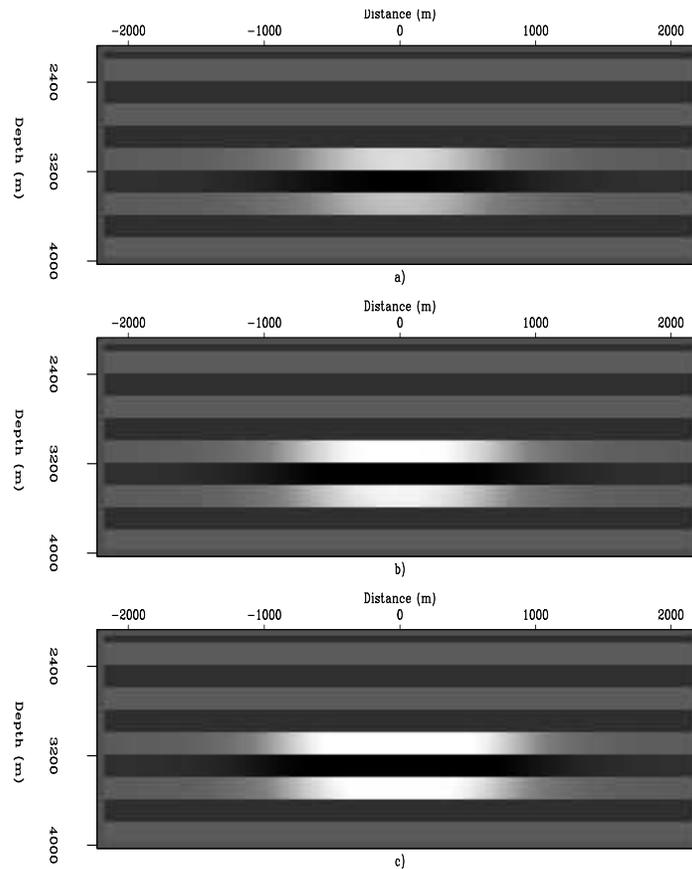


Figure 2: Reservoir models for the (a) baseline, (b) monitor, and (c) monitor 2 conditions around target area (distance, $x = -2200$ to $2200m$ and depth, $z = 2500$ to $4000m$). It is assumed that the reflectivity increases with time due to production. [ER] `gayeni1/. velsaltall`

Figure 3: Joint target-oriented Hessian for two surveys (one baseline and one monitor) for the subsalt reservoir models in Figure 1. The dimension of the square matrix here and in Figure 4 is equal to the number of surveys times the size of the model space. This figure corresponds to the operator in equation 12. Note however that the zeros (light regions in the matrices) were neither computed nor stored. [CR] `gayeni1/. hesssalt2`

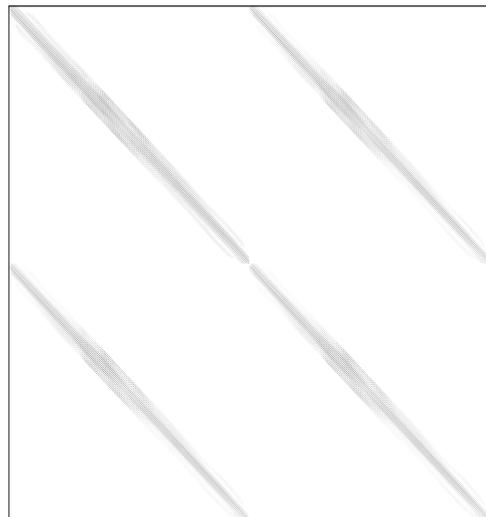


Figure 4: Joint target-oriented Hessian for one baseline and two monitor surveys for the reservoir models in Figure 1. See caption in Figure 3 for further description. [CR] gayeni/. hesssalt3

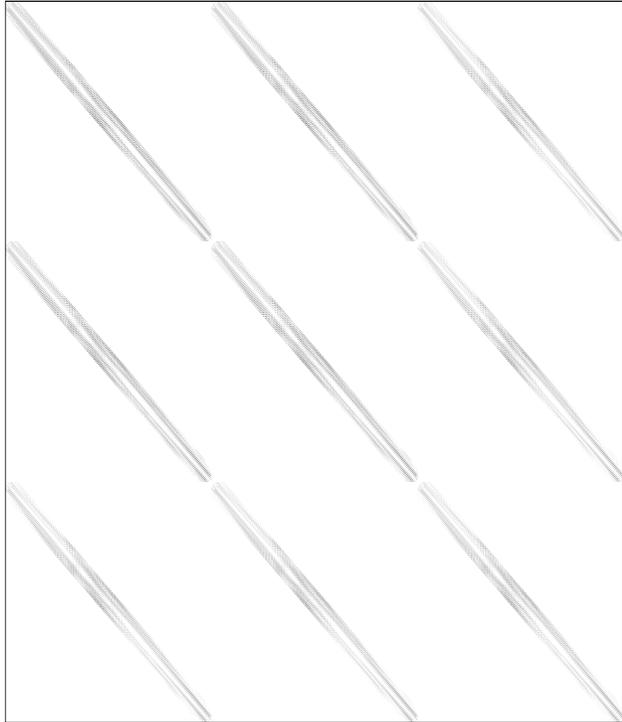


Table 1: Modeling parameters for synthetic datasets

	Geometry A	Geometry B
No. of shots	91	91
No. of receivers	271	271
Shot/receiver depth	0m	200m
Shot spacing	60m	60m
Receiver spacing	20m	20m
Shot/receiver spread	-2700 to 2700m	-3000 to 2400m

consider that reflectivity change is most influenced by a change in the density within the reservoir and that there is not a significant change in the background velocity model between surveys. This is needed in order to keep reflectors at lower depths than the reservoir top from significantly changing positions after migration with the baseline velocity.

The illumination maps (the diagonal of the Hessian) for both survey geometries are shown in Figure 6, while the migrated baseline and monitor images are shown in Figure 5. Note the irregular illumination pattern in Figure 6, which explain the uneven amplitudes of reflectors below the salt in Figure 5. The migrated results in this paper are obtained using a one-way shot-profile migration code. In the following sections we summarize some of the important results from the numerical tests.

Synthetic I: Born data

To demonstrate the methodology for noise-free datasets where all events are completely defined by the one-way wave-equation, we simulated single scattering (Born) datasets for

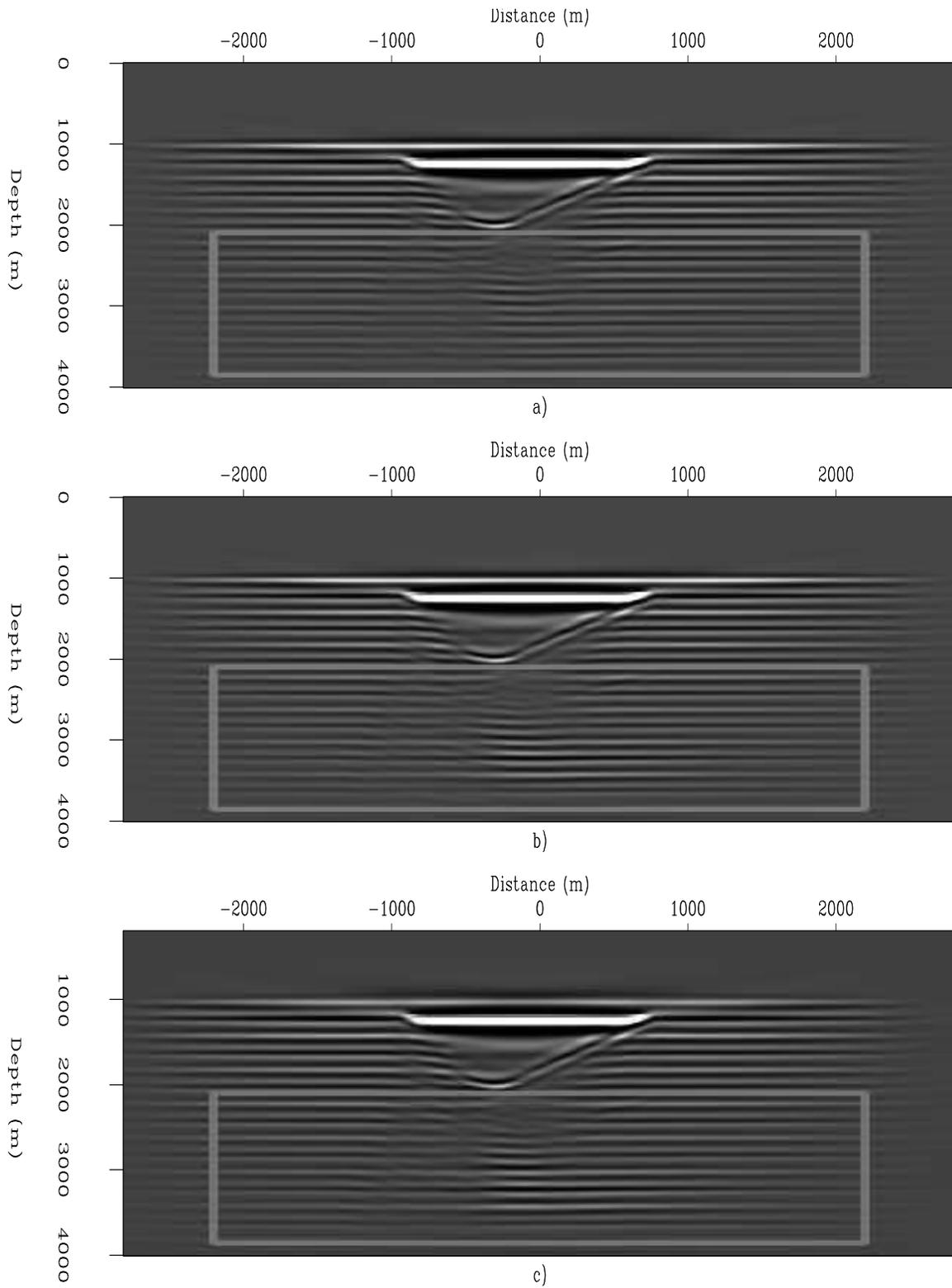


Figure 5: Migrated images for the (a) baseline, and monitor surveys with (b) same geometry as baseline, (c) different geometry from baseline. Note the irregular amplitude pattern below the salt structure. The box indicates the target area. [CR] `gayeni/. migsalt`

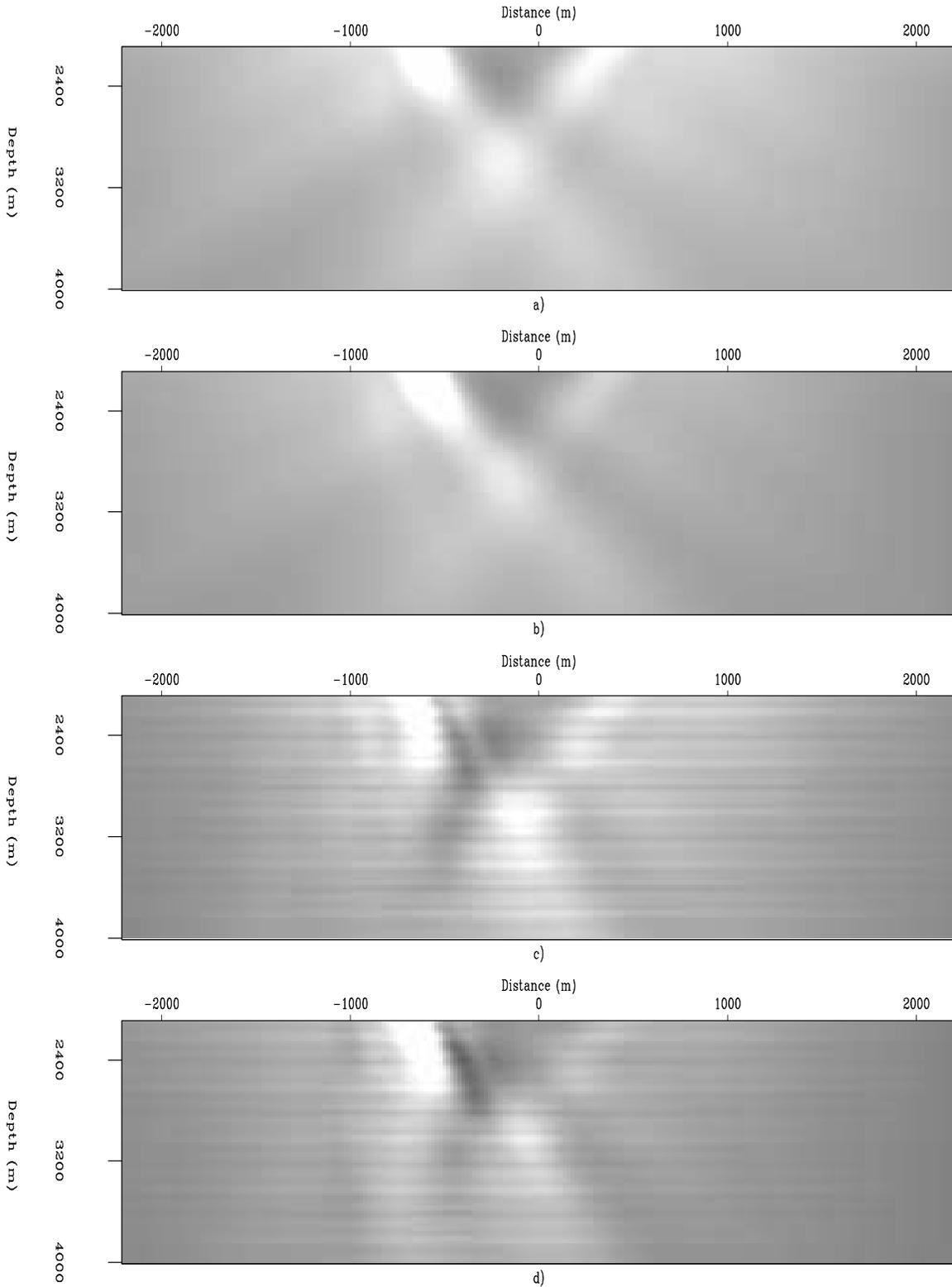


Figure 6: Illumination maps: (a) and (b) are the diagonal of the Hessian for the two geometries (A and B respectively) described in the text. (c) and (d) are respectively, the mass-lumped approximations to the Hessian (derived by summing a few off-diagonal terms in the Hessian into its diagonal) for geometries A and B. The light colors indicate high and the dark low illumination. [CR] `gayeni1/. illumb`

the baseline and monitor surveys. Because this is equivalent to the convolution of the one-way Hessian with the subsurface reflectivity, the wave-propagation and scattering are completely defined by the the Hessian. Figure 7 shows time-lapse images obtained from shot-profile migration and inversion, while Figure 8 shows a comparison of the time-lapse amplitudes along the top of the reservoir (at depth $z = 3000\text{m}$). For comparison, we include the time-lapse images obtained by weighting the migrated images with the diagonal of the Hessian (Duquet et al., 2000) and with a mass-lumped approximation (Chavent and Plessix, 1999) in both Figures 7 and 8. Only one inversion result is shown in both figures as the separate and joint inversions yield exactly the same result. With a different monitor geometry, B , the inversion goals are simultaneous corrections for both the illumination and geometry artifacts. Figures 9 and 10 show the corresponding the time-lapse results obtained using different geometries for the two surveys.

Next, we investigate the effect of uncorrelated noise on both the separate and joint inversion schemes, since in practice the migrated image is not entirely noise-free. We added different levels of uniformly-distributed random noise to the migrated images and try to invert the using the same operators as those in the noise-free case. The inversion results from this exercise are presented in Figure 11. Note that for different amount of noise in the migrated images, the separately inverted time-lapse images are consistently higher than those obtained from joint inversion.

Synthetic II: Two-way modeling

We modeled synthetic two-way data with a time-domain variable density acoustic wave-equation. Since one-way operators were used in the computation of the Green's functions, the Hessian contains no information regarding secondary events such as multiples. No pre-processing was done to remove multiples from the data and hence we expect these to have some impact on the inversion. For perfectly repeated surveys (i.e., using geometry A for both baseline and monitor surveys), the time-lapse images obtained from differences of migrated images, weighted migrated images (using diagonal of the Hessian and the mass-lumped approximation), and inversion are shown in Figure 12. Both the separate and joint inversion results were obtained with regularization incorporating structural information of the reservoir. Note that there is little improvement in the joint inversion result over separate inversion. With non-repeated survey geometries (i.e. geometry A for the baseline and B for the monitor survey), the images obtained from migration, weighted migration and inversion are shown in Figure 13. Note that the image from joint inversion has less artifacts compared than other methods.

DISCUSSION

Results from inversion of the Born datasets suggest that it is possible to recover the true time-lapse reflectivity if the Hessian accurately models the propagation and scattering of waves in seismic experiments. As shown in Figures 7 and 8, the images recovered from weighting migrated images with either the diagonal of the Hessian or its mass-lumped equivalent are inferior to those from inversion. This is because these weights contain no information regarding the shape of the filter, resulting from the finite frequency of the seismic waveform, direction of illumination and local dips of reflectors — information captured in

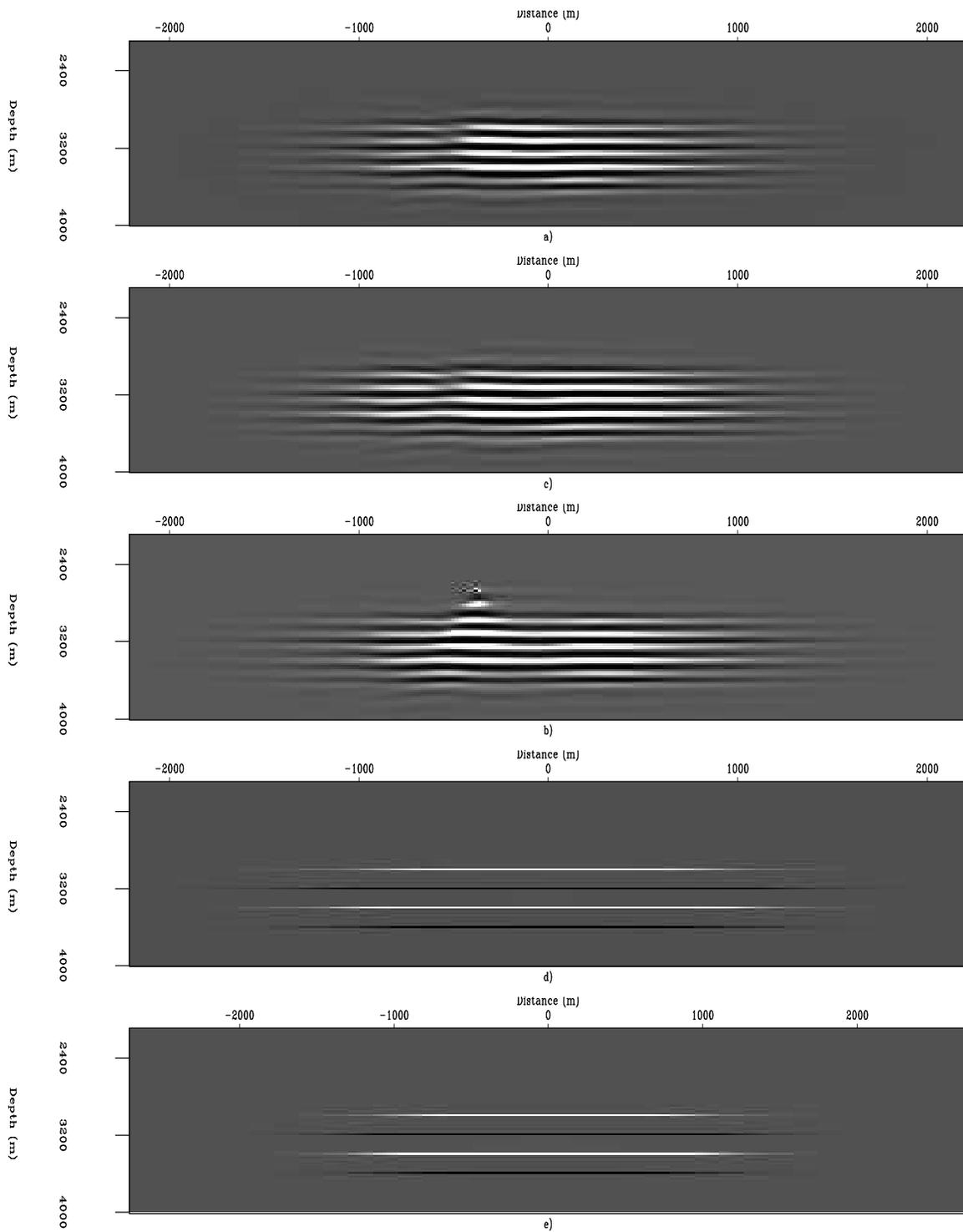


Figure 7: Time-lapse reflectivity amplitudes for perfectly repeated surveys derived from (a) migration, migration weighted by the (b) diagonal of the Hessian, and (c) diagonal matrix from the mass-lumping approximation. (d) shows the image difference from joint inversion — result from separate inversion is not show because the results are the same as joint inversion— and (e) is the true reflectivity change. [CR] `gayeni1/.comparemod1`

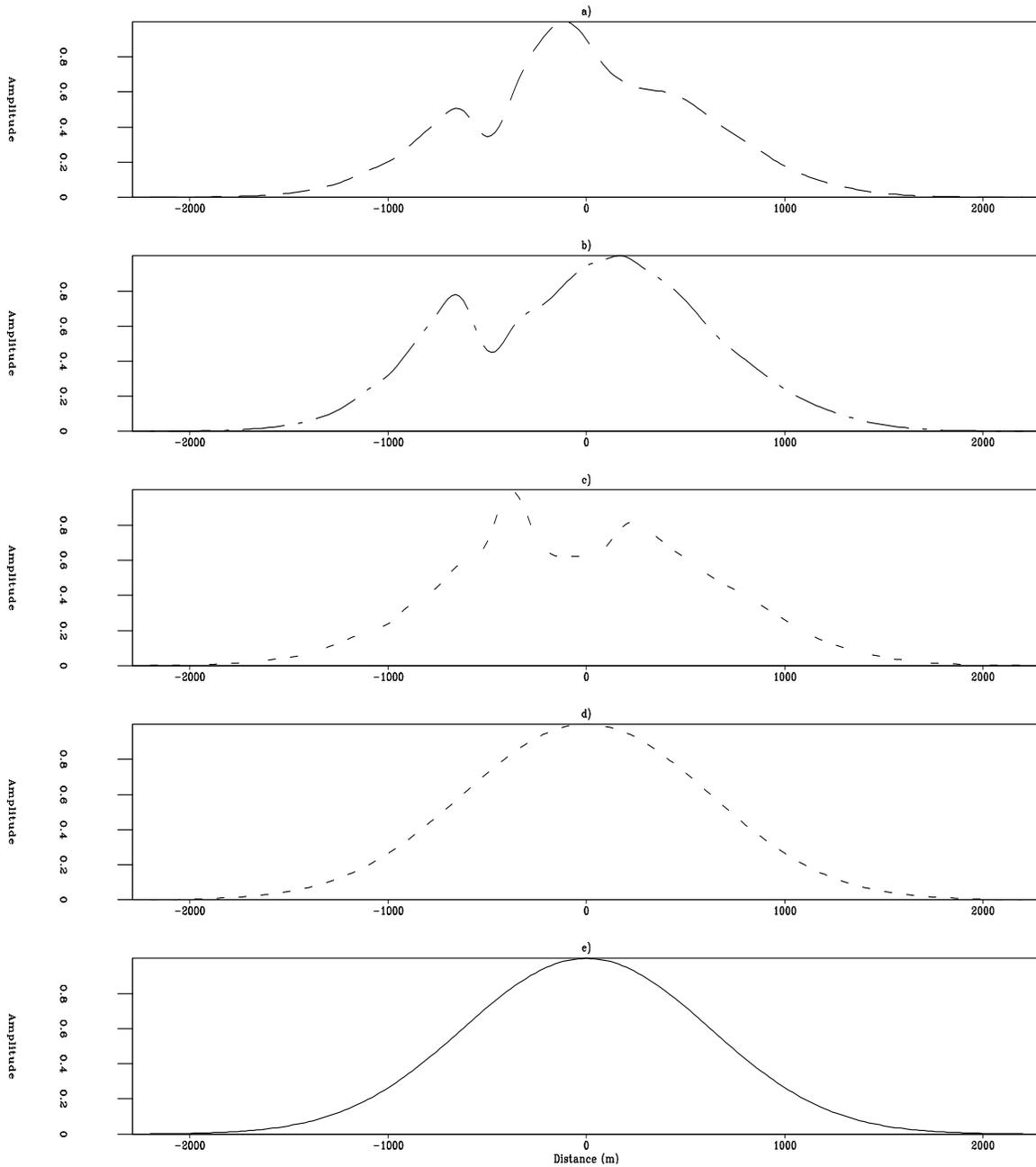


Figure 8: Normalized time-lapse amplitudes above the reservoir (depth 3000m) for perfectly repeated surveys. Each plot corresponds to the images in Figure 7. Note that in the migrated and weighted images, the reflectivity change is strongly distorted and that inversion perfectly reproduces the true reflectivity change. [CR] `gayeni1/. comparemod3000`

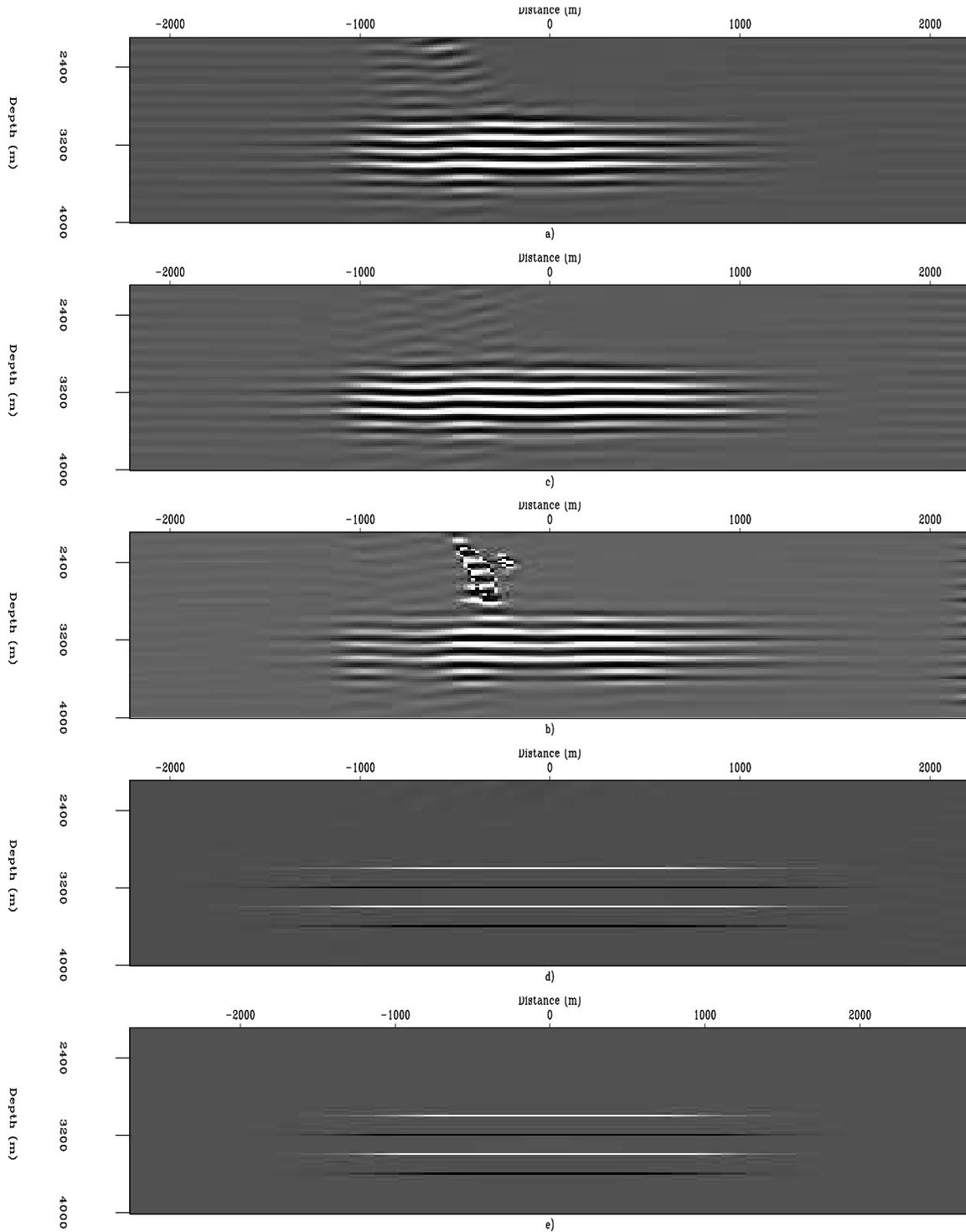


Figure 9: Time-lapse reflectivity amplitudes for different baseline and monitor geometries. See caption in Figure 7. By comparing (a) to Figure 7a, note that above the reservoir, there is a large image difference caused by both uneven illumination and the different survey geometries. Also note the poor performance of the diagonal weighting in (b) and (c), compared with inversion (d), which perfectly reproduces the true reflectivity change shown in (e). [CR] `gayeni/. comparemod2`

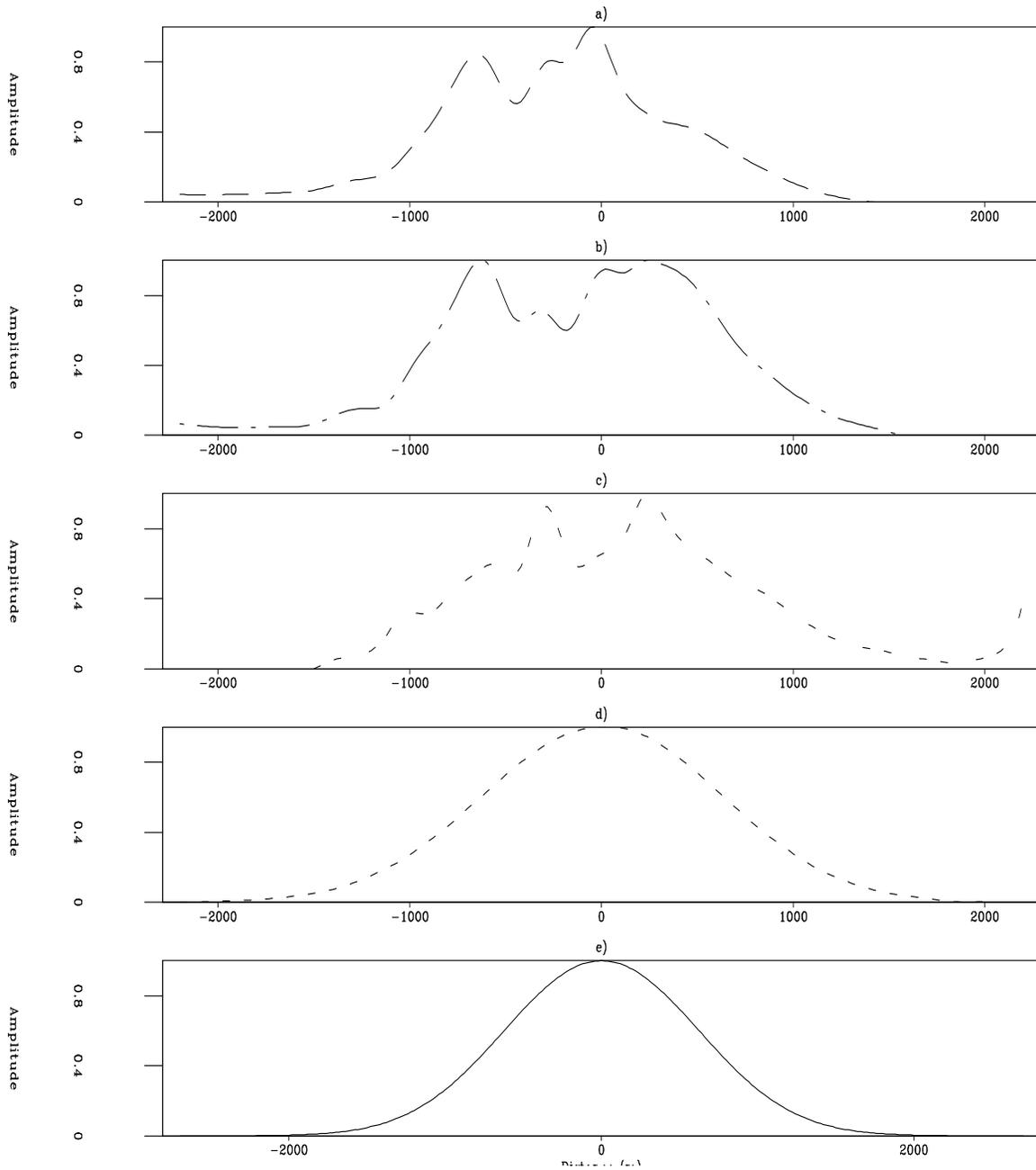


Figure 10: Normalized time-lapse amplitudes at top of the reservoir, (depth 3000m) for perfectly repeated surveys. Each plot corresponds to the images in Figures 9. Compare these to Figures 8. Note that in the migrated and weighted images, the reflectivity change is distorted and that inversion perfectly reproduces the true reflectivity change. [CR]

gayeni/. comparemodshift3000

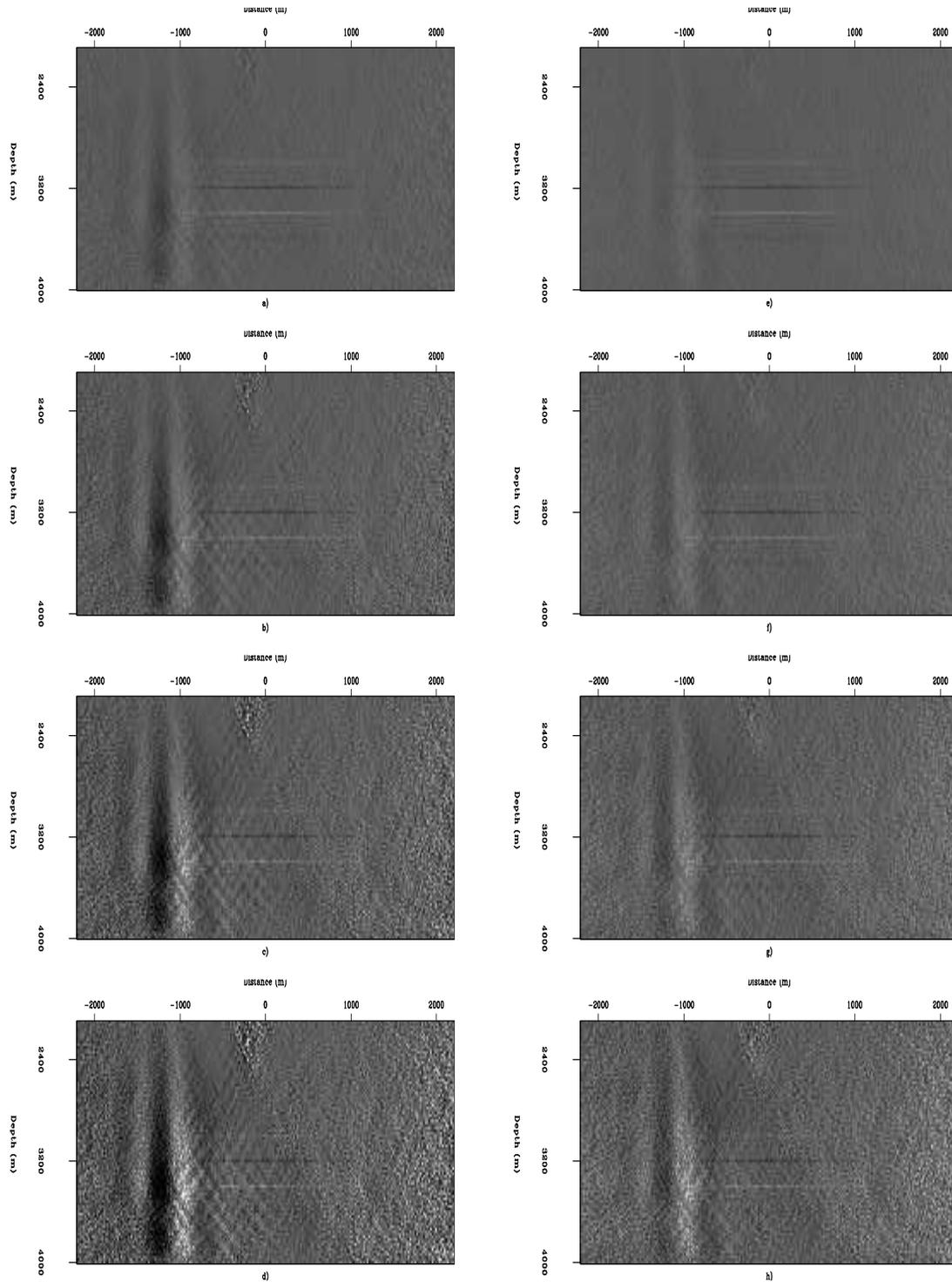


Figure 11: Time-lapse images obtained from separate, (a) - (d) and joint, (e) - (h) inversion with different (from top, 2.5, 5.0, 7, 5, and 10.0%) levels of uniform random noise added to the migrated images. The baseline and monitor surveys were modeled with the different geometries (A and B respectively). Note that the noise level is higher in the separately inverted images relative to the jointly inverted images ones, especially in parts of the model with poor illumination. [CR] [gayeni1/. comparenoise](#)

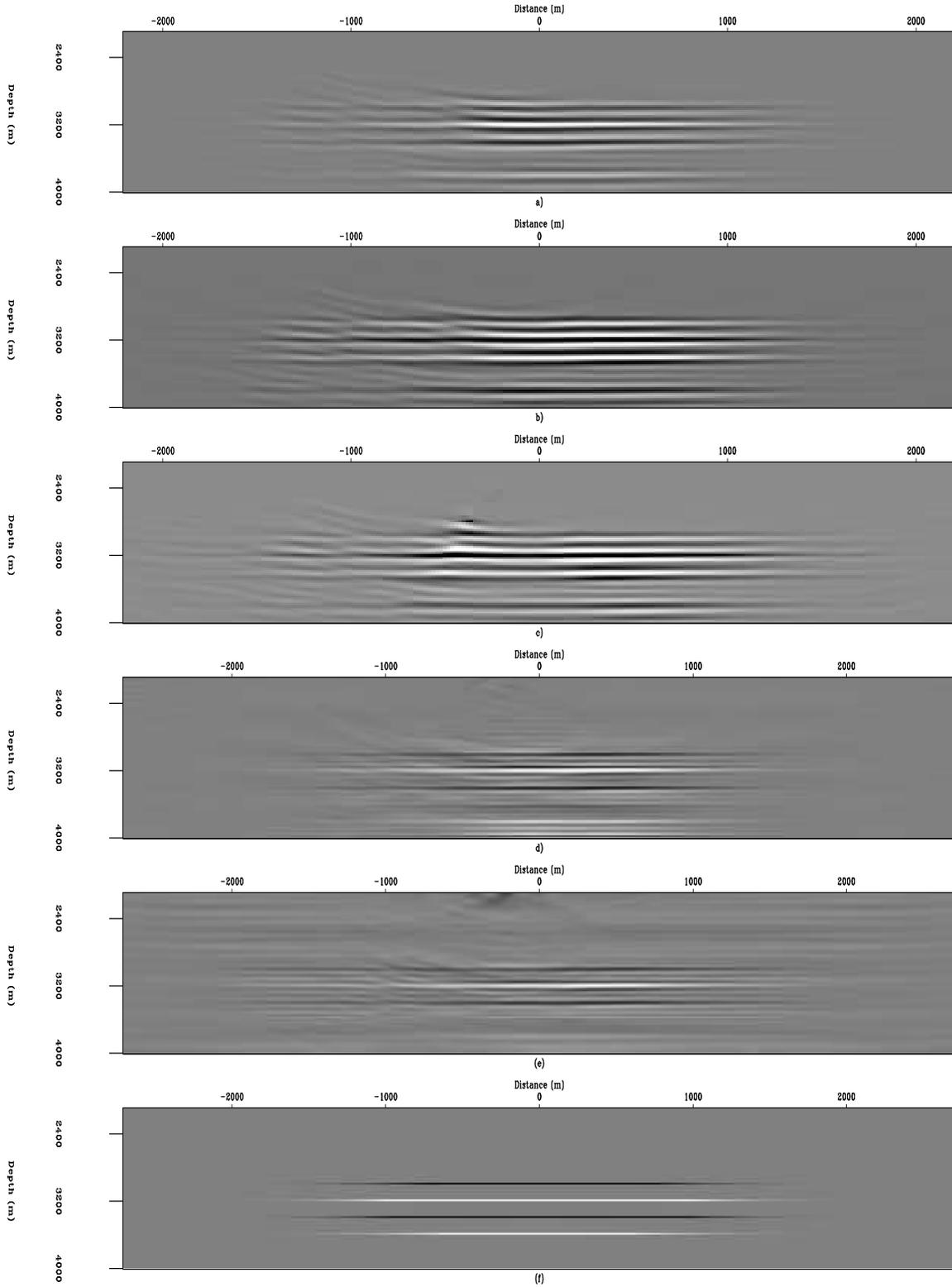


Figure 12: Time-lapse reflectivity for perfectly repeated survey geometries. The images are derived from (a) migration, weighted migrated images using the (b) diagonal, and (c) mass-lumping approximation, (d) separate inversion, (e) joint inversion, and (f) is the true reflectivity change. [CR] [gayeni1/. comparemiginv](#)

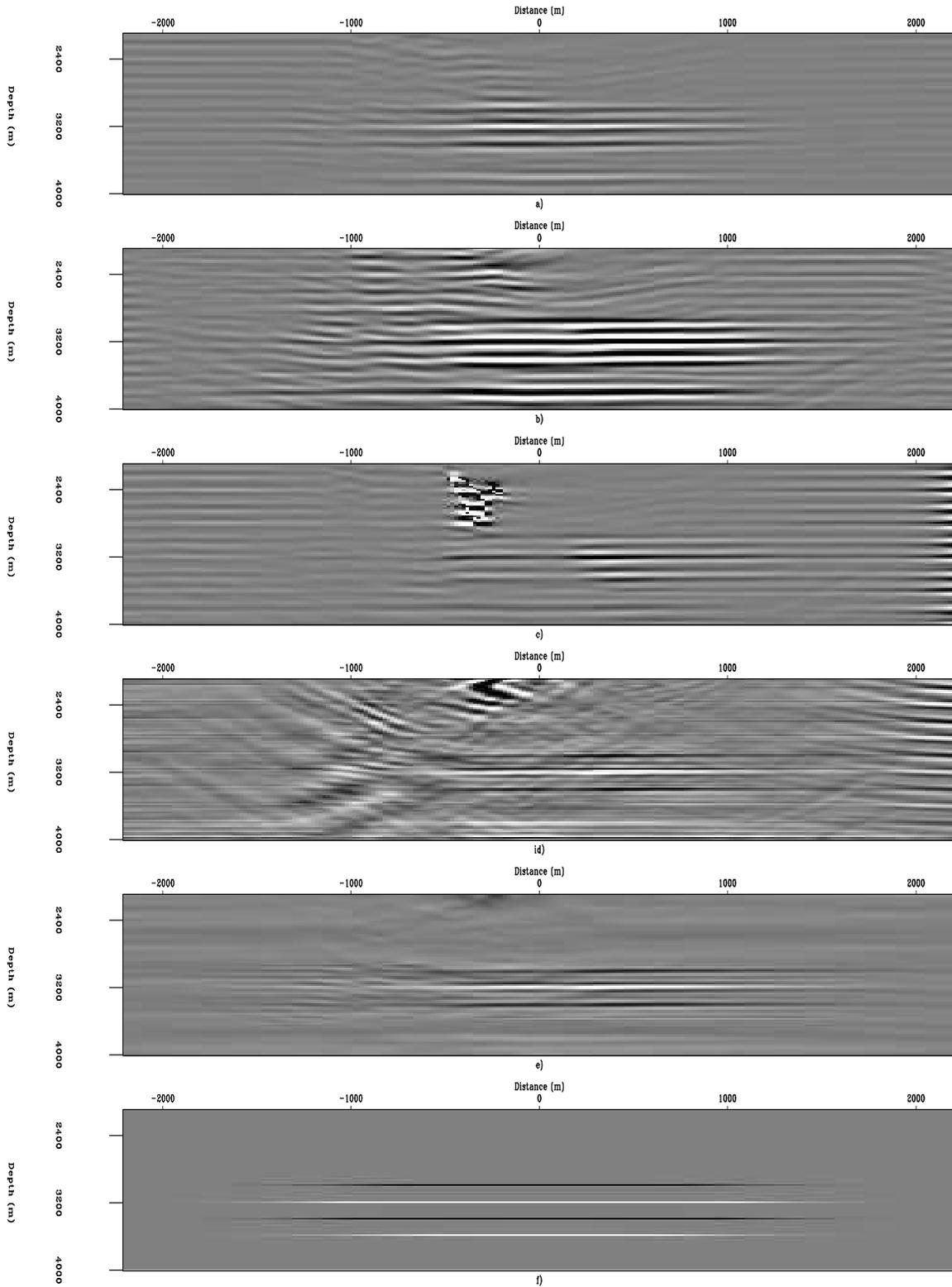


Figure 13: Time-lapse reflectivity using different survey geometries for baseline (geometry A) and monitor (geometry B). The images were derived from (a) migration, weighted migrated images using the (b) diagonal, and (c) mass-lumping approximation, (d) separate inversion, (e) joint inversion, and (f) is the true reflectivity change. [CR]

gayenil/. comparemigshiftinv

the off-diagonal elements of the target-oriented filter used in the inversion. The limitation of such weights become even more apparent when the survey geometries are not repeated as shown in Figures 9 and 10, where the inversion recovers the true reflectivity change but the weighting operators perform poorly. In the absence of noise both separate and joint inversion yield comparable results; however with uncorrelated random noise present the joint inversion yields superior images as shown in Figure 11. An intuitive explanation for such improvement is that by using a combination of operators over the same model, the overall effect of the noise would reduce, since the iterative search becomes more robust. The inversion results from the two-way data shows the improvement derivable from both separate and joint inversion as shown in Figures 12 and 13. Note that, though there is limited improvement in the joint inversion image relative to that from separate inversion when the geometries are the same (Figures 12d and e), there is remarkable improvement over separate inversion when the geometries are different as shown in Figures 13d and e. Weighting operators (e.g. diagonal of the Hessian, a mass-lumping, or other approximation) are an over-simplification of the inversion and therefore yield inaccurate results as shown in Figures 13b and c. As shown in all the inverted images, the seismic wavelet has been deconvolved from the time-lapse, thus increasing the resolution.

CONCLUSIONS

Seismic monitoring of reservoirs under complex overburden is challenging due to a combination of factors including uneven illumination and non-repeatability of acquisition geometries. We have proposed a joint inversion scheme to solve both the imaging and monitoring problems typical of such complex geological environments. The numerical examples show that we can address illumination and geometry problems, that cause contamination the time-lapse seismic responses below complex overburden. We speculate this scheme could make reservoir monitoring in subsalt (and other complex) reservoirs feasible. This scheme is also more robust than migration and inversion using the most common approximations to the Hessian. In the presence of random uncorrelated noise, this inversion scheme would yield better results than separately inverting the migrated images. We concede that the overall cost of the Hessian is still outside reach for most reasonably large surveys, but by using the target-oriented approach discussed in the paper, particular areas of interest can be sufficiently handled. We do believe that with new computational methods such as phase- or plane-wave encoding, increased computational speed and storage capacity of computers, and the need to monitor reservoirs in complex frontiers, the joint inversion method proposed here would become practical. Finally, although separate inversion is computationally less expensive, joint inversion provides more flexibility in model styling/regularization during inversion.

ACKNOWLEDGMENTS

The first author acknowledges contributions from Alejandro Valenciano and Yaxun Tang, whose previous and current research have helped him through understanding wave-equation inversion and efficient computational methods for the Hessian. Discussions with Bob Clapp and Claudio Cardoso helped shape this research.

REFERENCES

- Ajo-Franklin, J. B., J. Urban, and J. M. Harris, 2005, Temporal integration of seismic traveltimes tomography: SEG Technical Program Expanded Abstracts, **24**, 2468–2471.
- Arogunmati, A. and J. Harris, 2007, Data integration for spatiotemporal imaging.
- Batzle, M. and Z. Wang, 1992, Seismic properties of pore fluids: *Geophysics*, **57**, 1396–1408.
- Calvert, R., 2005, Insights and methods for 4D reservoir monitoring and characterization: SEG/EAGE DISC (Distinguished Instructor Lecture Course).
- Chavent, G. and R. E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: *Geophysics*, **64**, 508–515.
- Duquet, B., K. J. Marfurt, and J. Dellinger, 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: *Geophysics*, **65**, 1195–1209.
- Eastwood, J., P. Lebel, A. Dilay, and S. Blakeslee, 1994, Seismic monitoring of steam-based recovery of bitumen: *The Leading Edge*, **13**, 242–251.
- Eiken, O., G. U. Haugen, M. Schonewille, and A. Duijndam, 2003, A proven method for acquiring highly repeatable towed streamer seismic data: *Geophysics*, **68**, 1303–1309.
- Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: *Geophysics*, **69**, 1017–1024.
- Johnston, D., 2005, Time-lapse 4D technology: Reservoir surveillance: AAPG Search and Discovery.
- Korneev, V. and A. Bakulin, 2006, On the fundamentals of the virtual source method: *Geophysics*, **71**, A13–A17.
- Laws, R. and E. Kragh, 2000, Rough seas and time-lapse seismic: 62nd Mtg., Session:L0015, Eur. Assn. Geosci. Eng.
- Lefeuvre, F., Y. Kerdraon, J. Peliganga, S. Medina, P. Charrier, R. L’Houtellier, and D. Dubucq, 2003, Improved reservoir understanding through rapid and effective 4D: Girassol field, Angola, West Africa: SEG Technical Program Expanded Abstracts, **22**, 1334–1337.
- Lumley, D. E., 1995, Seismic time-lapse monitoring of subsurface fluid flow: PhD thesis, Stanford University, <http://sepwww.stanford.edu/public/docs/sep91/>.
- Muerdter, D. and D. Ratcliff, 2001, Understanding subsalt illumination through ray-trace modeling, Part 3: Salt ridges and furrows, and the impact of acquisition orientation: *The Leading Edge*, **20**, 803–816.
- Mulder, W. and R.-E. Plessix, 2004, Frequency-domain finite-frequency amplitude-preserving migration: *Geophysical Journal International*, **157**, 975–985.
- Plessix, R.-E. and W. Mulder, 2002, Amplitude-preserving finite-difference migration based on a least-squares formulation in the frequency domain: SEG Technical Program Expanded Abstracts, **21**, 1212–1215.
- Rickett, J., 2003, Illumination-based normalization for wave-equation depth migration: *Geophysics*, **68**, 1371–1379.
- Rickett, J. and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: *Geophysics*, **66**, 1015–1025.
- Ross, C. P. and M. S. Altan, 1997, Time-lapse seismic monitoring: Some shortcomings in nonuniform processing: *The Leading Edge*, **16**, 931–937.
- Sarkar, S., W. P. Gouveia, and D. H. Johnston, 2003, On the inversion of time-lapse seismic data: SEG Technical Program Expanded Abstracts, **22**, 1489–1492.
- Sava, P., 2006, Subsalt exploration and development: Imaging, interpretation, and drilling—what have we learned? 2006 SEG/EAGE summer research workshop: *The Leading Edge*, **25**, 1370–1376.

- Shin, C., S. Jang, and D.-J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: *Geophysical Prospecting*, **49**, 592–606.
- Symes, W. W., 2008, Approximate linearized inversion by optimal scaling of prestack depth migration: *Geophysics*, **73**, R23–R35.
- Tang, Y., 2008, Wave-equation hessian by phase encoding: SEP-134.
- Tarantola, A., 1987, *Inverse problem theory: Methods for data fitting and model parameter estimation*: Elsevier Science Publication Company, Inc.
- Valenciano, A., 2008, *Imaging by Wave-Equation Inversion*: PhD thesis, Stanford University.
- Valenciano, A. A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: *Geophysics*, **71**, A35–A38.
- Whitcombe, D. N., J. M. Marsh, P. J. Clifford, M. Dyce, C. J. S. McKenzie, S. Campbell, A. J. Hill, R. S. Parr, C. Pearse, T. A. Ricketts, C. P. Slater, and O. L. Barkved, 2004, The systematic application of 4D in BP's North-West Europe operations — 5 years on: *SEG Technical Program Expanded Abstracts*, **23**, 2251–2254.
- Zhang, M., M. D. Sacchi, and D. R. Schmitt, 2005, Simultaneous inversion of time-lapse seismic data: *CSEG Conference*, 228–231.
- Zou, Y., L. R. Bentley, L. R. Lines, and D. Coombe, 2006, Integration of seismic methods with reservoir simulation, Pikes Peak heavy-oil field, Saskatchewan: *The Leading Edge*, **25**, 764–781.

APPENDIX A

JOINT INVERSION FORMULATION FOR MULTIPLE SURVEYS

This appendix shows an extension of the joint inversion formulation to an arbitrary number of surveys. For three surveys, the data modeling procedure is given by

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 & \mathbf{0} \\ \mathbf{L}_2 & \mathbf{L}_2 & \mathbf{L}_2 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \Delta\mathbf{m}_1 \\ \Delta\mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \quad (\text{A-1})$$

where \mathbf{d}_0 , \mathbf{d}_1 and \mathbf{d}_2 are respectively datasets for the baseline, first and second monitor, \mathbf{m}_0 is the baseline reflectivity and the time-lapse reflectivities $\Delta\mathbf{m}_1$ and $\Delta\mathbf{m}_2$ are defined as

$$\begin{aligned} \Delta\mathbf{m}_1 &= \mathbf{m}_1 - \mathbf{m}_0, \\ \Delta\mathbf{m}_2 &= \mathbf{m}_2 - \mathbf{m}_1, \end{aligned} \quad (\text{A-2})$$

where \mathbf{m}_1 and \mathbf{m}_2 are respectively the monitor reflectivities at the times data \mathbf{d}_0 and \mathbf{d}_1 were acquired (with survey geometries defined by the linear \mathbf{L}_1 and \mathbf{L}_2). The least-squares solution to equation A-1 is given as:

$$\begin{bmatrix} \mathbf{L}_0^\dagger\mathbf{L}_0 + \mathbf{L}_1^\dagger\mathbf{L}_1 + \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_1^\dagger\mathbf{L}_1 + \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_2^\dagger\mathbf{L}_2 \\ \mathbf{L}_1^\dagger\mathbf{L}_1 + \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_1^\dagger\mathbf{L}_1 + \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_2^\dagger\mathbf{L}_2 \\ \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_2^\dagger\mathbf{L}_2 & \mathbf{L}_2^\dagger\mathbf{L}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{L}_0^\dagger & \mathbf{L}_1^\dagger & \mathbf{L}_2^\dagger \\ & \mathbf{L}_1^\dagger & \mathbf{L}_2^\dagger \\ & & \mathbf{L}_2^\dagger \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \quad (\text{A-3})$$

where the symbol \dagger denotes transposed complex conjugate. We rewrite equation A-3 as

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_2 & \mathbf{H}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_2 \end{bmatrix}, \quad (\text{A-4})$$

where $\tilde{\mathbf{m}}_i$ is the migrated image from the i_{th} survey, and \mathbf{H}_i is the corresponding Hessian matrix. Introducing a styling/regularization term that incorporates prior knowledge of the reservoir geometry and location as well as constraints on the inverted time-lapse images into equation A-4 we obtain

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_1 + \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_2 & \mathbf{H}_2 \\ \hline \epsilon_0\Lambda_0\mathbf{R}_0 & \mathbf{0} & \mathbf{0} \\ -\epsilon_1\Lambda_1\mathbf{R}_0 & \epsilon_1\Lambda_1\mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & -\epsilon_2\Lambda_2\mathbf{R}_0 & \epsilon_2\Lambda_2\mathbf{R}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 \\ \tilde{\mathbf{m}}_2 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (\text{A-5})$$

With four surveys, the time-lapse images can be obtained by solving the equation

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_3 \\ \mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_1 + \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_3 \\ \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_2 + \mathbf{H}_3 & \mathbf{H}_3 \\ \mathbf{H}_3 & \mathbf{H}_3 & \mathbf{H}_3 & \mathbf{H}_3 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \\ \Delta\hat{\mathbf{m}}_3 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 + \tilde{\mathbf{m}}_3 \\ \tilde{\mathbf{m}}_1 + \tilde{\mathbf{m}}_2 + \tilde{\mathbf{m}}_3 \\ \tilde{\mathbf{m}}_2 + \tilde{\mathbf{m}}_3 \\ \tilde{\mathbf{m}}_3 \end{bmatrix}, \quad (\text{A-6})$$

and for N surveys

$$\begin{bmatrix} \mathbf{H}_0 + \dots + \mathbf{H}_N & \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \mathbf{H}_2 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_N & \mathbf{H}_N & \mathbf{H}_N & \dots & \mathbf{H}_N \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta \hat{\mathbf{m}}_1 \\ \Delta \hat{\mathbf{m}}_2 \\ \vdots \\ \Delta \hat{\mathbf{m}}_N \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \dots + \tilde{\mathbf{m}}_N \\ \tilde{\mathbf{m}}_1 + \dots + \tilde{\mathbf{m}}_N \\ \tilde{\mathbf{m}}_2 + \dots + \tilde{\mathbf{m}}_N \\ \vdots \\ \mathbf{m}_N \end{bmatrix}. \quad (\text{A-7})$$

By solving equation A-7, it is possible to obtain the inverted baseline and time-lapse images from multiple surveys. We note that because different surveys would have some coincident shot or receiver locations computational cost of the Hessian for different surveys is reduced. Finally, introducing a regularization operator into equation A-7, the deconvolution operator becomes

$$\begin{bmatrix} \mathbf{H}_0 + \dots + \mathbf{H}_N & \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_1 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \mathbf{H}_2 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \mathbf{H}_2 + \dots + \mathbf{H}_N & \dots & \mathbf{H}_N \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_N & \mathbf{H}_N & \mathbf{H}_N & \dots & \mathbf{H}_N \\ \hline \epsilon_0 \Lambda_0 \mathbf{R}_0 & - & - & - & - \\ -\epsilon_1 \Lambda_1 \mathbf{R}_0 & \epsilon_1 \Lambda_1 \mathbf{R}_1 & - & - & - \\ - & -\epsilon_2 \Lambda_2 \mathbf{R}_1 & \epsilon_2 \Lambda_2 \mathbf{R}_2 & - & - \\ - & - & \vdots & \vdots & - \\ - & - & \vdots & \vdots & - \\ - & - & -\epsilon_{N-1} \Lambda_{N-1} \mathbf{R}_{N-2} & \epsilon_{N-1} \Lambda_{N-1} \mathbf{R}_{N-1} & - \\ - & - & - & -\epsilon_N \Lambda_N \mathbf{R}_{N-1} & \epsilon_N \Lambda_N \mathbf{R}_N \end{bmatrix}, \quad (\text{A-8})$$

where, \mathbf{R}_i is the spatial regularization for survey i , Λ_i is the temporal regularization between surveys, and ϵ_i determines the relative importance of model changes between surveys.

Automatic wave-equation migration velocity analysis

Biondo Biondi

ABSTRACT

I present a general framework to formalize a wide class of wave-equation velocity-estimation methods that search for the velocity function that optimally focuses the migrated image. Two well-known methods, the Differential Semblance Optimization (DSO) method and the Wave-Equation Migration Velocity Analysis (WEMVA), are particular instances of velocity-estimation methods that can be formalized in this framework. I define a WEMVA-like algorithm that exploits the potential of residual migration for velocity estimation but does not require the picking of residual-migration parameters. This result enables the derivation of a closed-form expression for the gradient of the objective function, making possible to use quasi-Newton methods to solve the optimization problem. I use a simple synthetic example to illustrate some of the characteristics of the method, and to evaluate its potential as a velocity estimation method.

INTRODUCTION

Tomographic velocity estimation based on wave-equation operators has the potential of improving the imaging in areas where wavefield-continuation migration is needed. Among the most promising methods to update velocity using wave-equation operators are the ones that formulate the velocity estimation problem as a maximization of the image focusing in the migrated domain. Two important examples of this approach are the Wave-Equation Migration Velocity Analysis (WEMVA) method (Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Sava, 2004) and the Differential Semblance Optimization (DSO) (Symes and Carazzone, 1991; Shen, 2004; Shen et al., 2005b). However, each iteration of wave-equation velocity updating is computationally expensive and convergence is often slow. Practical applications are still rare and small in scale (Sava, 2004; Shen et al., 2005a; Albertin et al., 2006). The goal of this paper is to present a general framework for formalizing wave-equation velocity estimation methods that comprise both the DSO and the WEMVA methods. I follow the idea presented by Sava and Symes (2002), but I further generalize and develop it to devise a velocity estimation method that should benefit from the relative advantages of both DSO and WEMVA methods and overcome some of their shortcomings.

The WEMVA method has the practical disadvantage over the DSO method that it includes a picking step of the residual-migration parameters ($\Delta\rho$), and thus it is less amenable to automatic implementations. Furthermore, as I will discuss in more detail in the next section, it does not lend itself to the application of efficient quasi-Newton optimization methods that may speed up convergence (Shen, 2004; Maharramov and Albertin, 2007). On the other hand, the WEMVA method has the fundamental advantage of harnessing the power of residual migration for the velocity estimation process. Residual migration has the potential of extracting velocity information from the image focusing, or lack thereof,

along the midpoint directions as well as the offset/angle directions. This can be useful, when the velocity resolution afforded by analyzing common image gathers is low because of poor angular coverage (e. g. subsalt). Sava et al. (2005) show a promising example of the application of residual migration to estimate velocity from the focusing of diffracted events.

Another, and more "technical", reason for relying on residual migration is that the wave-equation tomographic operators used to invert image perturbations (Biondi and Sava, 1999; Shen, 2004) are derived as the first-order Born linearization of wavefield-continuation migration. Therefore, the image perturbations (data residuals in inversion terminology) created by residual migration are more consistent with the wave-equation tomographic operator (modeling operator) than the image perturbations created by either applying residual moveout (RMO) or DSO. For example, in presence of a uniform velocity perturbation the wave-equation tomographic operator predicts a bulk shift of the migrated events at all reflection angles in migrated angle-domain common image gathers. This bulk shift of the whole events is inconsistent with the image perturbations inverted when applying the DSO method; these image perturbations are different from zero only for wide aperture angles where there is substantial differential shift between the image at adjacent angles.

The general framework that I introduce in this paper allows the formulation of a method to perform wave-equation MVA that does not include a picking step of the residual-migration parameters ($\Delta\rho$). The vector of residual-migration parameters $\Delta\rho$ that is needed by residual migration is estimated by computing the gradient direction of an auxiliary optimization problem formulated as a function of $\Delta\rho$. This explicit evaluation of $\Delta\rho$ also enables the derivations of a closed-form expression for the gradient of the primary objective function, which now can be minimized by efficient quasi-Newton algorithms. Because of the ability of the proposed method to avoid the picking step, I will call it Automatic WEMVA, or for short, AWEMVA.

The new general framework should also enable to easily substitute in the AWEMVA procedure other useful residual-focusing operators, such as a multi-parameter residual-migration operators, for the single-parameter residual migration I use in this paper. This improvement should help the convergence at the later stages of the estimation process, when the spatial variability of the residual velocity errors causes the migrated gathers to have complex moveout as a function of the aperture angle(s). Other possible research directions that are made possible by the general framework is the tailoring of the objective function to specific velocity problems, such as the ones related to shallow-velocity anomalies or velocity estimation in presence of shadow zones.

GENERAL FRAMEWORK FOR WAVE-EQUATION MVA

A general formulation of the velocity estimation problem, that comprises both the DSO inversion as well as the WEMVA method, is the minimization of the following objective function:

$$J_{\mathbf{s}} = \frac{1}{2} \|\mathbf{R} - \mathbf{F}(\mathbf{R}(\mathbf{s}))\|_2, \quad (1)$$

where \mathbf{R} is the prestack image obtained by wavefield continuation migration, \mathbf{s} is the slowness function, and \mathbf{F} is a differential residual-focusing operator. For example, we could define the differential residual-focusing operator using the DSO operator, \mathbf{D} , as:

$$\mathbf{F}(\mathbf{R}) = (\mathbf{I} - \mathbf{D}) \mathbf{R}, \quad (2)$$

where \mathbf{I} is the identity operator. The DSO operator can be either applied in the subsurface-offset domain or in the angle domain.

In the particular case of DSO, the objective function in equation 1 simplifies into the following well-known form (Shen, 2004):

$$J_s = \frac{1}{2} \|\mathbf{DR}(\mathbf{s})\|_2. \quad (3)$$

The more general idea behind the objective function in equation 1 is that the optimally-focused image is achieved when no more residual focusing is possible, and thus the application of the residual-focusing operator is equivalent to the application of an identity operator. In the particular example of DSO (equations 2 and 3), the result of applying DSO to an image is null when all the events in the image are either focused at zero subsurface offset (subsurface-offset domain image), or they are aligned in the angle domain (angle-domain image).

In this paper I am mostly concerned with the choice of a differential residual-focusing operator that takes advantage of the potential of residual migration to improve the focusing of the image. Sava and Biondi (2004a) introduced a differential residual-migration operator based on residual prestack Stolt migration (Sava, 2003). I am using a similar differential residual-migration operator (presented in the Appendix) that is not based on Stolt migration but on the derivation of residual moveout in the angle domain introduced by Biondi and Symes (2004).

The residual migration introduced in the Appendix is less accurate than the one introduced by (Sava, 2003) for large velocity errors because it is based on a stationary raypath linearization. However, the differential residual migration I introduce should be practically equivalent to the one introduced by Sava and Biondi (2004a) but have substantial efficiency advantages. It can be directly applied to the 3D prestack image transformed into the angle domain, and thus does not require a large transpose of the prestack image cube, which can be an expensive and cumbersome operation. Furthermore, the residual migration introduced here is more amenable to potentially useful generalizations that use several parameters to describe the residual moveout across aperture angles. Other residual-focusing operators could be used as well; for example, in one of the numerical test shown in the next section I use the residual-moveout operator introduced by Biondi and Symes (2004).

I use differential residual-focusing operator, and not a conventional "finite" residual-focusing operator because they are consistent with the first-order Born linearization of wavefield-continuation migration, as amply discussed by Sava and Biondi (2004a) and Sava (2004). Such differential residual-focusing operators can be generally defined as

$$\mathbf{F}(\mathbf{R}) = (\mathbf{I} + \mathbf{K}[\Delta\rho]) \mathbf{R}, \quad (4)$$

where the linear operator \mathbf{K} applies different phase rotations to the image for different reflection angles and geological dips. The operator is a function of a vector of residual-focusing parameters $\Delta\rho$ that is defined in the (z, \mathbf{x}) of the spatial coordinates depth, z , and horizontal location, \mathbf{x} .

When the input image is fixed the differential residual-focusing operators are also linear functions of the vector $\Delta\rho$. It is thus useful to introduce the following notation to represent

the linear operator \mathbf{M} that, for a fixed image \mathbf{R} , has the domain defined in the space of the vectors $\Delta\rho$ and the range defined in the image space,

$$\mathbf{M}[\mathbf{R}] = \frac{\partial \mathbf{K}[\Delta\rho] \mathbf{R}}{\partial \Delta\rho}. \quad (5)$$

One of the main obstacles to the application of WEMVA as defined by Sava and Biondi (2004a) and Sava (2004) is the estimation of the vector of residual-focusing parameters $\Delta\rho$ by an automatic, efficient, and robust procedure. I propose to estimate $\Delta\rho$ by (approximately) solving an auxiliary optimization problem defined as the minimization of the following objective function

$$\begin{aligned} J_{\Delta\rho} &= \frac{1}{2}\epsilon_{\mathbf{D}} \|\mathbf{DF}(\Delta\rho, \mathbf{R}(\mathbf{s}))\|_2 - \frac{1}{2}\epsilon_{\mathbf{S}} \|\mathbf{SF}(\Delta\rho, \mathbf{R}(\mathbf{s}))\|_2 \\ &= \frac{1}{2}\epsilon_{\mathbf{D}} \|\mathbf{DR} + \mathbf{DM}[\mathbf{R}] \Delta\rho\|_2 - \frac{1}{2}\epsilon_{\mathbf{S}} \|\mathbf{DR} + \mathbf{SM}[\mathbf{R}] \Delta\rho\|_2, \end{aligned} \quad (6)$$

where \mathbf{S} is a "stacking" operator that actually stacks over reflection angles when the prestack image is defined in the angle domain, or extracts the zero-offset cube when the prestack image is defined in the subsurface-offset domain. The weights $\epsilon_{\mathbf{D}}$ and $\epsilon_{\mathbf{S}}$ provide a trade-off between a minimization of the differential-semblance criterion, which has appealing properties for global convergence, and a maximization of the stacking power, which has attractive high-resolution characteristics close to the solution.

It would be relatively straightforward and inexpensive to find the minimum of the objective function in equation 6 by an iterative optimization scheme. However, to provide a closed-form representation of the gradient of the main objective function (equation 1), I will use, as an estimate of the optimal $\Delta\rho$, the gradient of the objective function in equation 6, evaluated at $\Delta\rho = 0$, and smoothed by a simple smoothing operator \mathbf{G} ; that is:

$$\widehat{\Delta\rho} = \mathbf{G}\nabla J_{\Delta\rho} = \mathbf{GM}'[\mathbf{R}] (-\epsilon_{\mathbf{D}}\mathbf{D}'\mathbf{D} + \epsilon_{\mathbf{S}}\mathbf{S}'\mathbf{S}) \mathbf{R} = \mathbf{GM}'[\mathbf{R}] \mathbf{CR}, \quad (7)$$

where, to simplify the expressions that describe the development that follows, I introduce the following identity

$$\mathbf{C} = -\epsilon_{\mathbf{D}}\mathbf{D}'\mathbf{D} + \epsilon_{\mathbf{S}}\mathbf{S}'\mathbf{S}. \quad (8)$$

WEMVA approximation of the gradient

To solve the optimization problem formalized in equation 1, we need to compute the gradient of the objective function with respect to the slowness model \mathbf{s} . According to the "classical" WEMVA solution (Sava and Biondi, 2004a), the gradient is evaluated by fixing both the focusing-parameters vector $\Delta\rho$ and the image to which the residual focusing operator is applied; that is, by ignoring their dependence on \mathbf{s} and assuming that the image is the background image \mathbf{R}_0 obtained by migrating the data with the current best-estimate of the slowness function, and that the vector $\Delta\rho$ is either picked or evaluated by equation 7.

With these assumptions, the residual-focusing operator is

$$\mathbf{F}_0 = \mathbf{F}(\mathbf{R}_0) = \mathbf{R}_0 + \mathbf{K}[\widehat{\Delta\rho}] \mathbf{R}_0, \quad (9)$$

and the objective function can be written as:

$$\begin{aligned} J_s &= \frac{1}{2} \left\| \mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right) \right\|_2 \\ &= \frac{1}{2} \left\langle \mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right), \mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right) \right\rangle, \end{aligned} \quad (10)$$

$$(11)$$

where with the notation $\langle \mathbf{x}, \mathbf{y} \rangle$ I indicate the inner product of the vectors \mathbf{x} and \mathbf{y} .

The perturbation δJ_s in the objective function caused by a perturbation in the slowness model $\delta \mathbf{s}$ is:

$$\delta J_s = \left\langle \mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right), \delta \mathbf{R} \right\rangle = \left\langle \mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right), \mathbf{L} \delta \mathbf{s} \right\rangle, \quad (12)$$

where \mathbf{L} is the wave-equation tomographic operator relating perturbations in slowness to perturbations in the image ($\delta \mathbf{R} = \mathbf{L} \delta \mathbf{s}$) that was introduced by Biondi and Sava (1999) for source-receiver migration, and by Shen (2004) for shot-profile migration.

The gradient of the objective function is therefore

$$\nabla J_s = \mathbf{L}' \left(\mathbf{R} - \left(\mathbf{R}_0 + \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right) \right). \quad (13)$$

When the gradient is evaluated for $\mathbf{R} = \mathbf{R}_0$ it becomes

$$\nabla J_s|_{\mathbf{R}=\mathbf{R}_0} = -\mathbf{L}' \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0, \quad (14)$$

as presented by Sava (2004).

Another way of reaching the same result is to assume linear perturbations in the image and substitute $\mathbf{R} = \mathbf{R}_0 + \mathbf{L} \delta \mathbf{s}$ in equation 10. I can then derive the quadratic objective function

$$J_s = \frac{1}{2} \left\| \mathbf{L} \delta \mathbf{s} - \mathbf{K} \left[\widehat{\Delta\rho} \right] \mathbf{R}_0 \right\|_2, \quad (15)$$

that has the gradient expressed in equation 14.

Complete gradient

As discussed in the previous section, the gradient provided by the WEMVA method is an approximation of the full gradient of the objective function with respect to slowness. It is based on the assumption that both inputs to the residual focusing operator are independent from the slowness. This approximation may be inconsequential if the velocity estimation problem is solved by nesting several linearized iterations within an outer loop of non-linear iterations. Following this approach, a background velocity model is defined at each non-linear iteration, and the quadratic objective function in equation 15 is optimized with several steps of a conjugate gradient algorithm.

Shen (2004) and Maharramov and Albertin (2007) report success with a different optimization strategy when minimizing the DSO objective function in equation 3. He used a quasi-Newton optimization method (limited memory BFGS) that builds up information on the Hessian with iterations. Quasi-Newton methods can be more efficient when solving an

ill-conditioned non-quadratic optimization problem; however, they are sensitive to approximations in the gradient computation. It is thus useful to derive a procedure to compute the gradient that overcomes the approximations discussed above. On the other hand, one disadvantage of employing a quasi-Newton method is that at each iteration we need to recompute the background wavefields that are needed to evaluate \mathbf{L} and its adjoint. This requirement of updating the background wavefields more than doubles the computational cost of the quasi-Newton iterations over the cost of the linearized iterations needed to solve the quadratic objective function defined in equation 15.

If the residual migration operator is applied to the current image \mathbf{R} , the objective function in equation 1 becomes

$$J_s = \frac{1}{2} \|\mathbf{K}[\Delta\rho] \mathbf{R}\|_2 = \frac{1}{2} \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -\mathbf{K}[\Delta\rho] \mathbf{R} \rangle. \quad (16)$$

If the dependence of both \mathbf{R} and $\Delta\rho$ from the slowness function are taken into account, the perturbations in J_s caused by image perturbations $\delta\mathbf{R}$ are formally expressed as

$$\delta J_s = \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -\mathbf{K}[\Delta\rho] \delta\mathbf{R} - \mathbf{K}[\Delta\rho + \delta\Delta\rho(\delta\mathbf{R})] \mathbf{R} \rangle \quad (17)$$

If the estimate $\widehat{\Delta\rho}$ for the optimal perturbation in the focusing parameters vectors expressed in equation 7 is used, then equation 17 becomes

$$\begin{aligned} \delta J_s &= \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -\mathbf{K}[\Delta\rho] \delta\mathbf{R} - \mathbf{K}[\Delta\rho + \mathbf{GM}'[\mathbf{R}]\mathbf{C}\delta\mathbf{R}] \mathbf{R} \rangle \\ &= \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -\mathbf{K}[\Delta\rho] \delta\mathbf{R} - \mathbf{M}[\mathbf{R}]\mathbf{GM}'[\mathbf{R}]\mathbf{C}\delta\mathbf{R} \rangle \\ &= \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -(\mathbf{K}[\Delta\rho] + \mathbf{M}[\mathbf{R}]\mathbf{GM}'[\mathbf{R}]\mathbf{C}) \delta\mathbf{R} \rangle, \end{aligned} \quad (18)$$

where \mathbf{M} is defined in equation 5.

The objective-function perturbations caused by slowness perturbations are written as:

$$\delta J_s = \langle -\mathbf{K}[\Delta\rho] \mathbf{R}, -(\mathbf{K}[\Delta\rho] + \mathbf{M}[\mathbf{R}]\mathbf{GM}'[\mathbf{R}]\mathbf{C}) \mathbf{L}\delta\mathbf{s} \rangle \quad (19)$$

and the gradient becomes

$$\nabla J_s = \mathbf{L}'(\mathbf{K}'[\Delta\rho] + \mathbf{C}'\mathbf{M}[\mathbf{R}]\mathbf{G}'\mathbf{M}'[\mathbf{R}]) \mathbf{K}[\Delta\rho] \mathbf{R}. \quad (20)$$

that can be rewritten as the sum of two terms; that is,

$$\nabla J_s = \mathbf{L}'\mathbf{K}'[\Delta\rho] \mathbf{K}[\Delta\rho] \mathbf{R} + \mathbf{L}'\mathbf{C}'\mathbf{M}[\mathbf{R}]\mathbf{G}'\mathbf{M}'[\mathbf{R}]\mathbf{K}[\Delta\rho] \mathbf{R}. \quad (21)$$

The first term in the gradient in 21 is related to the image perturbations that are directly caused by perturbations in slowness. The second term is related to the image perturbations caused indirectly through perturbations in the optimal estimate of residual-focusing parameters $\widehat{\Delta\rho}$. When the residual focusing operator is not defined in parametric form but it is fixed, as for the important case of DSO, the second term vanishes, and we get the well-known expression for the gradient of the objective function in equation 3 (Shen, 2004):

$$\nabla J_s = \mathbf{L}'\mathbf{D}'\mathbf{DR}. \quad (22)$$

Notice that the application of the adjoint operator of the differential residual-focusing operator in both the first term of the gradient in equation 21 and in the DSO gradient in equation 22 removes the phase rotation from the image that is back-projected by the wave-equation tomographic operator. Intuitively, this lack of phase information in the image perturbations may slow the convergence towards a better focused image. The numerical tests I show in the next section are too preliminary to shed a light on the validity of this intuition.

The second term of the gradient in equation 21 maintains the phase rotation and may help the convergence. However, in conducting the numerical tests shown in the next section I have not included this term. The balance between the two terms in the gradient requires a careful implementation of the relative-amplitude response of the linear operators included in equation 21, which is beyond the limitations of my current implementation of these operators.

NUMERICAL EXAMPLES

I conducted some preliminary tests of the theory presented in the previous section. These tests cannot be considered conclusive because of the immature state of the software I am using. In particular, I have not yet included a regularization term in the objective functions that I optimize. Consequently, the iterative estimation did not converge to meaningful results and below I show the results after only three iterations. As explained at the end of the previous section, I also neglected the second term in the expression of the full-gradient in equation 21. Finally, a quasi-Newton optimization scheme should probably be employed to fully benefit from the full-gradient expression in equation 21, whereas I applied a simple conjugate-gradient optimization.

The test dataset is a simple synthetic with two flat reflectors in a constant velocity background. The migration velocity is the correct one (2 km/s) above the shallow reflector and 5% too low between the shallow reflector and the deep reflector. Figure 1 shows the prestack migrated cube: the whole cube in panel a), and a zoom around the deep reflector in panel b).

I computed the $\Delta\rho$ using equation 7 with $\epsilon_S = 0$; that is, I used only the DSO criterion to define a well-focused image. Figure 2 shows the image perturbations computed by applying differential residual-migration operator to the image cube shown in Figure 1: (a) using the gradient expression in equation 14, (b) using the first term of the gradient expression in equation 21. The figure shows the image perturbations before the application of the back-projection operator \mathbf{L} that is at the front of the operator chain in both equation 14 and equation 21. The phases of the two image-perturbation cubes are different. The phase of the cube shown in Figure 2b is similar to the phase of the original migrated cube shown in Figure 1b. In contrast, the image-perturbation cube shown in Figure 2a has a 90 degrees upward phase rotation. This phase rotation may speed up the inversion, as discussed at the end of the previous section.

Figure 3 shows the comparison of the image perturbations computed starting from the image cube shown in Figure 1: and applying: (a) the DSO gradient in equation 22, (b) the differential residual-moveout operator in the first term of the gradient expression in equation 21. Because the velocity error was constant in the horizontal direction, the two

image perturbations are similar. They have similar phase and no image perturbations around normal incidence. The inversion of these two image-perturbations cube can thus be expected to produce similar results.

The next three figures show the result of the iterative inversion of the image perturbations shown in Figure 2 and Figure 3a. Figure 4 shows the normalized slowness perturbations computed by the differential residual-migration operator and the gradient expression in equation 14: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 2a: (b) estimation result after three iterations. Similarly, Figure 5 shows the normalized slowness perturbations computed by the differential residual-migration operator and the first term of the gradient expression in equation 21: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 2b: (b) estimation result after three iterations. Comparing the slowness perturbations shown in Figure 4b with the slowness perturbations shown in Figure 5b, one might conclude that the use of the approximate WEMVA-like gradient might be more stable. However, I do not think that general conclusion can be drawn from this preliminary results.

Finally, Figure 6 shows the normalized slowness perturbations computed by the DSO method: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 3a: (b) estimation result after three iterations. Small differences can be observed between the slowness perturbations after the first and the third iteration. The first three iterations of the conjugate-gradient optimization are mostly spent dealing with the artifacts at the edges of the reflectors, and the erroneous velocity perturbations placed above the shallow reflector by the simple back-projection are not corrected by the following three iterations. This result is also too preliminary to support general conclusions.

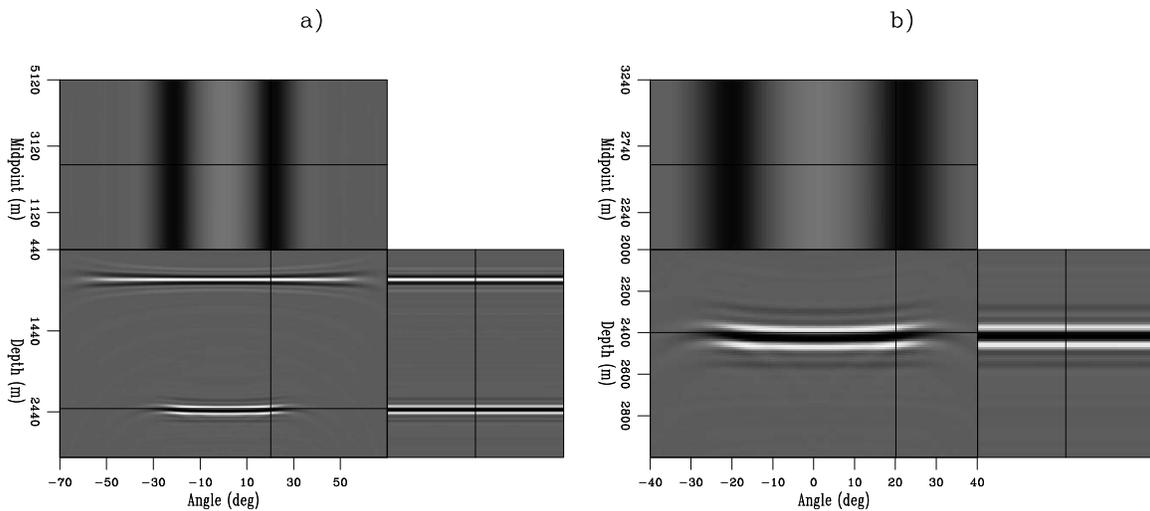


Figure 1: Angle-domain prestack image used to test the velocity-updating process: (a) whole migrated cube, (b) zoom around the deep reflector. The migration velocity model is 5% too low between the shallow reflector and the deep reflector. [CR] biondo1/. ANG-both-overn

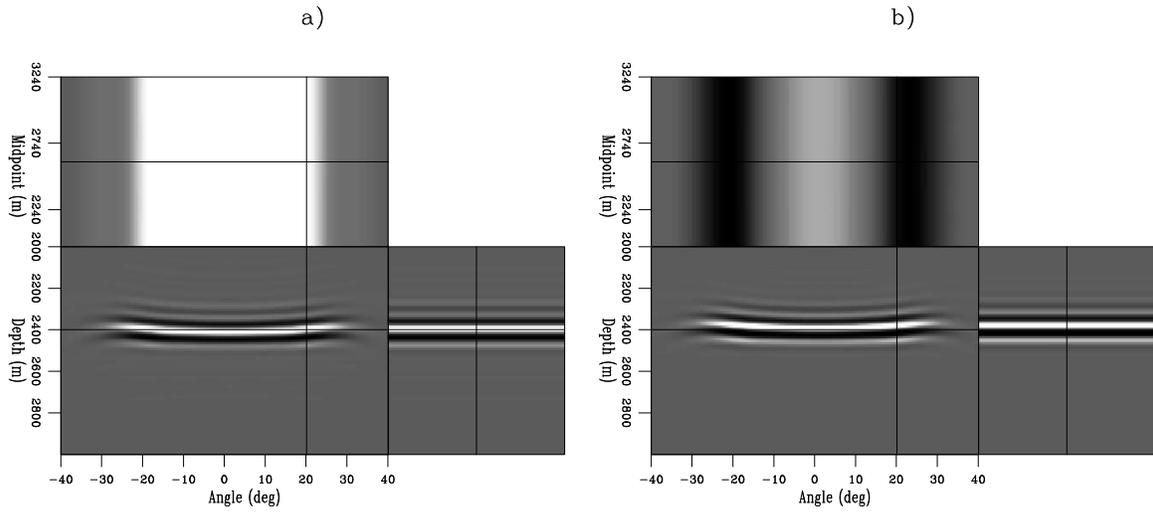


Figure 2: Image perturbations computed by applying differential residual-migration operator to the image cube shown in Figure 1: (a) using the gradient expression in equation 14, (b) using the first term of the gradient expression in equation 21. Notice the phase rotation between the two panels. [CR] `biondo1/. Dmig-both-rmig-overn`

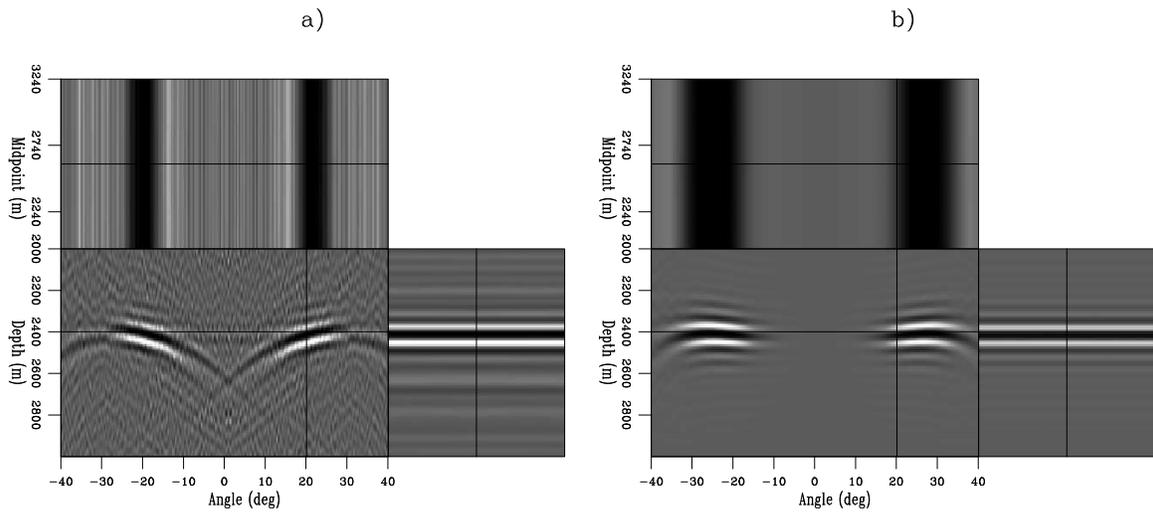


Figure 3: Comparison of the image perturbations computed starting from the image cube shown in Figure 1 and applying: (a) DSO gradient in equation 22, (b) the differential residual-moveout operator in the first term of the gradient expression in equation 21. Notice the similarities. [CR] `biondo1/. Dmig-Dso-overn`

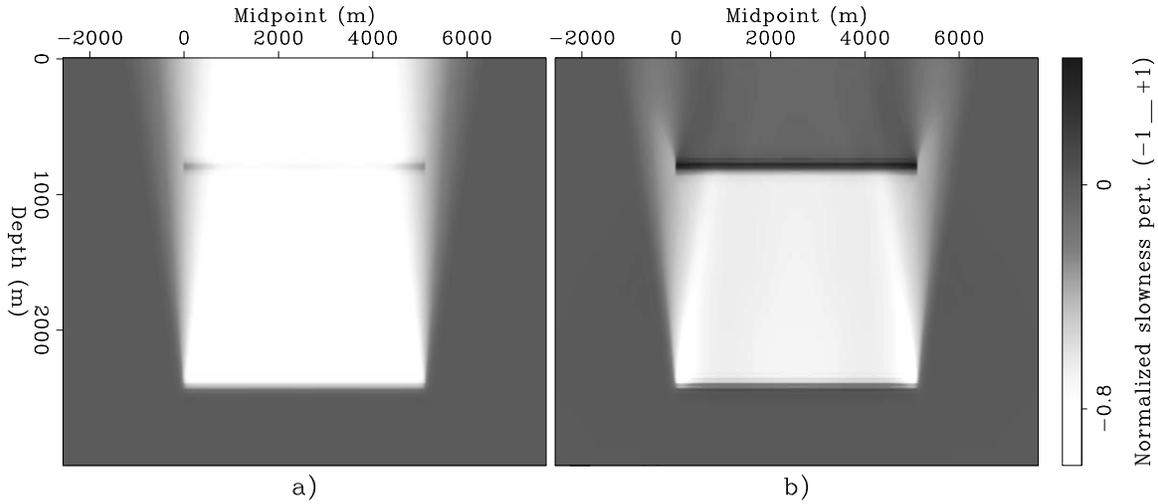


Figure 4: Normalized slowness perturbations computed by the differential residual-migration operator and the gradient expression in equation 14: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 2a, (b) estimation result after three iterations. [CR] `biondo1/. Dslow-rmig-both-overn`

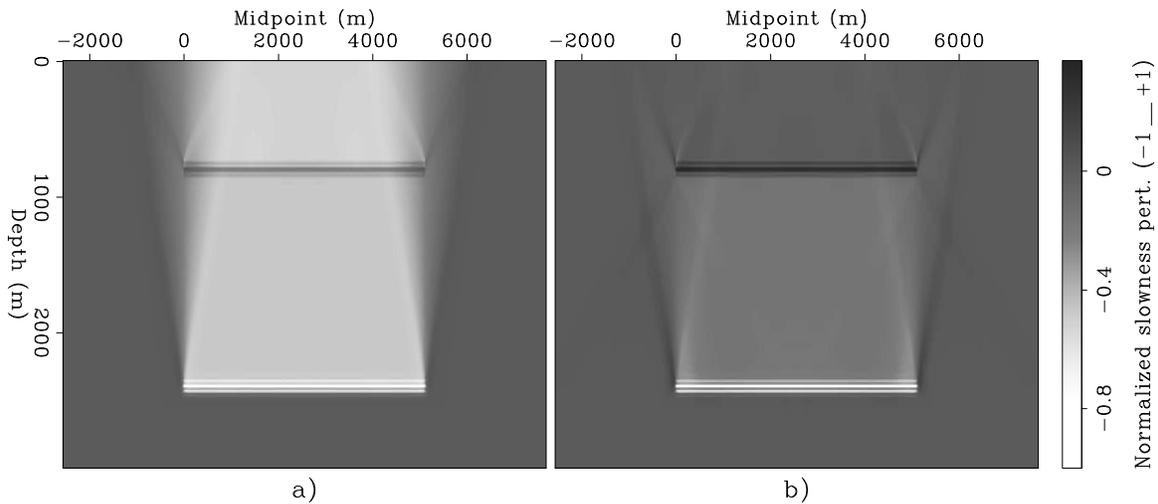


Figure 5: Normalized slowness perturbations computed by the differential residual-migration operator and the first term of the gradient expression in equation 21: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 2b: (b) estimation result after three iterations. [CR] `biondo1/. Dslow-rmig-adj-both-overn`

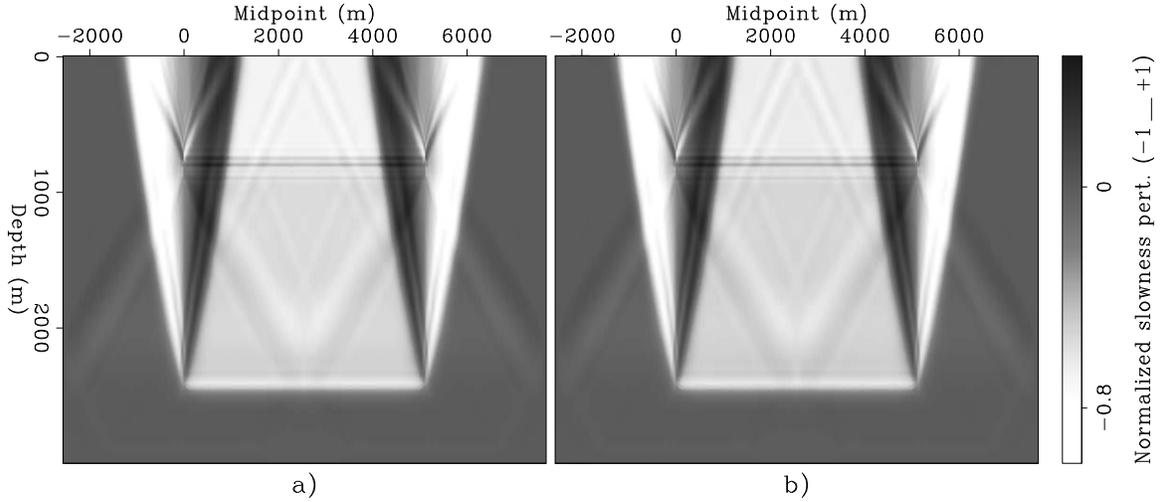


Figure 6: Normalized slowness perturbations computed by the DSO method: (a) gradient at first iteration; that is, back-projection of image perturbations shown in Figure 3a: (b) estimation result after three iterations. [CR] `biondo1/. Dslow-dso-both-overn`

CONCLUSIONS

I presented a general framework for wave-equation velocity estimation based on the criterion of maximizing the focusing of prestack images after migration, or more precisely, of minimizing the amount of unfocused energy. This framework enabled the definition of an automatic velocity estimation method that benefits from residual migration, but does not require picking residual-migration parameters.

The numerical examples illustrate the basic elements and the concepts underlying the proposed method. They confirm that the method has the potential of updating the velocity model in the right direction; they are, however, too preliminary to demonstrate its usefulness in practical applications.

APPENDIX A

DIFFERENTIAL RESIDUAL MIGRATION

In this appendix I define the differential residual-migration and residual-moveout operators that are used in the paper, starting from the expression of the residual migration and residual moveout for angle-domain common image gathers introduced in 2D by Biondi and Symes (2004) and in 3D by Biondi and Tisserant (2004). For simplicity, I will derive the operator in 2D, but the 3D extension should be straightforward.

Combining equations D-3 and D-6 in Appendix D of Biondi and Symes (2004), the residual-migration shift $\Delta \mathbf{n}$ along the normal to the reflector \mathbf{n} is given by

$$\Delta \mathbf{n} = -\Delta \rho \frac{\cos \alpha}{(\cos^2 \alpha - \sin^2 \gamma)} z_0 \mathbf{n} = -\Delta \rho S_{\text{Rmig}}(\alpha, \gamma) z_0 \mathbf{n}, \quad (\text{A-1})$$

where α is the geological dip of the reflector, γ is the aperture angle, and z_0 is the depth of the reflector in the migrated domain with the incorrect velocity. The parameter $\Delta\rho$ depends on the velocity error and is defined as

$$\Delta\rho = \frac{S - S_\rho}{S}, \quad (\text{A-2})$$

where S is the correct migration slowness, and S_ρ is the actual, and incorrect, migration slowness.

If we assume that $\Delta\rho$ is a constant of the spatial coordinates, and ignore the dependency of $\Delta\mathbf{n}$ from the reflector depth z_0 by fixing z_0 at an average value \bar{z}_0 , the residual migration shift can be applied efficiently in the wavenumber domain, by taking advantage of the well-known relationship between geological dip and the wavenumber:

$$\tan \alpha = \frac{k_x}{k_z}, \quad (\text{A-3})$$

where k_x is the inline midpoint wavenumber and k_z is the depth wavenumber. In the wavenumber domain the residual migration is a simple phase shift linear function of $\Delta\rho$; that is

$$\text{Rmig}(\gamma, k_z, k_x) = \exp[i\Delta\rho S_{\text{Rmig}}(k_z, k_x, \gamma)(k_x \sin \alpha + k_z \cos \alpha)]. \quad (\text{A-4})$$

To derive the differential residual-migration operator we simply take the derivative of Rmig with respect to $\Delta\rho$. Following the definitions given in the main body we have

$$\mathbf{K}[\Delta\rho] = i\Delta\rho S_{\text{Rmig}}(k_z, k_x, \gamma)(k_x \sin \alpha + k_z \cos \alpha), \quad (\text{A-5})$$

and

$$\mathbf{M} = iS_{\text{Rmig}}(k_z, k_x, \gamma)(k_x \sin \alpha + k_z \cos \alpha) \mathbf{R}(k_z, k_x, \gamma). \quad (\text{A-6})$$

So far I assumed that $\Delta\rho$ is constant, but in reality we would like to have the parameter $\Delta\rho$ varying, may be smoothly, as a function of the spatial coordinates z and x . That is a potential implementation challenge since I defined the differential residual-migration operator in the wavenumber domain.

For the numerical examples shown in the paper I used the following approximations when applying \mathbf{K} and \mathbf{M} to a spatially varying $\Delta\rho(z, x)$:

$$\begin{aligned} \mathbf{K}[\Delta\rho] \mathbf{R} &= \mathbf{M}[\mathbf{R}] \Delta\rho = \\ &\Delta\rho(z, x) \mathbf{FFT}^{-1} \{iS_{\text{Rmig}}(k_z, k_x, \gamma)(k_x \sin \alpha + k_z \cos \alpha) \mathbf{FFT}[\mathbf{R}(z, x, \gamma)]\}, \end{aligned} \quad (\text{A-7})$$

where \mathbf{FFT} and \mathbf{FFT}^{-1} represent forward FFT and inverse FFT, respectively. The error related to this approximation is probably minor when $\Delta\rho(z, x)$ is a smooth function of the spatial coordinates, since the differential residual-migration operator is a simple differentiator (multiplication by ik_x and ik_z), and thus very compact in the space domain.

In some cases it might be useful to use a differential residual-moveout operator instead of a differential residual-migration operator. A differential residual-moveout operator can

be defined following the same procedure outlined above for differential residual migration, but with the shift function S_{Rrmo} derived from equation D-7 in Appendix D of Biondi and Symes (2004) as

$$S_{\text{Rrmo}}(\alpha, \gamma) = -\frac{\Delta\rho}{\cos\alpha} \frac{\sin^2\gamma}{(\cos^2\alpha - \sin^2\gamma)}. \quad (\text{A-8})$$

Figure 3 shows an example of the application of equation A-8.

REFERENCES

- Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3345–3349.
- 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.
- Biondi, B. and T. Tisserant, 2004, 3-D angle-domain common-image gathers for migration velocity analysis: *Geophysical Prospecting*: accepted for publication, **52**, 575–591.
- Maharramov, M. and U. Albertin, 2007, Localized image-difference wave-equation tomography: SEG Technical Program Expanded Abstracts, **27**, 3009–3013.
- Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.
- 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 W. W. Symes, 2002, A generalization of wave-equation migration velocity analysis: SEP-Report, **112**, 27–36.
- Sava, P. C., 2003, Prestack residual migration in frequency domain: *Geophysics*, **68**, 634–640.
- Sava, P. C., B. Biondi, and J. Etgen, 2005, Wave-equation migration velocity analysis by focusing diffractions and reflections: *Geophysics*, **70**, U19–U27.
- Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Shen, P., E. Menyoli, and H. Calandra, 2005a, Automatic sub-salt velocity analysis: an integrated strategy of ray-based tomography and wave-equation migration velocity inversion: SEG, Expanded Abstracts, **24**, 1978–1980.
- Shen, P., W. W. Symes, S. Morton, and H. Calandra, 2005b, Differential semblance velocity analysis via shot profile migration: SEG, Expanded Abstracts, **24**, 2249–2252.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.

Prestack exploding reflector modeling: The crosstalk problem

Claudio Guerra and Biondo Biondi

ABSTRACT

The recently introduced prestack exploding reflector modeling aims to model a small dataset comprised by areal shots, while keeping the correct kinematics to be used in iterations of migration velocity analysis. To achieve this goal, the modeled areal data must be combined into sets. This procedure generates data which is subjected to crosstalk during migration. Here, using a simple constant velocity model, we describe two different origins for the crosstalks and we show that by applying the concept of random phase-encoding during the modeling of the areal shots we achieve an almost complete elimination of the crosstalks.

INTRODUCTION

Oil exploration is being conducted in increasingly more geologically complex areas. The imaging challenges, along with the continuous growth of computer power, have augmented the relative importance of depth migration – in particular, wave-equation migration. A variety of wave-equation migration algorithms have been developed to balance the competing constraints of cost and final image quality. Velocity model definition for depth migration, however, strongly depends on ray-based methods because of their flexibility and low cost when compared with wave-equation migration velocity analysis. In areas where ray theory fails and the wavefields become too complex, this dependence might result in inaccurate velocity models that yield poor quality or unreliable images even when using state-of-the-art wave-equation algorithms. Thus, in these areas it is desirable to use wave-equation migration to generate image gathers in order to evaluate the accuracy of the velocity model.

Every iteration of migration velocity analysis for depth migration includes: 1) migrating with the current velocity model; 2) measuring properties that diagnoses the current velocity model accuracy; and, 3) updating the velocity model. Conventionally, Kirchhoff migration is used in step 1 and steps 2 and 3 are accomplished by measuring the departures of reflectors from the horizontal in the common image gathers and by back-projecting them along rays onto the model space, respectively. In areas where ray theory fails, multi-pathing and shadow zones occurs, causing problems in updating the velocity model using ray methods.

Wave-equation methods for both migration and velocity update overcome the limitations of ray-based methods (Sava and Biondi, 2004a,b), but at the expenses of a high cost. The practical implementation of wave-equation methods, therefore, still depends on an efficient migration/modeling scheme.

Based on the linearity of wavefield extrapolators, different strategies have been formulated to reduce the computational cost of wave-equation migration. The synthesis of plane

waves was first introduced by Schultz and Claerbout (1978). Rietveld et al. (1992) used pre-defined wavefields of arbitrary shape to synthesize areal shots after propagation back to the surface. Rietveld and Berkhout (1992) proposed a depth migration method using the concept of controlled illumination, which can be considered as a generalization of Schultz and Claerbout's method. Migration of phase-encoded shot-records (Romero et al., 2000) is a technique that provides the ability to migrate combined shots with minimal cross-talk. The combination of phase-encoded and plane-wave areal shots can be used to drastically decrease the computational effort (Sun et al., 2002). Liu et al. (2006) provide a general framework to evaluate plane-wave composition in prestack source plane-wave migration and show the equivalence of shot profile and plane-wave migrations.

Recently, Biondi (2006, 2007) introduced the concept of the prestack exploding reflector modeling. This method synthesizes source and receiver wavefields along the entire survey at the surface, in the form of areal data, starting from a prestack migrated image cube represented by subsurface-offset domain common-image gathers (SODCIGs). For the case of migration velocity analysis, the aim is to generate a smaller dataset than the one used in the initial migration while maintaining the kinematics necessary to perform migration velocity analysis.

If the migration velocity is correct and assuming sufficiently good illumination, the reflectors are focused around the zero subsurface offset and the method reduces to the conventional exploding reflector modeling. However, when the migration velocity is inaccurate and in the presence of geologic dip, a pre-processing step is needed before modeling. Biondi (2007) describes how to generate dip-independent image gathers to be used as initial condition for the upward propagation of source and receiver wavefields. This is achieved by rotating the SODCIGs according to the geological dip in the wavenumber domain.

Conceptually, in the prestack exploding reflector modeling, the synthesized areal data is computed by upward propagating source and receiver wavefields using as initial condition one SODCIG. Depending on the relationship between shot and SODCIG sampling, this procedure actually can generate as many individual experiments as the existing in the original data. However, we take advantage of the linearity of the wave propagation to combine several experiments into a set composite records, therefore decreasing the amount of data to migrate.

Combining several experiments, though, gives rise to crosstalk during imaging. Biondi (2006) uses a decorrelation distance between SODCIGs sufficiently large to prevent crosstalk. Using this criterion, however, limits the maximum data reduction to the ratio between the number of SODCIGs and the decorrelation distance. Furthermore, he mentions that an additional saving can be achieved by using concepts similar to phase encoding (Romero et al., 2000) during the modeling step. This allows us to combine more shots into one set. From now on, upward propagated areal shots combined into one single areal shot will be called a set.

Biondi (2006, 2007) uses the two-way wave equation to propagate the wavefields. Here, we use the one-way wave equation to achieve additional computational saving. Furthermore, under the one-way framework, we may introduce a phase-encoding-like scheme during the modeling to reduce cross-talk when applying the imaging condition.

PRESTACK EXPLODING REFLECTOR MODELING

The modeling of 2D source and receiver wavefields at the surface, using the one-way wave equation and starting from a prestack image at a selected position, x_ξ , can be described by:

$$\begin{aligned} S(x, \omega) &= G(z_\xi, x_\xi - h_\xi; x, z = 0, \omega) * I_s(z_\xi, x_\xi, h_\xi), \\ R(x, \omega) &= G(z_\xi, x_\xi + h_\xi; x, z = 0, \omega) * I_r(z_\xi, x_\xi, h_\xi), \end{aligned} \quad (1)$$

where $S(x, \omega)$ is the source wavefield at $z = 0$; $R(x, \omega)$ is the receiver wavefield at $z = 0$; $I_s(z_\xi, x_\xi, h_\xi)$ and $I_r(z_\xi, x_\xi, h_\xi)$ are the prestack images used as initial condition for the source and receiver wavefield extrapolation, respectively; $G(z_\xi, x_\xi \pm h_\xi; x, z = 0, \omega)$ represents the one-way operator which extrapolates the wavefields from the subsurface to the surface; h_ξ is the subsurface offset; z_ξ is depth; ω is the temporal frequency and x is the spatial coordinate in the data space coinciding with x_ξ . Notice that for a perfectly focused image around $h_\xi = 0$, equation 1 reduces to the conventional zero-offset exploding reflector modeling. The prestack images used as initial condition for the source and receiver wavefield extrapolation are supposed to be dip-independent gathers computed by changing the dip along the offset direction according to the apparent geological dip (Biondi, 2007).

The number of individual experiments described in equation 1 equals the number of shots used in the initial migration if the CMP spacing is equal to the shot spacing. Therefore, to decrease the amount of input data to migration, areal shots need to be combined into sets after the upward propagation. We achieve this by regularly selecting individual experiments and adding them up into their set, after being upward propagated, according to:

$$\begin{aligned} \widetilde{S}_n(x, \omega) &= \sum_{n=1}^k \sum_{i=n,k,N} S_i(x, \omega), \\ \widetilde{R}_n(x, \omega) &= \sum_{n=1}^k \sum_{i=n,k,N} R_i(x, \omega), \end{aligned} \quad (2)$$

where $\widetilde{S}_n(x, \omega)$ and $\widetilde{R}_n(x, \omega)$ contains k -sets of summed areal sources and areal receivers, respectively; N is the number of SODCIGs. Every other k areal data are selected to compose one set. Pairs of $\widetilde{S}_k(x, \omega)$ and $\widetilde{R}_k(x, \omega)$ are to be used as the areal source function and the areal receiver wavefield, respectively, in areal shot migration.

As we show in the next section, unrelated events pertaining to the same pair of sets of areal source and areal receiver correlate together during imaging by cross-correlation, generating crosstalk. The crosstalk degrades the image, obscuring the kinematics we are interested in to perform migration velocity analysis.

CROSTALK GENERATION AND ATTENUATION

Wavefield propagation is a linear process. This allows us to linearly combine wavefields before the propagation effects are removed by migration. Several migration methods explore this linearity to decrease the amount of data to migration (Rietveld and Berkhout, 1992; Romero et al., 2000; Liu et al., 2006). However, while applying the imaging condition, energy from unrelated source and receiver wavefields crosscorrelate generating crosstalk. For the present case, let us consider the migration of a set, m , of added areal shots containing a

single pair of individual experiments. The prestack image is formed by crosscorrelating the source and receiver wavefields

$$\widetilde{I}_m(z_\xi, x_\xi, h_\xi) = \sum_{\omega} \widetilde{S}_m^*(z_\xi, x_\xi - h_\xi, \omega) \widetilde{R}_m(z_\xi, x_\xi + h_\xi, \omega), \quad (3)$$

where $*$ represents complex conjugation. If $\widetilde{S}_m(x, \omega)$ and $\widetilde{R}_m(x, \omega)$ in equation 2 are comprised by two summed areal shots, the image $\widetilde{I}_m(z_\xi, x_\xi, h_\xi)$ will be given by:

$$\begin{aligned} \widetilde{I}_m(z_\xi, x_\xi, h_\xi) = & I_1(z_\xi, x_\xi, h_\xi) + I_2(z_\xi, x_\xi, h_\xi) + \\ & \sum_{\omega} S_1^*(z_\xi, x_\xi - h_\xi, \omega) R_2(z_\xi, x_\xi + h_\xi, \omega) + \\ & \sum_{\omega} S_2^*(z_\xi, x_\xi - h_\xi, \omega) R_1(z_\xi, x_\xi + h_\xi, \omega). \end{aligned} \quad (4)$$

In equation 4, the last two summation terms represent the crosstalk. Romero et al. (2000) attenuate these migration artifacts by multiplying the source and the receiver wavefields by a function of frequency and space. By doing so, the crosstalks are dispersed (random phase encoding) throughout the image after stacking over frequency, during imaging, or shifted outside of the image space (linear phase encoding). Here, we use a similar approach while modeling the areal shots.

To illustrate crosstalk problem in the prestack exploding reflector strategy let us make use of a simple constant velocity model of 2 km/s with two intersecting reflectors, one horizontal and the other dipping 15°. The dataset is comprised of 200 split-spread shots spaced every 20 m with a maximum offset of 2000 m. Figure 1 shows the shot-profile migration results using the correct velocity. The trace spacing in the migrated result is 20 m and the number of subsurface offsets in the prestack image is 41 sampled every 20 m. The front panel corresponds to the zero-subsurface offset section, the side panel is a SODCIG selected at $x=1.5$ km and the upper panel is a constant depth slice. We use this prestack image to model the areal data.

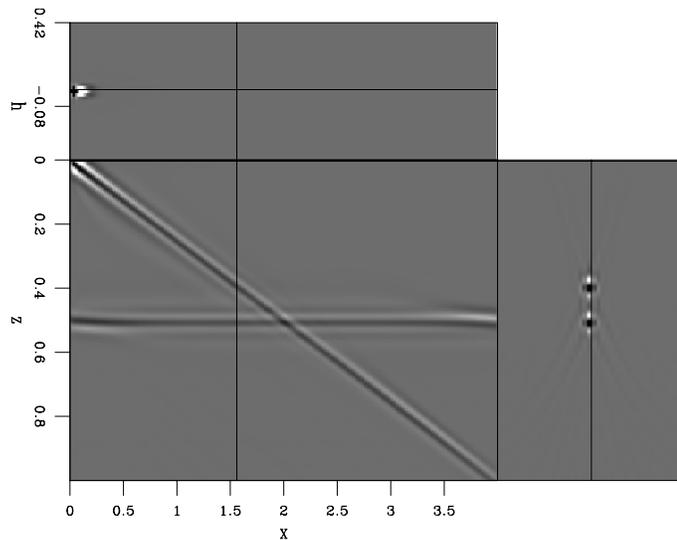


Figure 1: Prestack cube obtained with shot profile migration with the correct velocity. [ER]

claudio2/. mig08

Because we performed migration with the correct velocity and the energy is concentrated at zero-offset, the pre-processing step of rotating the SODCIGs according to the geological

dip was not necessary before modeling. Therefore, to model the areal source and receiver, SODCIGs are upward propagated using the one-way wave equation (equation 1) without any pre-processing step. The result of one modeling experiment is collected at the surface as an areal shot and assigned to its specific set.

Figures 2 and 3 show sets of time-domain areal sources and areal receivers computed with different distances between SODCIGs. Since the areal source wavefield is backward propagated it exists for negative times. For both figures, for the set of areal data labeled a) the distance between SODCIGs is 51 traces (1000 m), for the set labeled b) is 41 traces (800 m) and for the set labeled c) is 11 (200 m). These distances in number of traces defines the number of sets to be migrated and, consequently, the migration efficiency. As we show next, this also influences the intensity of the crosstalk.

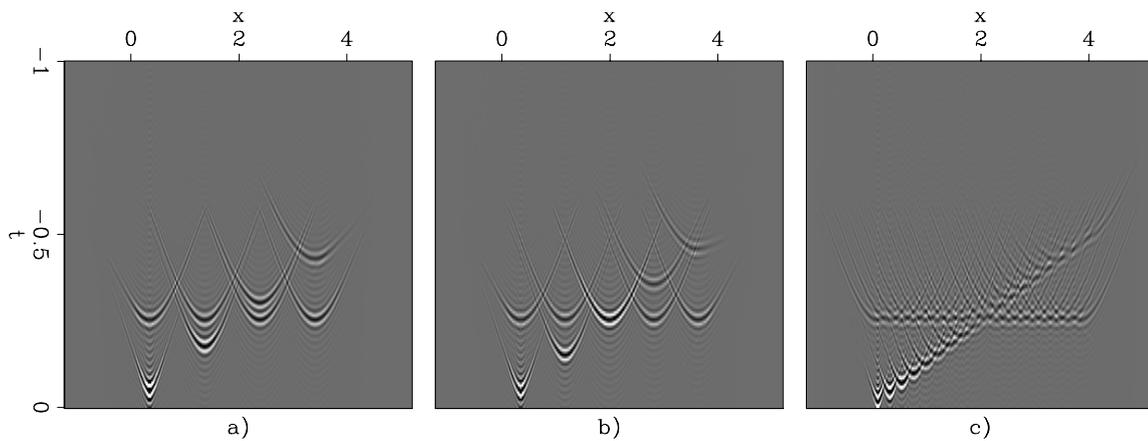


Figure 2: Sets of areal sources computed with different distances between SODCIGs. a) 51 traces; b) 41 traces; and c) 11 traces. [ER] [claudio2/. arsource](#)

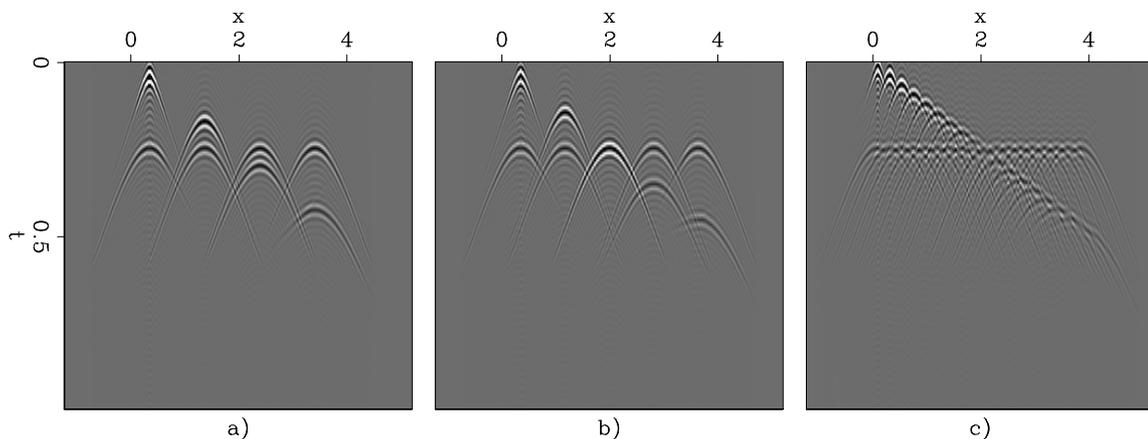


Figure 3: Sets of areal receivers computed with different distances between SODCIGs. a) 51 traces; b) 41 traces; and c) 11 traces. [ER] [claudio2/. arrecev](#)

The crosstalk intensity on a subsequent migration depends on the distance between SODCIGs making up a set of areal data. Biondi (2006) shows that, using a decorrelation

distance greater than twice the maximum subsurface offset, the crosstalks can be eliminated for the case of non-intersecting reflectors. For the case of reflectors with different geological dip and, moreover, intersecting, this procedure does not guarantee the absence of crosstalk.

Figures 4, 5 and 6 show the areal shot migration using cross-correlation imaging condition of the datasets containing the sets shown in Figures 2 and 3. The number of subsurface offsets is 41. Figure 4 shows the migration of 51 sets (Figure 2a and 3a), Figure 5 shows the migration of 41 sets (Figure 2b and 3b) and Figure 6 shows the migration of 11 sets (Figure 2c and 3c). The migration of the data with 51 sets shows no crosstalk in the SODCIG because the distance of 51 traces is bigger than the decorrelation distance. The migration of the data with 41 sets shows an acceptable level of crosstalk in the SODCIG. However, the crosstalk is strong in the SODCIG of the data with 11 sets. Notice that for all of the three results a strong crosstalk occurs at zero-subsurface offset (x,z -domain).

Figure 4: Areal shot migration of 51 sets, each set containing the summation of areal shots initiated at every 51th SODCIG. Original figure was substituted. [ER] claudio2/. armig51W

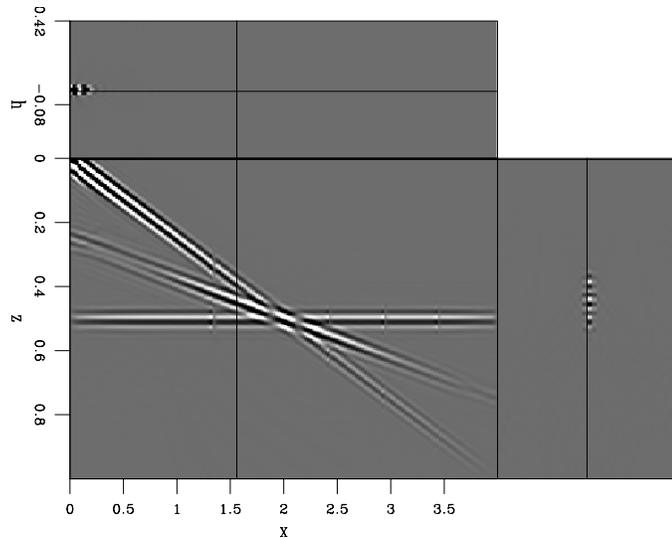
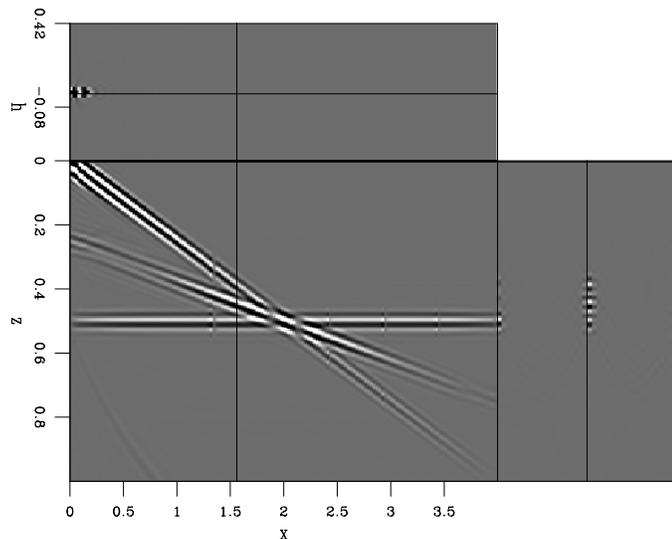


Figure 5: Areal shot migration of 41 sets, each set containing the summation of areal shots initiated at every 41th SODCIG. Original figure was substituted. [ER] claudio2/. armig41W



The crosstalk described in the previous paragraph have two distinct origins. The one occurring in the zero-subsurface offset section as a reflector with an intermediate dip results

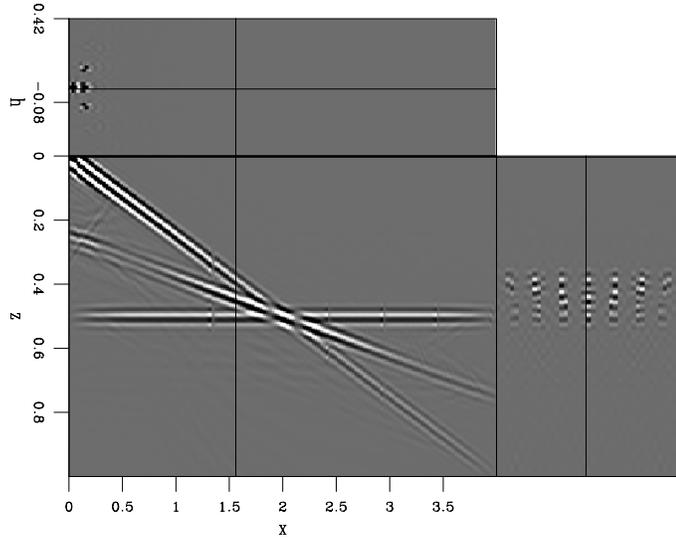


Figure 6: Areal shot migration of 11 sets, each set containing the summation of areal shots initiated at every 11th SODCIG. Original figure was substituted. [ER] claudio2/. armig11W

from the crosscorrelation of reflections in the source wavefield with reflections in the receiver wavefield from a single upward propagation experiment. The cross-correlation occurs at times different from zero propagation time. Therefore, one possible strategy is to use a time-windowed imaging condition about zero propagation time. The crosstalk present in non-zero offsets in the SODCIGs has similar origin as the one described by Sava (2007). In the case of the prestack exploding reflector modeling, they are related to reflections in the source and receiver wavefields pertaining to different areal data summed together into the same set. Sava (2007) uses an imaging condition that crosscorrelates decomposed source and receiver wavefields as a function of local slope at every position and time. This procedure yields good results, but at the expense of a higher computational cost.

The fact that the areal source wavefield also contains reflections originates crosstalk. According to the exploding reflector model (Loewenthal et al., 1976), reflectors explode at time zero. Therefore, the reflections in the areal source wavefield should be focused at zero time of wavefield propagation. We can use this information to restrict the cross-correlation in a time-window about zero time of wavefield propagation. As Biondi (2007) shows, the cross-correlation of areal wavefields within a small window around zero time of wavefield propagation can diminish the crosstalk at zero-subsurface offset. Equation 5 shows that, using the time-windowed imaging condition, crosstalk at the zero-subsurface offset only occurs between events having similar traveltimes.

$$\begin{aligned} \tilde{I}(z_\xi, x_\xi, h_\xi) = & I_1(z_\xi, x_\xi, h_\xi) + I_2(z_\xi, x_\xi, h_\xi) + \\ & \sum_{-\Delta t \leq t=0 \leq \Delta t} S_1^*(z_\xi, x_\xi, t) R_2(z_\xi, x_\xi, t) + \\ & \sum_{-\Delta t \leq t=0 \leq \Delta t} S_2^*(z_\xi, x_\xi, t) R_1(z_\xi, x_\xi, t). \end{aligned} \quad (5)$$

Applying this imaging condition using one-way propagators in frequency domain, requires storing both wavefields in memory for all the frequencies, which increases the memory requirements. On the other hand, applying the imaging condition for a small number of time samples saves much computational effort. Figures 7, 8 and 9 show the areal shot

migration using the time-windowed cross-correlation imaging condition from equation 4. Again, the number of subsurface offsets is 41. Figure 7 is the migration of 51 sets, Figure 8 is the migration of 41 sets and Figure 9 is the migration of 11 sets. Notice how the crosstalk in the zero-subsurface offset sections have been largely attenuated in each migration result.

Figure 7: Areal shot migration using time-windowed imaging condition of 51 sets, each set containing the summation of areal shots initiated at every 51th SODCIG. Compare with Figure 4. Original figure was substituted. [ER] `claudio2/. armig51T`

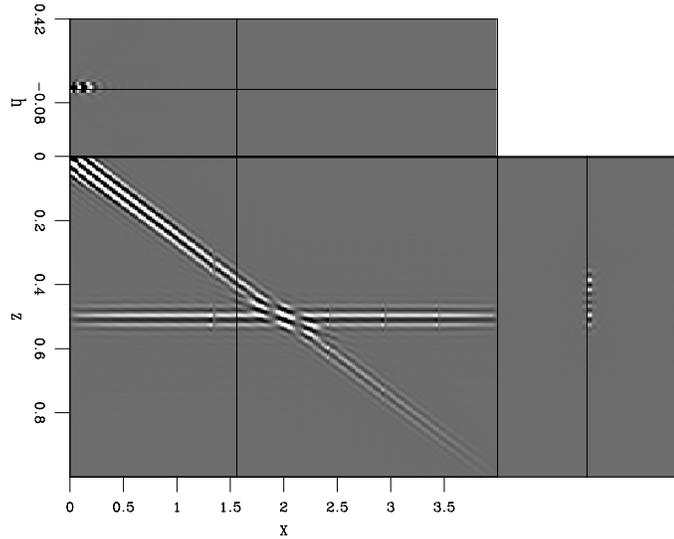
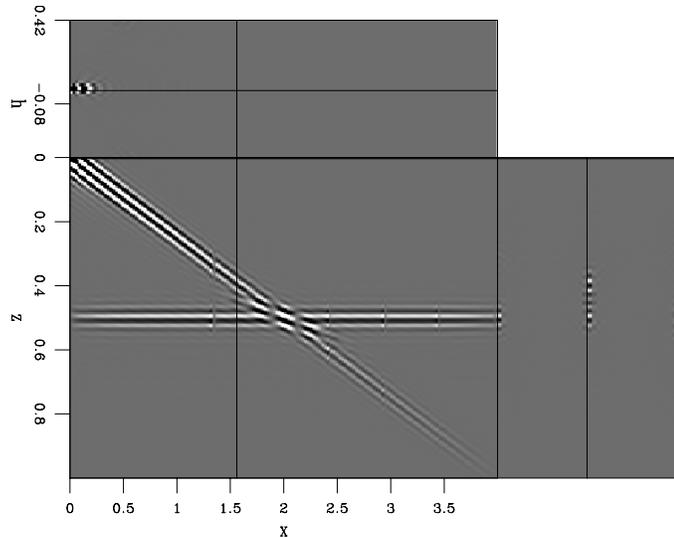


Figure 8: Areal shot migration using time-windowed imaging condition of 41 sets, each set containing the summation of areal shots initiated at every 41th SODCIG. Compare with Figure 5. Original figure was substituted. [ER] `claudio2/. armig41T`



For the migration of Figure 9, the time-windowed imaging condition, however, does not attenuate the severe crosstalk in the SODCIG. Ideally, the areal shots pertaining to the same set should be uncorrelated. A possible way to decrease the correlation between areal shots is by encoding them with a random phase function, $a(z_\xi, x_\xi, \omega)$, in such a way that the crosstalk is dispersed throughout the image. As Romero et al. (2000) show, for the random phase encoding we can choose $a(z_\xi, x_\xi, \omega) = e^{if(z_\xi, x_\xi, \omega)}$, where $f(z_\xi, x_\xi, \omega)$ is a random sequence as a function of z_ξ, x_ξ and ω .

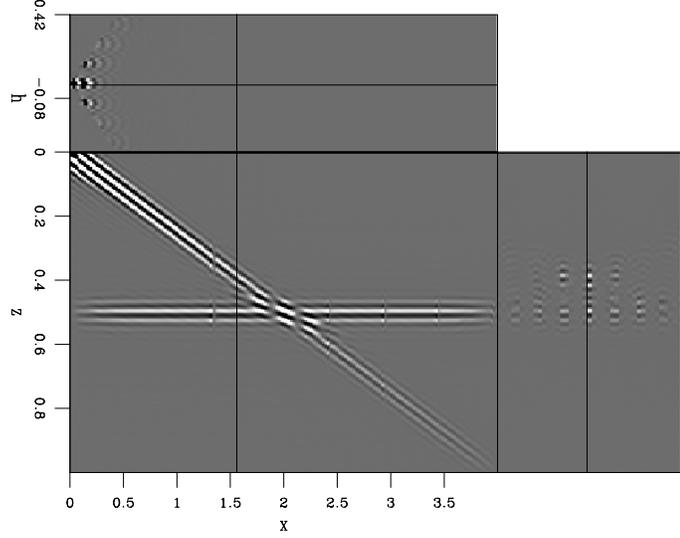


Figure 9: Areal shot migration using time-windowed imaging condition of 11 sets, each set containing the summation of areal shots initiated at every 11th SODCIG. Compare with Figure 6. Original figure was substituted. [ER] `claudio2/. armig11T`

The modeling with phase encoding synthesizes data according to

$$\begin{aligned} S_a(x, \omega) &= a(z_\xi, x_\xi, \omega)G(z_\xi, x_\xi - h_\xi; x, z = 0, \omega) * I_s(z_\xi, x_\xi, h_\xi), \\ R_a(x, \omega) &= a(z_\xi, x_\xi, \omega)G(z_\xi, x_\xi + h_\xi; x, z = 0, \omega) * I_r(z_\xi, x_\xi, h_\xi), \end{aligned} \quad (6)$$

and the sets of random phase encoded areal sources and receivers, $\widetilde{S}_n^a(x, \omega)$ and $\widetilde{R}_n^a(x, \omega)$, respectively, are given by:

$$\begin{aligned} \widetilde{S}_n^a(x, \omega) &= \sum_{n=1}^k \sum_{i=n, k, N} S_i^a(x, \omega), \\ \widetilde{R}_n^a(x, \omega) &= \sum_{n=1}^k \sum_{i=n, k, N} R_i^a(x, \omega), \end{aligned} \quad (7)$$

Figures 10a and 10b show a set of random phase-encoded areal source and receiver wavefields, respectively. Alternatively, we can encode just in x_ξ and ω . Figures 11a and 11b show a set of (x_ξ, ω) -random phase-encoded areal source and receiver wavefields, respectively.

The image, $\widetilde{I}_m(z_\xi, x_\xi, h_\xi)$, after areal shot migration of one set, m , comprised by two randomly phase-encoded areal sources and receivers is

$$\begin{aligned} \widetilde{I}_m(z_\xi, x_\xi, h_\xi) &= I_1(z_\xi, x_\xi, h_\xi) + I_2(z_\xi, x_\xi, h_\xi) + \\ &\sum_{\omega} a_1^* a_2 S_m^{a_1^*}(z_\xi, x_\xi - h_\xi, \omega) R_m^{a_2}(z_\xi, x_\xi + h_\xi, \omega) + \\ &\sum_{\omega} a_2^* a_1 S_m^{a_2^*}(z_\xi, x_\xi - h_\xi, \omega) R_m^{a_1}(z_\xi, x_\xi + h_\xi, \omega). \end{aligned} \quad (8)$$

The cross-correlation between the phase-encoding functions $a_1(z_\xi, x_\xi, \omega)$ and $a_2(z_\xi, x_\xi, \omega)$ must provide random values that disperse the crosstalk throughout the image.

Figures 12 and 13 show the areal shot migration results of 11 sets of randomly encoded areal data using encoding functions varying in (x_ξ, ω) and (z_ξ, x_ξ, ω) , respectively. The

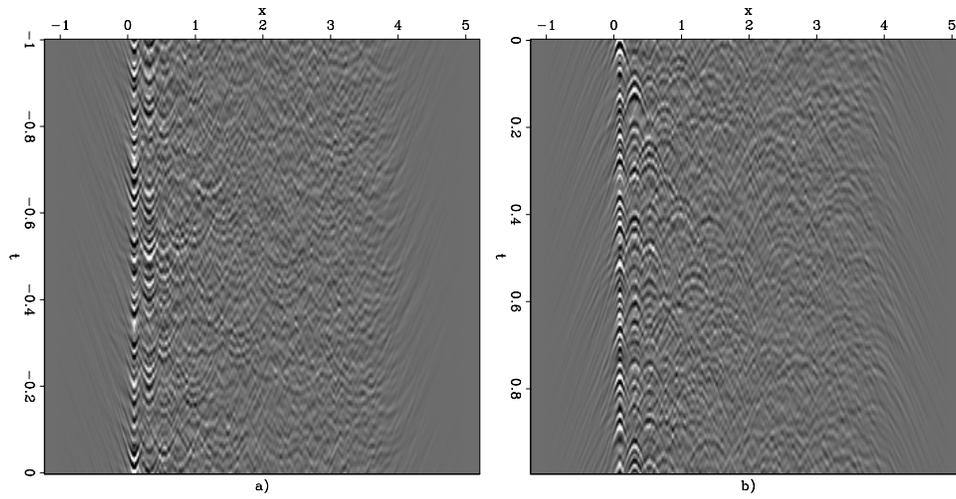


Figure 10: Areal source (a) and receiver (b) wavefields computed with random phase encoding, $a(z_\xi, x_\xi, \omega)$. [ER] [claudio2/. arrphz](#)

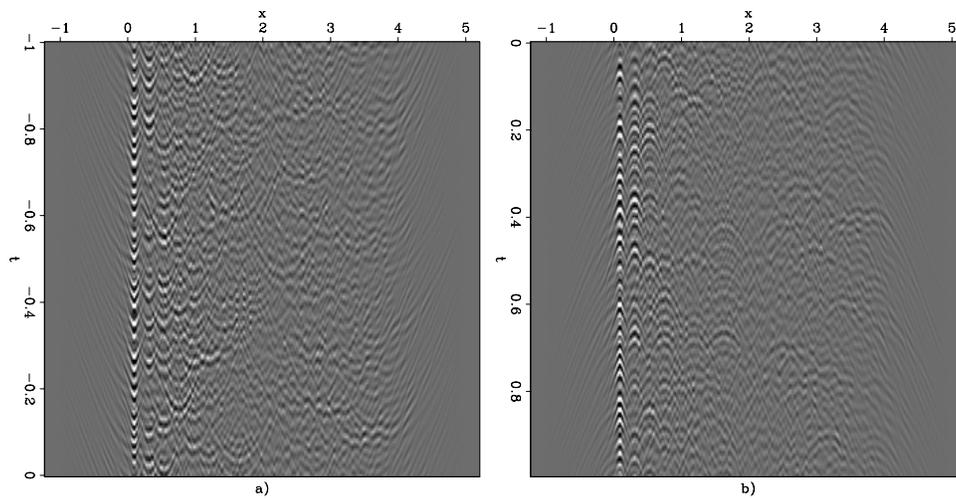


Figure 11: Areal source (a) and receiver (b) wavefields computed with random phase encoding, $a(x_\xi, \omega)$. [ER] [claudio2/. arrphw](#)

comparison between Figure 12 and Figure 6 shows that encoding only in (x_ξ, ω) yields a significant reduction of the SODCIG crosstalk. The crosstalk in the zero-subsurface offset, however, is much less affected. The comparison between Figure 13 and Figure 12 shows that encoding in (z_ξ, x_ξ, ω) performs as efficiently as encoding in (x_ξ, ω) in dispersing the SODCIG crosstalk. However, for the crosstalk in the zero-subsurface offset section, caused by cross-correlation of reflections in the areal source and receiver wavefields originated from the upward propagation of the same SODCIG, encoding in (z_ξ, x_ξ, ω) performs much better than encoding in (x_ξ, ω) . This is because the additional encoding in (z_ξ) decorrelates the exploding reflectors.

Since the number of phase-encoded areal data is small, the noise dispersed through the image is not negligible. By migrating more sets of (z_ξ, x_ξ, ω) -randomly phase-encoded areal data, these amplitudes can be largely attenuated. Figure 14 shows the superior result when compared to Figures 6, 12 and 13, obtained by migrating 4 different realizations of 11 sets of (z_ξ, x_ξ, ω) -randomly encoded areal data. Besides virtually eliminating the crosstalk, the speckled noise has lower amplitudes than in Figure 13.

Figure 12: Areal shot migration of randomly encoded data. 11 sets of (x_ξ, ω) -randomly encoded areal data were migrated. Compare with Figure 6. Original figure was substituted. [ER] [claudio2/. arm11phw](#)

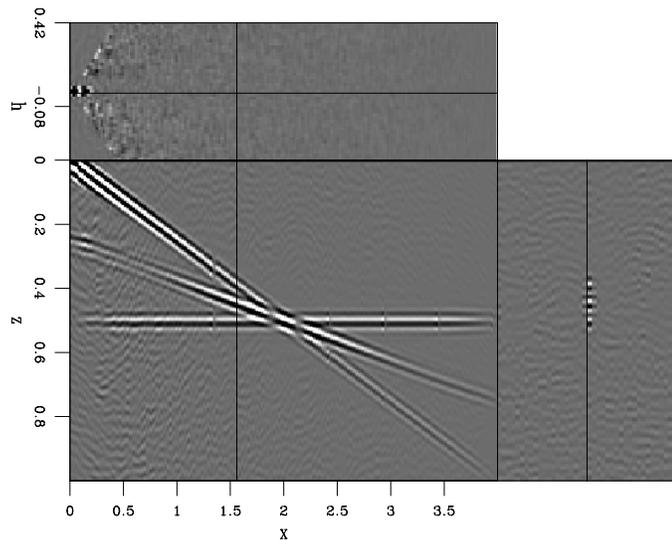


Figure 13: Areal shot migration of randomly encoded data. 11 sets of (z_ξ, x_ξ, ω) -randomly encoded areal data were migrated. Compare with Figure 6. Original figure was substituted. [ER] [claudio2/. arm11phz](#)

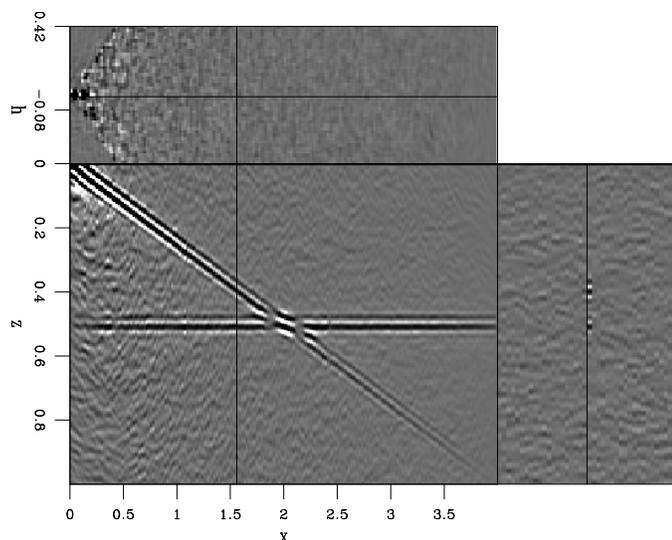
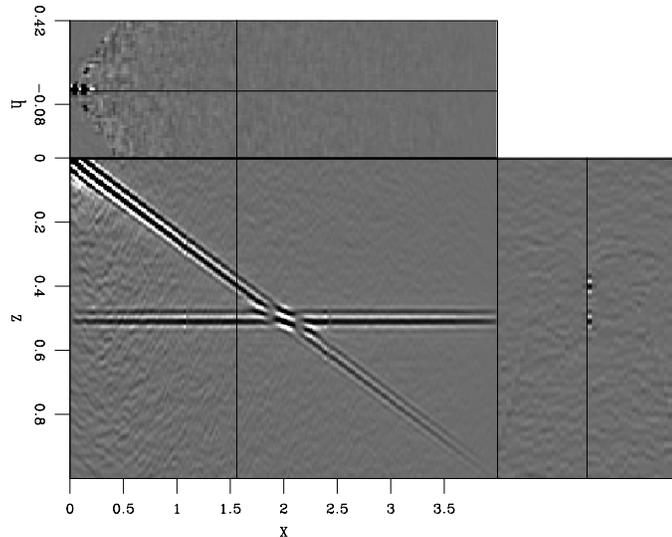


Figure 14: Areal shot migration of randomly encoded data. 4 groups of 11 sets of (z_ξ, x_ξ, ω) -randomly encoded areal data were migrated. Compare with Figures 6, 12 and 13. Original figure was substituted. [ER]

claudio2/. arm11phz4



One drawback of applying random phase encoding is that it makes the reflectors randomly “explode” at times different from the zero time of the wavefields. Consequently, the time-windowed imaging condition can not be applied. Perhaps, the linear phase encoding will enable us to use the time-windowed imaging condition. This deserves future research.

CONCLUSIONS

The prestack exploding reflector method potentially can decrease data volume required to perform migration velocity analysis. To be cost efficient, the method relies on the linearity of the wave-propagation to combine modeled areal shots into areal data super-sets. This process gives rise to crosstalk when migrating the areal data.

We show that applying a time-windowed imaging condition efficiently mitigates the crosstalk in the zero-subsurface offset section, but does little to attenuate SODCIG crosstalk. We present a strategy to perform random phase encoding during the modeling. The results are promising and much better than the time-windowed imaging condition with respect to attenuating SODCIG crosstalk. The next steps will be the application of this strategy on data migrated with an inaccurate velocity model as well as on data from complex geology.

Because of the random encoding, the reflectors do not focus anymore about the zero time of the wave propagation, disabling the use of the time-windowed imaging condition. Further research will investigate the use of linear phase encoding associated to the application of this imaging condition.

REFERENCES

- Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: SEP-Report, **124**, 45–60.
 ———, 2007, Prestack modeling of image events for migration velocity analysis: SEP-Report, **131**.

- Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: *Geophysics*, **71**, 129–139.
- Loewenthal, D., L. Lu, R. Roberson, and J. Sherwood, 1976, The wave equation applied to migration: *Geophysical Prospecting*, **24**, 380–399.
- Rietveld, W. E. A. and A. J. Berkhout, 1992, Prestack depth migration by means of controlled illumination: *Geophysics*, **59**, 801–809.
- Rietveld, W. E. A., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Optimum seismic illumination of hydrocarbon reservoirs: *Geophysics*, **57**, 1334–1345.
- Romero, L., D. Ghiglia, C. Ober, and S. Morton, 2000, Phase encoding of shot records in prestack migration: *Geophysics*, **65**, 426–436.
- Sava, P., 2007, Stereographic imaging condition for wave-equation migration: *Geophysics*, **72**, A87–A91.
- 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. Subsalt imaging examples: *Geophysical Prospecting*, **52**, 607–623.
- Schultz, P. S. and J. Claerbout, 1978, Velocity estimation and downward continuation by wavefront synthesis: *Geophysics*, **43**, 691–714.
- Sun, P., S. Zhang, and F. Liu, 2002, Prestack migration of areal shot records with phase encoding: 72nd Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1172–1175.

Ray tracing modeling and inversion of light intensity under a water surface

Abdullah Al Theyab

ABSTRACT

Light propagating in a water-filled pool is perturbed by the water surface, creating patterns on the pool floor. In this report, I use ray tracing to compute an approximation of the light intensity field on the pool floor using point source and exploding surface models. The ultimate goal is to infer the water surface from the intensity field. In a seismic imaging, it is similar to imaging using amplitudes rather than travel-times. I present a geometric approach to invert for a discretized surface from a ray-count representation of the intensity field. With this formulation, the inversion becomes a combinatorial problem, which can be solved using non-deterministic search techniques. The formulation has a large inherent null space. The low cost of the technique allows a large number of iterations to be applied.

INTRODUCTION

Light propagation is similar in both nature and theory to seismic wave propagation in the subsurface. It is not surprising to see a large number of common problems between the fields of seismology and optics. Claerbout (2007) poses a question about the relationship between light patterns under water and seismology. One prominent problem in reflection seismology is estimating seismic velocities in the heterogeneous subsurface. Velocity is a measure that depends on travel time; i.e. it is determined by the moveout of seismic events. It is always costly to estimate subsurface velocities, whether using conventional velocity analysis or inversion methodologies. Unfortunately, imaging algorithms rely on the accuracy of velocity models. Therefore, a curious scientist might ask: *Can we use amplitudes alone for imaging?*

For the sake of simplicity, let us consider a simple imaging problem involving a single wavefield caustic: a single reflector including a syncline. Using the exploding reflector concept, we can simulate a zero-offset section (Claerbout, 1985). What we will observe in the section is a bowtie that is caused by the syncline. The syncline causes many ray paths between the exploding reflector to the receiver to converge at some point before reaching the surface, forming a caustic. Given the correct velocity model, migration algorithms like Kirchhoff migration will resolve the bowtie into a syncline (Yilmaz, 2001).

Now, let us consider a swimming-pool experiment where the goal is to infer the water surface from the light patterns on the floor of the pool like the ones shown in Figure 1. We might encounter similar difficulties as in seismic imaging. It is hard to estimate the exact speed of light in the pool. In this case the question is as follows: *Can we use the light intensity field at the bottom of the pool to infer the surface of water?*

The two questions are closely related, and the pool experiment is easier to comprehend intuitively. In this paper, I implement the forward simulation using ray theory to build a light intensity field at the pool floor that can be used for inferring the surface. I will not follow the exact physics as it is done in optical simulations: I assume infinite frequencies for ray theory and representing light with a finite number of rays. The forward simulation is done using ray tracing and beam tracing. I finally discuss a Monte Carlo ray tracing inversion that is based on ray counting.

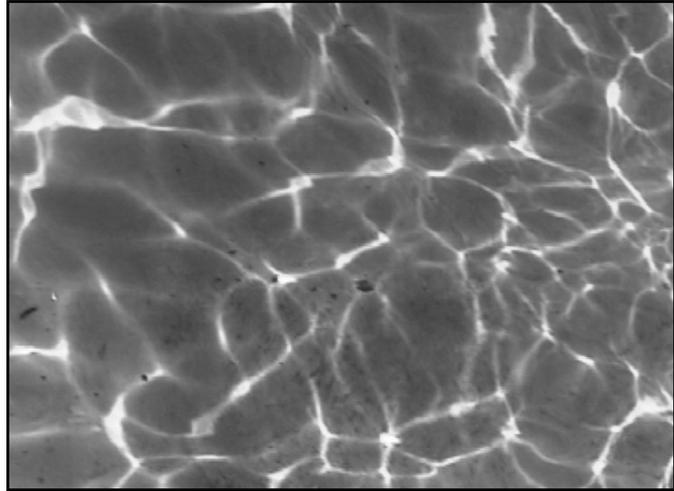


Figure 1: Light patterns at the bottom of a pool (Claerbout, 2007).

[NR] `altheyab1/. fig-jonpool`

FORWARD SIMULATION

I simulate the light intensity field at the pool floor using ray tracing. For our purposes, we assume that light travels with an infinite velocity; i.e. any ray refracting through the water surface will project instantly onto another point on the floor. If multiple rays for any reason converge at a point before reaching the floor, they form a refracted caustic (Shah et al., 2007).

This simplistic view of the experiment has received considerable attention in the field of computer graphics. Shah et al. (2007), for example, introduce a real-time technique for rendering caustics from reflective and refractive surfaces. Ray tracing techniques have a major deficiency, in that we need an infinite number of rays to simulate a realistic distribution of light. This shortcoming is discussed by Watt (1990) who suggests using beam tracing as an alternative. The vast majority of the techniques developed in computer graphics, however, are developed for a 3D space having point light source(s), polygonal objects, and a single observation point.

Our simplified case inherits only a minor subset of the wide range of techniques used in computer graphics. To start, we will not have an observer point, and therefore do not need to ray trace from the pool floor to the observer point. Also, we have only a single smooth surface of moderate relief, and the incident light direction is limited to rays arriving from above the surface. In other words, we have a unique refraction from each point on the surface. Because we have moderate topology in the water surface, we will assume that the refracted rays do not re-intersect the water surface.

Surface Representation

In (x, y, z) -space, the surface of the water is given as a discretized function on a uniformly sampled grid in one or two dimensions. The function has points that represent the depth of the surface from the xy -plane. If we have a surface $S = f(x, y)$ in 3D space, the normal vector to that surface is

$$\vec{n} = \frac{-\frac{\partial S}{\partial x} \hat{\mathbf{i}} - \frac{\partial S}{\partial y} \hat{\mathbf{j}} + \hat{\mathbf{k}}}{\sqrt{\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + 1}} \quad , \quad (1)$$

and for a surface $S = f(x)$, in 2D space, it is

$$\vec{n} = \frac{-\frac{\partial S}{\partial x} \hat{\mathbf{i}} + \hat{\mathbf{k}}}{\sqrt{\left(\frac{\partial S}{\partial x}\right)^2 + 1}} \quad . \quad (2)$$

where $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are unit vectors along the x -, y -, z -axes, respectively. To compute the first partial derivatives, we can use finite-divided-difference approximations (Chapra and Canale, 2002). A design decision must be made at this point regarding where the refracted rays will start with respect to the points on the surface function. For a 2D simulation with $S(x)$, we can have the ray coming out of the points. In that case, the preferred scheme is centered finite-divided-difference:

$$\frac{\partial}{\partial x} S(x_i) \approx \frac{S(x_{i+1}) - S(x_{i-1}))}{2\Delta x} \quad . \quad (3)$$

The same differencing scheme can be used for the partial derivatives with respect to x and y for the 3D simulation. Using this scheme, it is not possible to compute the partial derivatives for the edges of the surface, since more points are needed for the computation.

The second design option is to have the rays launching from segments of the surface between the points in the 2D simulation. For the 3D case, we divide every quad – i.e. square between four adjacent points– using two diagonals. For every possible triangle within a quad, we compute the normal and have a ray starting from the center of the triangle. The advantage of this approach is that we have many more rays than the number of surface samples. For computational stability, the surface must be smooth locally with respect to a segment or a quad.

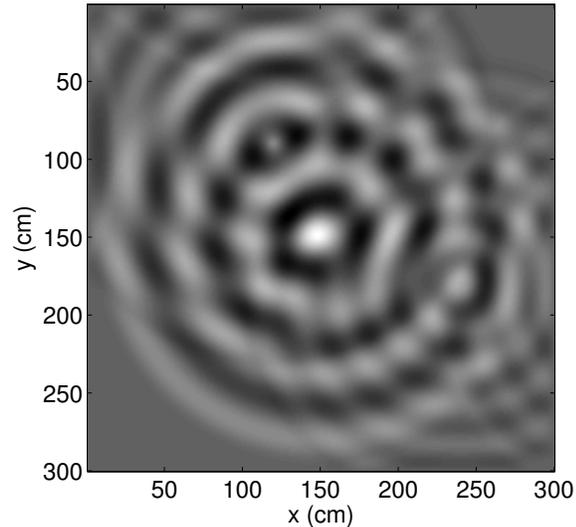
An easy way to construct a synthetic water surface with a sinusoidal wave with a decaying factor as a function of the distance from the source:

$$S(x, y) = c_0 \cos\left(c_2 |\vec{d}|\right) e^{-c_1 |\vec{d}|} \quad , \quad (4)$$

where \vec{d} is the distance from the source, and c_0 , c_1 , and c_2 are arbitrary positive constants to customize amplitudes, decay rate, and angular frequency of the ripples, respectively. Waves originating from several source points can be summed to form a more complex surface. This method of construction is for a time-invariant surface without reflection at the surface boundaries. A more realistic water surface can be obtained using finite differencing of the 2D wave equation. For the sake of simplicity, I use the explicit differencing scheme in the (x, y, t) -domain. Figure 2 is the result of finite differencing with three sources that are Gaussian wavepackets of different amplitudes.

Figure 2: Finite-difference simulation of a water surface with one strong and two weaker sources disturbing the water surface. [ER]

`altheyab1/. fig-finitdiff`



Light Source

The light refracting into the pool can be from a fixed point source from which light is spreading radially, or ambient light scattering from all directions. Usually, the light incident on a real water surface is a combination of both. Because I use rays for modeling light, I classify sources into two categories. The first category includes point sources, from which light is represented as rays traveling to the surface and refracting onto the pool floor (specular-to-specular transport (Watt, 1990)). In the second category, ambient light refracts through the water surface and forms rays (diffuse-to-specular transport (Watt, 1990)). I approach the second case with the notion of an exploding surface.

For a point source, I start by computing the incident ray \vec{s} on each refraction point on the surface. Snell's law can be used in the plane containing both \vec{s} and \vec{n} to find the refracted ray \vec{r} into the water. First, however, the incident ray must be checked to insure that it is not coming from below the surface. This occurs when the source is located very close to the surface creating shadow zones behind high water ripples where there are no rays incident on the surface. Figure 3 shows the refraction of rays into water as well as some shadow zones.

As mentioned earlier, I use the exploding surface concept for the diffuse-to-specular mechanism of light transport through the surface. This concept is analogous to the exploding reflector concept that Claerbout (1985) uses as an introductory model for imaging. The refracted ray from each point on the exploding surface travels in the same direction as the surface normal at that point; i.e. $\vec{r} = \vec{n}$. This is an easier approach to the modeling than working with a point source, but it is less accurate.

Ray Representation

Having developed representations of the surface and of the light entering the surface, the next step is mapping the refracted rays from the surface to the pool floor. At this stage, a

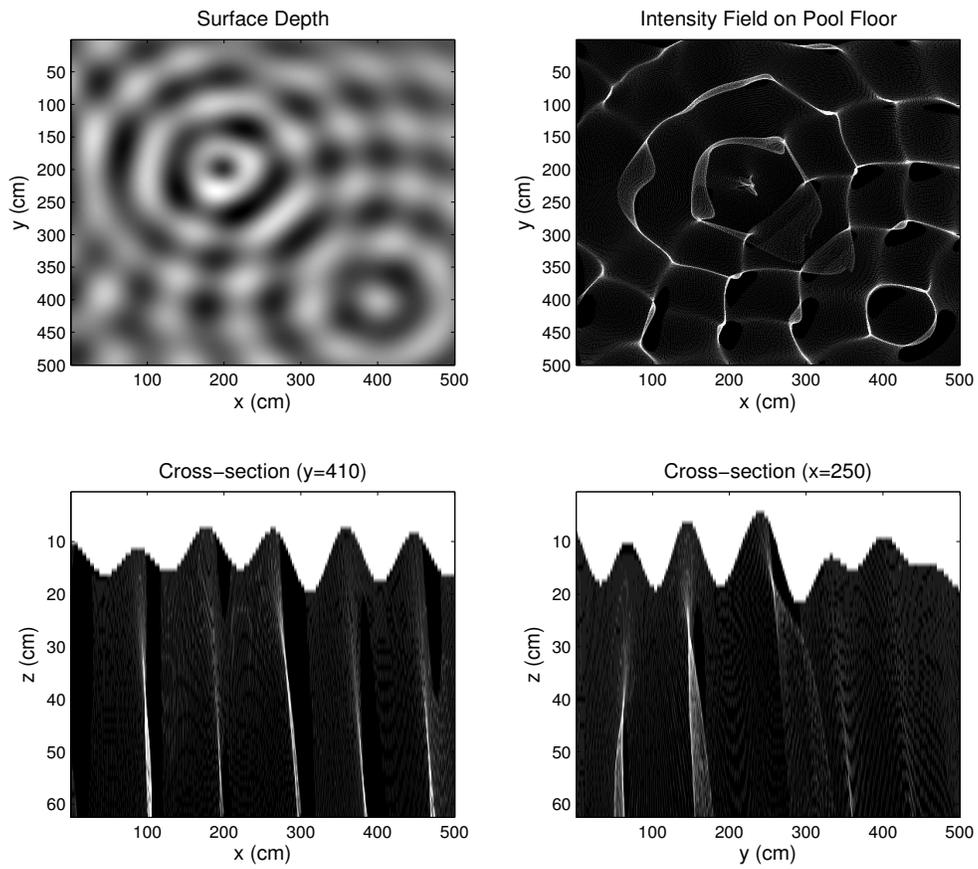


Figure 3: 3D ray tracing simulation with a point source. [ER] `altheyab1/.fig-snells3d`

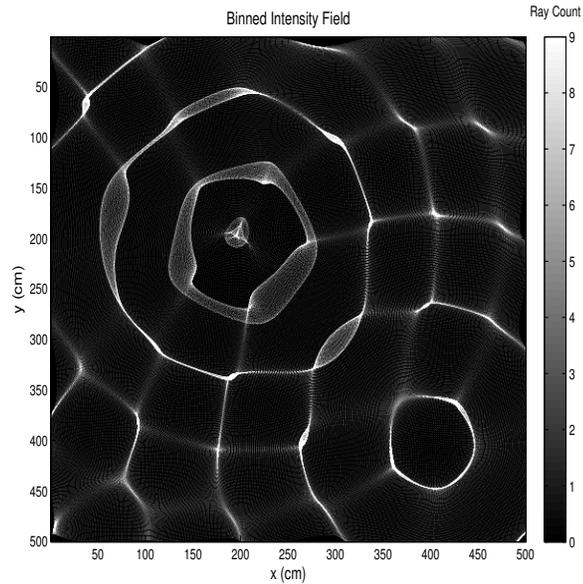
ray is simply a normalized vector \vec{r} at a surface position pointing downward. This vector controls the contribution to the light intensity field. The contribution can be a simple binned ray or a beam.

It is a simple exercise of trigonometry to project the refracted ray \vec{r} to a point \vec{p} the bottom of the pool.

$$\vec{p} = \frac{\text{depth}}{r_z} \vec{r} \quad (5)$$

The point \vec{p} is unlikely to fall on a grid point, and therefore the ray amplitude distribution can be binned. If each ray is given a unit amplitude, the sum of amplitudes is the ray count for each bin. Figure 4 shows the result obtained using binning. The observable caustics have sharp, though not continuous boundaries. Zones of low light intensity can be confused with shadow zones, simply because the light is not distributed continuously. Some of the image degradation might be attributed to the nearest-neighbor interpolation; however, even with the best interpolation algorithms, the result will not improve significantly. In Figure 3, the aliasing is due to an insufficient number of rays and binning. de Ridder (2008) suggests using a Gaussian distribution as a function of the lateral distance from the ray projection point on the pool floor. The proposed method yields more stable results. Nevertheless, the symmetry of the distribution neglects the tilting of rays.

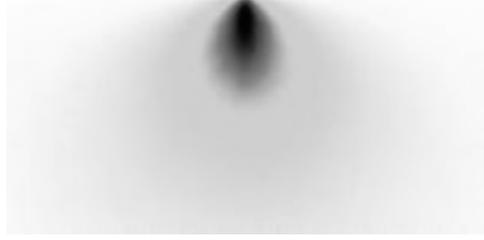
Figure 4: The result of binning is a good approximation but not realistic. [ER] [altheyab1/. fig-binning3d](#)



So far I have used the specular-transport mechanism within water and ignored light diffusion, which is a shortcoming of the ray-tracing. It would take orders of magnitude more rays for ray tracing to simulate diffusion, which is prohibitively expensive. Also, it will be inefficient, because diffusion does not change much from one pixel to the next (Watt, 1990). Figure 5 shows the diffusion of light incident from a point on a disordered medium (ICMM, 2008). It is obvious that light intensity within a beam is dependent on the angle of refraction. It can be modeled with a normal distribution as a function of the angle, with a decay factor that is a function of distance from the refraction point. In vector notation, the contribution of the diffusion to a point is

Figure 5: Diffusion of a light beam in an isotropically disordered medium (ICMM, 2008). [NR]

altheyab1/. fig-diffusion



$$\rho(\vec{p}) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(|\vec{r} \times \vec{p}|)^2}{2\sigma^2|\vec{p}|^2} - \alpha|\vec{p}| \right\} , \quad (6)$$

where \vec{p} is the position vector from the refraction point (x_0, y_0, z_0) on the surface that is defined as

$$\vec{p}(x, y, z) = (x - x_0)\hat{\mathbf{i}} + (y - y_0)\hat{\mathbf{j}} + (z - z_0)\hat{\mathbf{k}} , \quad (7)$$

σ determines how narrow the distribution is, and α is the rate of exponential decay.

Beams¹ demonstrate some of the characteristics of both specular and diffusive transport mechanisms. One way of handling diffusive and specular transport is to model them separately and stack the results. However, I use beams to model the two simultaneously.

In designing the beam shape, one can choose to favor either the diffusive or specular distribution. Favoring diffusion by using equation 6 can negatively affect the resolution of the light patterns. The following equation can be used to describe the distribution of light within the beam:

$$\rho(\vec{p}) = \frac{1}{b} \exp \left\{ -\frac{|\vec{r} \times \vec{p}|}{2b} - \alpha|\vec{p}| \right\} , \quad (8)$$

where b controls the width of the beam, and α is the rate of exponential decay. Figures 6 and 7 show the beam coming from a single point on the surface and the section obtained by stacking all the beams coming from the surface respectively.

The disadvantage of using beam tracing is that every beam contributes to a very large number of points, which can be computationally more expensive $-O(N^2)-$ than binning $-O(N)$, where N is the number of samples on the surface function. As mentioned earlier, the boundaries of the caustics are not as sharp. However, they become more realistic, as in Figure 1, where the details of caustics are not usually seen with sharp boundaries. Figure 8 is the result of the 3D beam tracing of the surface in Figure 2. Figure 9 shows the light intensity field of a surface distorted by many ripples. Although it is more expensive, beam tracing works better than ray tracing. We can distinguish shadow zones and the elongated caustics that usually connect two point caustics.

Beam tracing has many of the limitations of ray tracing with binning. The prominent limitation is that beam tracing needs a large number of beams to model light intensity under a very narrow trough on the surface. Also, for depths where many rays do not project on

¹Beams discussed here are different from beams used in seismic imaging.

Figure 6: A single beam refracting into the pool. [ER]
`altheyab1/. fig-beam`

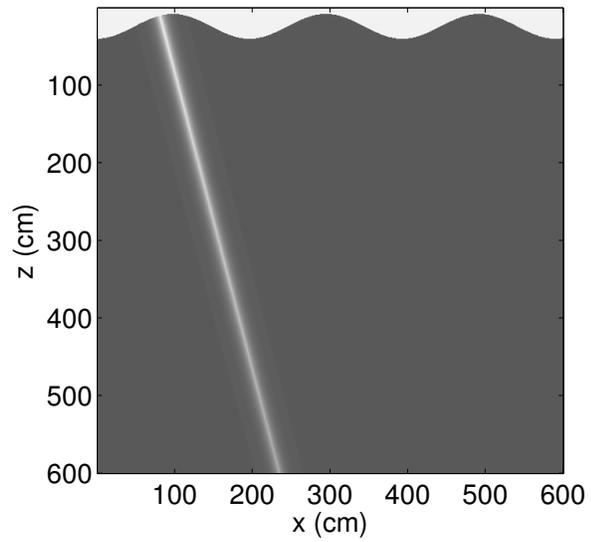
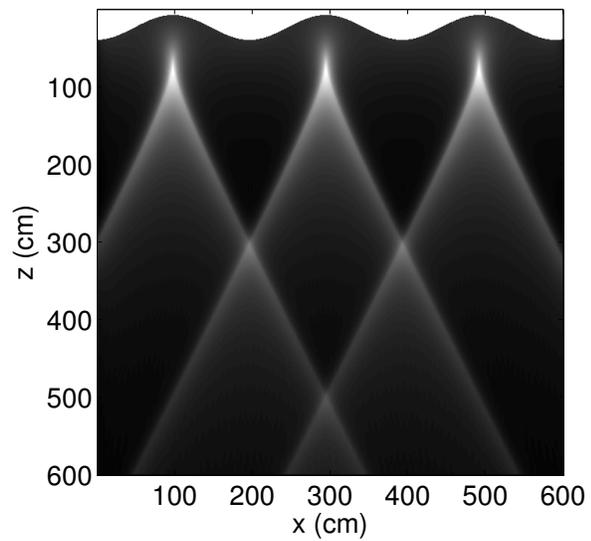


Figure 7: 2D forward modeling using beam tracing. The refracted light focuses into caustics. [ER]
`altheyab1/. fig-section2d`



the pool floor, there are too few beams to construct the correct intensity field. Moreover, the quality of the forward modeling results is sensitive to the beam width chosen; too wide beam hinders the resolution, and too narrow beams behave like rays and can produce aliased results.

Figure 8: Intensity field under the surface at Figure 2. [CR]
[altheyab1/. fig-finitdiffproj](#)

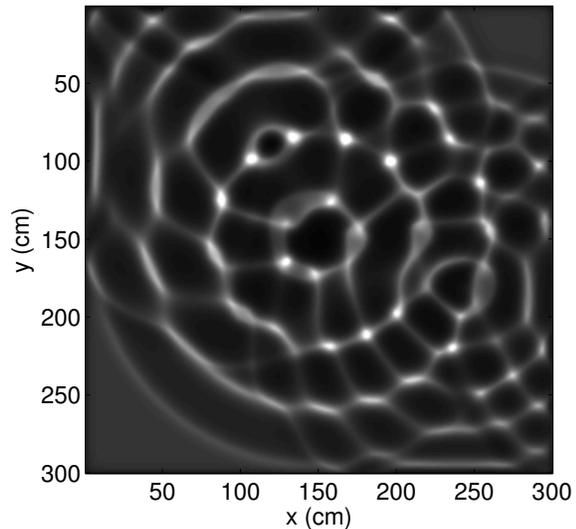
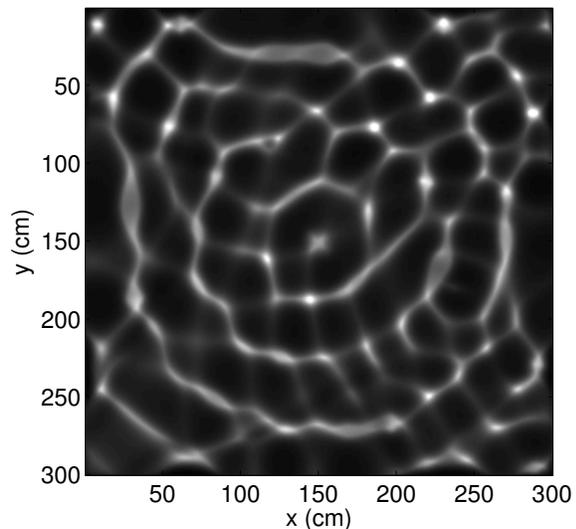


Figure 9: A result from the exploding surface model. [CR]
[altheyab1/. fig-exploding3d](#)



RAY-TRACING INVERSION

As shown earlier, ray tracing with binning can give a quick but crude approximation of the light intensity field at the pool floor. If we have a ray count for light intensity created using the exploding surface model, we can use a simple geometric approach for inversion. In two dimensions, the ray count in a single bin corresponds to the number of surface segments whose normal vectors point to that bin. If no rays hit the side-walls of the pool, the surface

can be reconstructed by assigning each surface segment to one unit of ray count in the intensity field; every assignment means the starting and the ending points of a refracted ray. However, there are many permutations of assignments that can produce the same intensity field. With this formulation, the inversion becomes a combinatorial optimization problem which could be implemented using a heuristic search method. The right panel of Figure 10 shows one possible surface obtained using simulated annealing. Although the obtained solution might be physically inaccurate, the ray counts produced by the two surfaces in the figure are identical. This numerical non-uniqueness is the result of simplified implementation of the theory – finite ray tracing with binning.

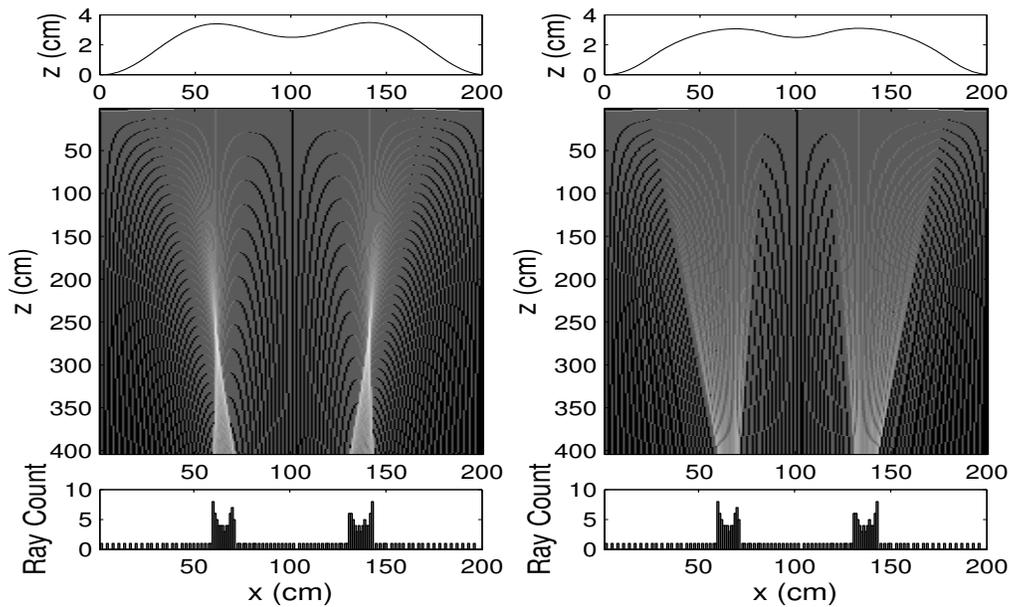


Figure 10: Left: ray tracing forward modeling for a surface. Right: One of the many surfaces obtained by inversion using ray-count. The upper graphs are the surface function, middle plots are the ray tracing of light under the surface, and the bottom plots show the ray count at the bottom of the pool. [CR] [altheyab1/.fig-subcaustics](#)

Despite the large null space, a large number of solutions can be tested at a minimal cost because the full forward modeling is not needed. The desired solution lies within the null space. Most of the solutions in the null space are physically impossible. Therefore, these solutions can be discarded using an energy function. For example, for a pre-caustic slice the energy function can be the sum of distances traveled by the rays:

$$E(\text{solution}) = \sum_i^N |\vec{p}_i| \quad . \quad (9)$$

Minimizing the energy function proposed yields the correct solution if given enough iterations (see Figure 11). In fact, simulated annealing can be reduced to the greedy algorithm, which moves only toward better solutions. Inversion for intensity fields below caustics needs a more elaborate energy function. The resolution that binning provides is insufficient to

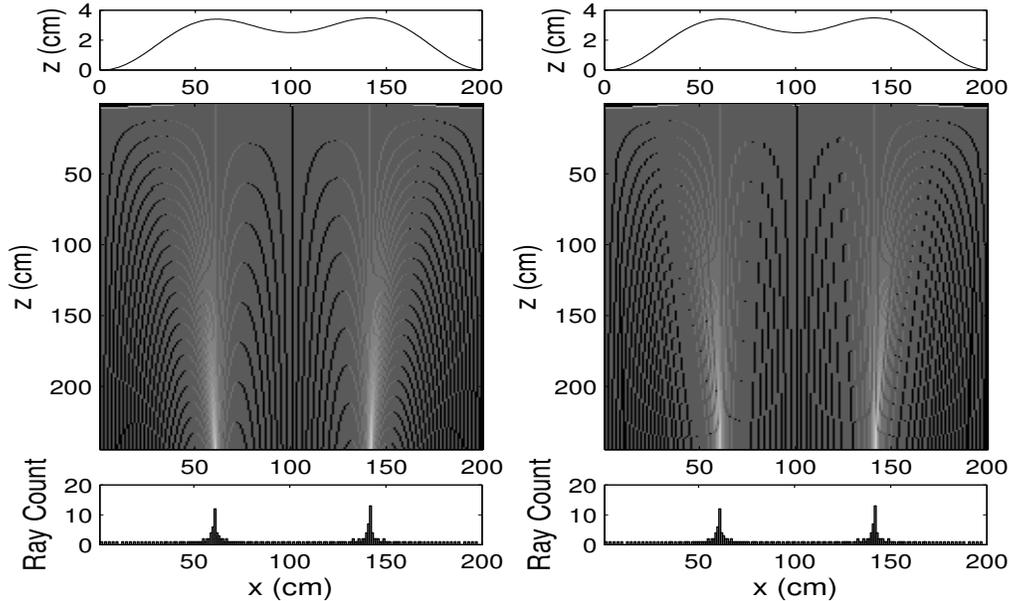


Figure 11: Left: ray tracing forward modeling of pre-caustics intensity field. Right: The inversion result by minimizing distances traveled by rays. [CR] [altheyab1/. fig-precaustics](#)

resolve the caustics, especially where the caustics overlap. In three dimensions, there are more constraints on the solutions because neighboring quads of the surface must match for the surface to be continuous. Therefore, I expect that the search space reduces radically, and it can be traversed using a tree search algorithm to find a good approximate solution. Using this geometric search for a solution might not obtain the result but it can provide very cheaply a good approximation that can be used as a starting model for other inversion of real function.

CONCLUSIONS AND FUTURE WORK

In this paper, I show how we can construct the light intensity field at the pool floor using approximate operators. A light intensity field can provide information about the geometry of the water surface. Inversion using patterns obtained from ray tracing and binning suffers from a large null space—i.e. many surfaces can produce the same patterns at the bottom of the pool. Future directions include using the proposed geometric inversion approach in three dimensions, and formulating and implementing the inversion problem using a realistic light intensity field.

ACKNOWLEDGMENTS

I thank Sjoerd de Ridder for sharing his ideas about stabilizing the forward modeling using distributions. Thanks to Jeff Shragge, Yaxun Tang, and the rest of SEP group for their contribution and for helping me during my struggle with the computing environment.

REFERENCES

- Chapra, S. and R. Canale, 2002, Numerical Methods for Engineers - Fourth Edition: The McGraw-Hill Companies Inc.
- Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific.
- , 2007, A tease for theoretical seismologists: WWW. (URL <http://sepwww.stanford.edu/sep/jon/poolwaves>).
- de Ridder, S., 2008, Stable simulations of illumination patterns caused by focusing of sunlight by water waves: Technical report, SEP-134.
- ICMM, P. C. G., 2008, Introduction to light diffusion: WWW. (URL http://luxrerum.icmm.csic.es/?q=node/research/light_diffusion).
- Shah, M., J. Kontinnen, and S. Pattanaik, 2007, Caustics mapping: An image-space technique for real-time caustics: IEEE Trans. Vis. Comput. Graph, **13**, 272–280.
- Watt, M., 1990, Light-water interaction using backward beam tracing: Computer Graphics, 377–385.
- Yilmaz, O., 2001, Seismic Data Analysis: Society of Exploration Geophysicists.

Stable simulations of illumination patterns caused by focusing of sunlight by water waves

Sjoerd de Ridder

ABSTRACT

Illumination patterns of underwater sunlight have fascinated various researchers in the past. I derive a set of equations that models these patterns at arbitrary depths from arbitrary surface topology functions. The rays are approximated by a statistical distribution function that is shifted to different lateral positions depending on the ray's refraction angle at the water surface. I perform simulations using either delta or Gaussian distributions. Using delta distributions proves unstable because of the Gibbs phenomenon, which can be suppressed using a low-pass filter. The simulations using a Gaussian distribution are stable and, apart from minor smoothing, do not suffer from additional artifacts.

INTRODUCTION

The topography of the water surface bends sunlight to form sharp stripes and exotic patterns under the surface. Claerbout (2007) relates these light patterns under water to seismology, bright spots in seismic sections usually associated with hydrocarbon reservoirs. Some of these bright spots do not reflect rock or reservoir characteristics, but are instead manifestations of wavefield focusing by lateral velocity variations in the overburden. This paper addresses accurate modeling illumination patterns caused by focusing of sunlight by water waves.

Much work has been done to understand and predict the amplitude spectra of light-field fluctuations with water depth. Schenck (1957) was the first to derive a model of underwater sunlight focusing, he derived the direction of light rays under the water surface in 2D using Snell's law and a function describing the water surface. His research and later work aimed at satisfying the needs of marine biologists to predict the photosynthetic production in the photic zone (Dera and Gordon, 1968). Snyder and Dera (1970) did a more careful analysis in 3D, deriving a set of equations that describe the focusing to first order but are valid only near to the water surface. This model could be used as a starting point for inversions or adjoint modeling for illumination patterns close to the water surface. Nikolayev and Khulapov (1976) derived a relation describing how the statistical properties of illumination fluctuations depend on average wave slopes. Stramski and Dera (1988) extended Schenck's gravity wave model to include shorter wavelength capillary waves. They also claimed that for practical purposes, only the statistical properties of the water surface can be derived from measured illumination fluctuation profiles. More recently, Zaneveld et al. (2001) estimated the diffuse attenuation coefficient of light intensity with depth, from real data. A study by Sabbah and Shashar (2006) considers how diffuse light sources, as opposed to direct sunlight, contribute to underwater light polarization and radiance fluctuations.

All previous simulations of underwater illumination patterns were performed by counting the ray density after ray tracing. This method suffers from instabilities caused by the finite number of rays used to calculate the densities. In this paper I formalize the ray-tracing method by approximating rays by delta distributions of illumination. These are extrapolated to depth and shifted, as a function of the angle of the ray, using the shift property of the delta function. I derive a close form expression of the illumination pattern in the spatial Fourier domain. I propose Gaussian distributions as an alternative to delta distributions for modeling incident light rays at the surface. They yield a more stable and smoother calculation in the space domain.

THEORY

First, I describe my theory for simulating underwater illumination patterns of sunlight. I define the Fourier transformation of a space-dependent function, $g = g(x)$ to be

$$\hat{g}(k) = \mathcal{F}\{g(x)\} = \int_{-\infty}^{\infty} g(x) \exp\{ikx\} dx, \quad (1)$$

where k is the wavenumber. The illumination as a function of distance x at a depth z is denoted as $f = f(x, z)$, with z positive downwards. I treat illumination at a point as being caused by the concentration of rays arriving from different directions.

Consider the special depth level, z_0 , at which the illumination at each point, $f_0(x) = f(x, z_0)$, is created by one light ray with a unique angular direction with respect to the vertical, $\theta = \theta(x)$. A straightforward derivation of the angle θ is given by Schenck (1957). Equation 2 gives the angle of a ray just under the water surface as a function of incidence angle $\theta^i = \theta^i(x)$, the function describing the water surface $S = S(x)$, and the indices of refraction of air and water, respectively η_a and η_w :

$$\theta(x) = \text{atan}\{\partial_x S(x)\} - \text{asin}\left[\eta_a \eta_w^{-1} \sin(\text{atan}\{\partial_x S(x)\} - \theta^i)\right]. \quad (2)$$

All angles are defined as positive clockwise. The intensity of sunlight, I_s , distributes over the surface as $f_0(x) = \sin\{\theta^i(x)\} I_s$. The illumination at depth z_0 is written as an integral of delta functions,

$$f_0(x) = f(x, z_0) = \int_{-\infty}^{\infty} f_0(r) \delta(x - r) dr. \quad (3)$$

Consider these delta functions as the contribution from each ray. Each ray is shifted to a different lateral position at greater depths; the shift is given by $\Delta z \tan\{\theta\}$, where $\Delta z = z - z_0$. In this way I write the illumination at a depth z as

$$f(x, z) = \int_{-\infty}^{\infty} f_0(r) \delta(x - [r + \Delta z \tan\theta(r)]) dr, \quad (4)$$

or, using equation 1,

$$\hat{f}(k, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_0(r) \delta(x - [r + \Delta z \tan\theta(r)]) \exp\{ikx\} dx dr \quad (5)$$

in the Fourier domain. Evaluating the integration over x modifies the argument of the exponent:

$$\hat{f}(k, z) = \int_{-\infty}^{\infty} f_0(r) \exp \{ik [r + \Delta z \tan \theta(r)]\} dr. \quad (6)$$

After inverse Fourier transformation, Equation 6 becomes

$$f(x, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_0(r) \exp \{ik [r + \Delta z \tan \theta(r)]\} \exp \{-ikx\} dr dk. \quad (7)$$

Calculation of the delta distribution in the wavenumber domain is unstable, because it would require infinite bandwidth. The result would suffer from the Gibbs phenomenon, as discussed in the next section. One can apply a smoothing filter in the wavenumber domain to suppress the worst of the ringing. An alternative approach is to replace the delta distribution with a different distribution. I propose a Gaussian distribution centered at $x = b$ of the following form:

$$\mathcal{G}(x, b) = \frac{1}{c\sqrt{2\pi}} \exp \left\{ -\frac{(x - b)^2}{2c^2} \right\}. \quad (8)$$

Note that the integral $\int_{-\infty}^{\infty} \mathcal{G}(x, b) dx = 1$. Smoothing the illumination at $z = z_0$ with a Gaussian distribution yields

$$f(x, z_0) = \int_{-\infty}^{\infty} f_0(r) \mathcal{G}(x, r) dr, \quad (9)$$

which projects to a depth z in a fashion similar to that in Equation 4,

$$f(x, z) = \int_{-\infty}^{\infty} f_0(r) \mathcal{G}(x, r + \Delta z \tan \theta(r)) dr. \quad (10)$$

Using Equation 10, this expands to

$$f(x, z) = \frac{1}{c\sqrt{2\pi}} \int_{-\infty}^{\infty} f_0(r) \exp \left\{ -\frac{(x - [r + \Delta z \tan \theta(r)])^2}{2c^2} \right\} dr. \quad (11)$$

In the next section I show examples of illumination patterns calculated using these formulas.

EXAMPLES

The shape of the water surface is modeled as a Gaussian function, similar to Equation 8 with $a = 5$, $b = 0$ and $c = 3$. This wavelet exhibits various focusing and bending aspects and is well suited to assess the accuracy to which the equations in the previous section can be computed. In addition, the sun is modeled as a point source at infinity, at an inclination of 85° to the horizon. Thus, the sunlight forms an incident plane wave at a 5° angle with the vertical. Let the intensity of the sun be $I_s = 1$, which yields an intensity of $f_0(x) = \cos 5^\circ$. Figure 1 shows a ray tracing example of the configuration, where the thickness of the water surface lens is not neglected. The illumination pattern at a depth of 75 m is calculated using Equation 7 and shown in the top panel of Figure 2. The Gibbs phenomenon is clearly visible and obscures the computed pattern. The real part of the spatial Fourier domain

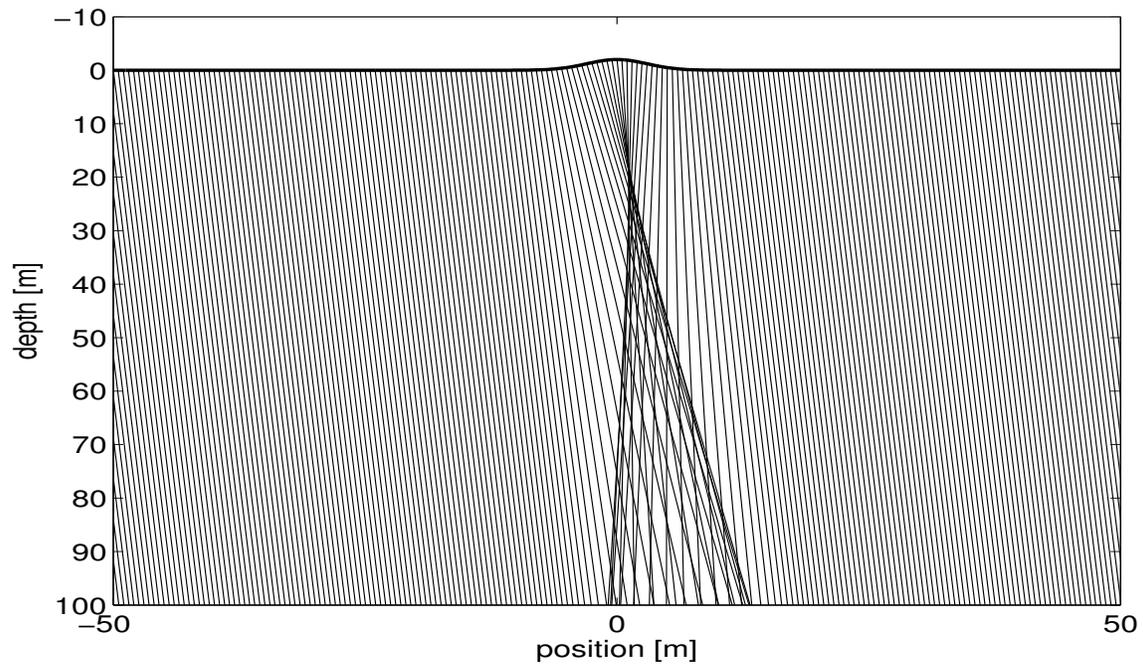


Figure 1: Ray tracing for sunlight passing through a Gaussian-shaped water wave. [ER] sjoerd2/. rays

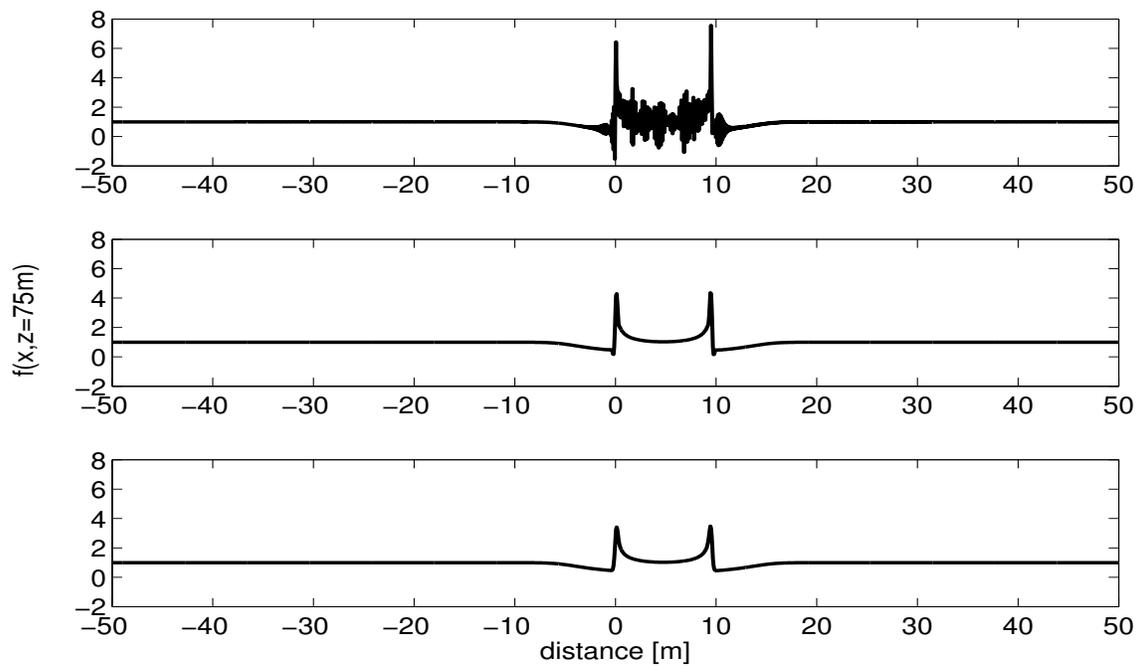


Figure 2: Various profiles at 75 meters deep. Top: rays modeled as delta distributions. Middle: Fourier-domain Hanning low-pass filtered version of the top panel. Bottom: rays approximated as Gaussian distributions. [ER] sjoerd2/. profiles

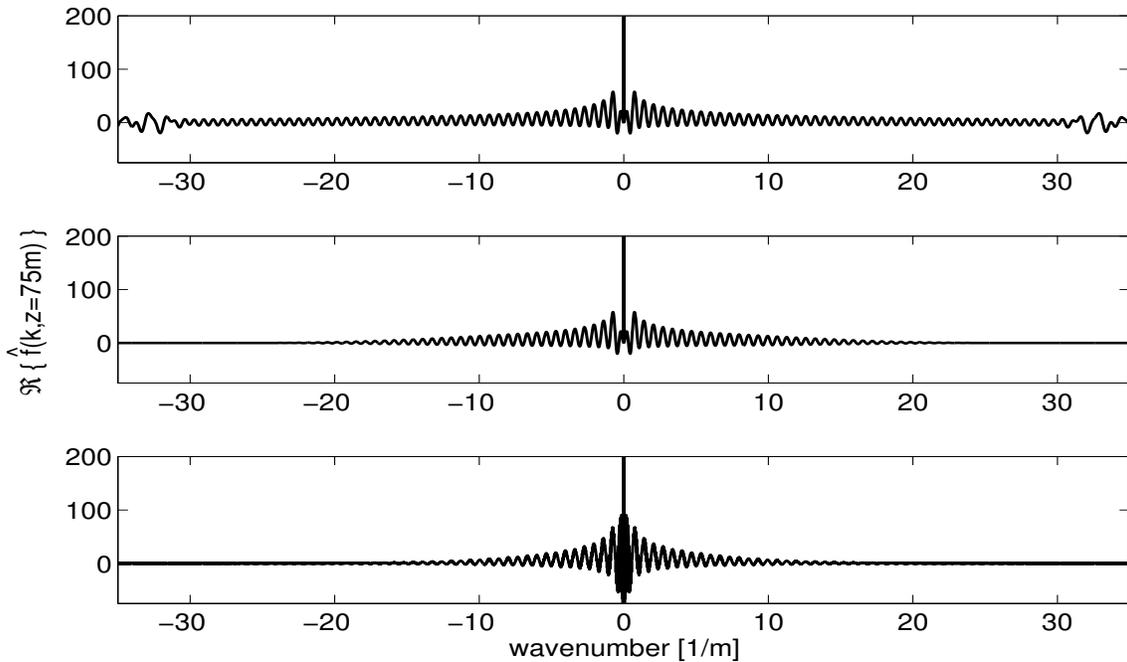


Figure 3: Real part of spatial Fourier domain spectrum for various profiles at 75 meters deep. Top: rays modeled as delta distributions. Middle: Fourier-domain Hanning low-pass filtered version of the top panel. Bottom: rays approximated as Gaussian distributions. [ER] sjoerd2/. filtering

spectrum is shown in the top panel of Figure 3. The Gibbs phenomenon is suppressed in the Fourier domain using a Hanning low-pass filter penalizing the spectrum above and under $k = \pm 18 \text{ m}^{-1}$. The filtered spectrum is shown in the middle panel of Figure 3. The obtained illumination pattern, seen in the middle panel of Figure 2, is now free of much of the ringing but still over- and under-shoots at the edges of the horn shape.

I use the Gaussian distributions to approximate the rays at the surface, and equation 11 to simulate the illumination pattern at a depth of 75 m. The resulting pattern is shown in the bottom panel of Figure 2. The real part of its spectrum is shown in Figure 3.

DISCUSSION AND CONCLUSIONS

The equations in this paper describe a new approach for modeling underwater illumination patterns of focused sunlight. The shift approach creates remarkably detailed results that are accurate in areas with low or high illumination intensity. Ray counting methods suffer in low coverage areas because of their discrete nature. Simulations using delta functions are unstable due to the Gibbs phenomenon, although applying a filter in the Fourier domain successfully improves the computed pattern. When we use a Gaussian distribution to approximate rays, there is no need to enter the Fourier domain. The standard deviation that modulates the width of the Gaussian should be matched to the sampling of the water surface function. I find that a standard deviation of $c = 2\Delta x$ generally produces good results. However, there is a limit on the second derivative of the function $\theta(x)$, for the ratio

of $c = 2\Delta x$ to remain valid.

Filtering is a processing step that needs adjusting, but so is the parameter c in the Gaussian distribution. My simulations become more stable but less accurate when a broader Gaussian is used to approximate rays. The mapping of $\theta(x)$ onto $f(x, z)$ is unique but non-linear. The map does not seem invertible, making processing by adjoint methods very difficult. Non-uniqueness is introduced by discretization of the presented integral equations. Posing a smoothness restriction on the function $\theta(x)$ will perhaps linearize the forward problem and reduce the number of non-unique solutions.

ACKNOWLEDGMENTS

The author would like to thank Jon Claerbout, Jeffrey Shragge, Abdullah Al Theyab and Ken Larner for helpful discussions and suggestions.

REFERENCES

- Claerbout, J. F., 2007, A tease for theoretical seismologists: WWW. (URL <http://sepwww.stanford.edu/sep/jon/poolwaves>).
- Dera, J. and H. R. Gordon, 1968, Light field fluctuations in the photic zone: *Limnol. Oceanogr.*, **13**, 697–699.
- Nikolayev, V. P. and M. S. Khulapov, 1976, Use of a nonstatistical model to explain the mechanism of underwater illumination fluctuations: *Izv. Acad. Sci. USSR Atmos. Oceanic Phys.*, **12**, 611–613.
- Sabbah, S. and N. Shashar, 2006, Underwater light polarization and radiance fluctuations induced by surface waves: *Applied Optics*, **45**, 4726–4739.
- Schenck, H., 1957, On the focussing of sunlight by ocean waves: *J. Opt. Soc. Am.*, **47**, 653–657.
- Snyder, R. L. and J. Dera, 1970, Wave-induced light-field fluctuations in the sea: *J. Opt. Soc. Am.*, **60**, 1072–1079.
- Stramski, D. and J. Dera, 1988, On the mechanism for producing flashing light under a wind-disturbed water surface: *Oceanologia*, **25**, 5–21.
- Zaneveld, J. R. V., E. Boss, and A. Barnard, 2001, Influence of surface waves on measured and modeled irradiance profiles: *Applied Optics*, **40**, 1442–1449.

Angle-domain common-image gathers in generalized coordinates

Jeff Shragge

ABSTRACT

The theory of angle-domain common-image gathers (ADCIGs) is extended to migrations performed in generalized coordinate systems and subsurface offset axes generated by nonlinear wavefield shifts. I develop an expression linking the definition of reflection opening angle to various geometric and nonlinear shifting factors. I demonstrate that, under certain circumstances, generalized coordinate ADCIGs can be calculated directly using Fourier-based offset-to-angle approaches. Cartesian and elliptic coordinate examples are given to validate the theory. A method for eliminating geometric factors from the ADCIG expression using judicious wavefield shifts is derived; however, this approach is not likely computationally advantageous in practice.

INTRODUCTION

Angle-domain common image gathers, or ADCIGs, are used increasingly in seismic imaging to examine migration velocity model accuracy (Biondi, 2006). The key idea is that migrating with correct velocity models leads to ADCIGs that do not shift vertically as a function of reflection opening angle (i.e., flat gathers). Migrating with incorrect velocity models, though, leads to inconsistent reflector depths and generates reflector smiles or frowns. ADCIGs are thus an effective analysis tool and have been incorporated in wave-equation-based inversion schemes to update velocity profiles (Sava and Biondi, 2004a,b).

ADCIGs can be generated during wave-equation imaging in a straightforward manner for both shot-profile (de Bruin et al., 1990; Sava and Fomel, 2003) and shot-geophone (Prucha et al., 1999; Mosher and Foster, 2000) migration approaches. In shot-profile migration, one can generate a subsurface offset axis at each depth step by correlating the source and receiver wavefields at a number of subsurface shifts. The second step involves computing an offset-to-angle domain transformation through a post-imaging, Fourier-based stretch (Sava and Fomel, 2003).

ADCIG theory is usually developed assuming horizontal and uniformly spaced wavefield shifts, largely because wavefield extrapolation and imaging are more commonly performed in Cartesian coordinates. The introduction of shot-profile migration in non-Cartesian coordinate systems (Shragge and Shan, 2008), though, warrants the development of a general ADCIG theory able to handle more arbitrary geometries and wavefield shifts. Generalizing ADCIG theory requires properly handling the effects of non-Cartesian geometry. For example, wavefield propagation in non-Cartesian coordinate systems induces local stretches, rotations and/or shearing of the wavenumbers (Shragge, 2008). Similarly, nonlinear and non-horizontal shifts can lead to angle-domain stretches.

The goal of this paper is to extend ADCIG theory to non-Cartesian geometries and nonlinear subsurface offset sampling. I demonstrate that ADCIG theory - as developed in a differential sense (Sava and Fomel, 2003) - remains valid for arbitrary geometry. Non-Cartesian coordinates do, though, introduce space-domain geometric factors that can render Fourier-based offset-to-angle methods unsuitable. I begin with a review of Cartesian ADCIG theory and provide an extension to generalized coordinate systems. I examine three canonical coordinate systems where the reflection angle can be explicitly calculated. I show how nonlinear shifting can modify the subsurface offset axis such that Fourier-based ADCIG calculation methods remain applicable.

ADCIG THEORY

The ADCIG theory presented in this section draws heavily from that presented in Sava and Fomel (2003). In the ensuing development, $\mathbf{x} = [x_1, x_3]$ denotes the Cartesian variables and $\boldsymbol{\xi} = [\xi_1, \xi_3]$ represents a generalized Riemannian coordinate system.

Cartesian Coordinates

For constant velocity media in Cartesian coordinates, a straightforward link exists between differential changes in the travel time, t , of rays connecting the source-reflector and reflector-receiver paths to changes in the subsurface offset, h_{x_1} , and depth, x_3 , coordinates. Figure 1a shows the geometry of the variables described above. Mathematically, these relationships

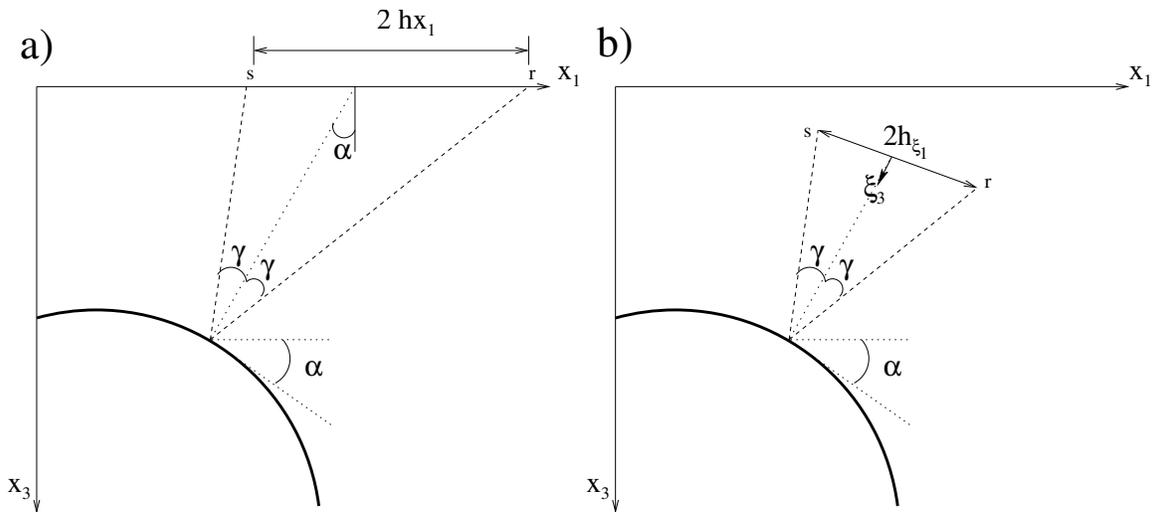


Figure 1: Zoomed in view illustrating reflection opening angle geometry. a) Cartesian coordinates. b) Generalized coordinates. Adapted from Sava and Fomel (2003). [NR]

jeff1/. rays

are expressed in the following two equations

$$\left[\begin{array}{c} \frac{\partial t}{\partial h_{x_1}} \\ \frac{\partial t}{\partial x_3} \end{array} \right] \Bigg|_{t, x_1} = 2 s \cos \alpha \begin{bmatrix} \sin \gamma \\ \cos \gamma \end{bmatrix}, \quad (1)$$

where s is slowness, α is reflector dip, and γ is the reflection opening angle. Using the implicit functions theory, equations 1 can be rewritten as

$$\left. \frac{\partial x_3}{\partial h_{x_1}} \right|_{t, x_1} = - \frac{\partial t}{\partial h_{x_1}} \bigg/ \frac{\partial t}{\partial x_3} = -\tan\gamma. \quad (2)$$

I introduce a negative sign in the right-hand-side of the equation above to be consistent with the notation of Sava and Fomel (2003). The left-hand side of equation 2 can be calculated in the frequency-wavenumber domain

$$\tan\gamma = - \frac{k_{h_{x_1}}}{k_{x_3}}, \quad (3)$$

where $k_{h_{x_1}}$ and k_{x_3} are the wavenumbers in the h_{x_1} and x_3 directions, respectively. Note that because equation 2 does not depend explicitly on \mathbf{x} , we may use Fourier-based methods to calculate the opening angle, γ , directly.

Generalized Coordinate Extension

Figure 1b illustrates a scenario similar to that illustrated in panel a, but for generalized coordinates. The reflection opening angle, γ , and the reflector dip, α , obviously remain unchanged; however, the orientations of the h_{ξ_1} and ξ_3 axes used to estimate γ now differ. The key question is which quantities in the ADCIG calculation are affected by this change of variables?

To answer these questions, I first assume that generalized coordinate systems are related to the Cartesian variables through a bijection (i.e., one-to-one mapping)

$$x_1 = f(\xi_1, \xi_3) \quad \text{and} \quad x_3 = g(\xi_1, \xi_3). \quad (4)$$

I also assert that the subsurface offset axes can be defined such that the implicit functions theory is valid

- h_{x_1} - a horizontal, but not necessarily linear, shift along the x_1 axis; and
- h_{ξ_1} - a shift in the direction of the ξ_1 axis commonly, though not always, sharing the same geometric qualities as the ξ_1 axis.

The bijection between the general and Cartesian coordinate systems allows us to rewrite equations 1 as

$$\left[\begin{array}{c} \frac{\partial t}{\partial h_{\xi_1}} \frac{\partial h_{\xi_1}}{\partial h_{x_1}} \\ \frac{\partial t}{\partial \xi_3} \frac{\partial \xi_3}{\partial x_3} \end{array} \right] \bigg|_{\xi_1, t} = 2s \cos\alpha \begin{bmatrix} \sin\gamma \\ \cos\gamma \end{bmatrix}. \quad (5)$$

Moving partial derivatives from the left side of the equation 5 to the right yields

$$\left[\begin{array}{c} \frac{\partial t}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial \xi_3} \end{array} \right] \bigg|_{\xi_1, t} = 2s \cos\alpha \begin{bmatrix} \frac{\partial h_{x_1}}{\partial h_{\xi_1}} \sin\gamma \\ \frac{\partial x_3}{\partial \xi_3} \cos\gamma \end{bmatrix}. \quad (6)$$

The generalized coordinate ADCIG is given by the division by the two expressions in equation 6

$$\tan\gamma = - \left. \frac{\partial\xi_3}{\partial h_{\xi_1}} \right|_{\xi_1,t} \left[\frac{\partial x_3}{\partial\xi_3} \middle/ \frac{\partial h_{x_1}}{\partial h_{\xi_1}} \right]. \quad (7)$$

The bracketed terms generally introduce a geometric dependence of the ADCIGs on coordinates ξ , which can preclude the use of Fourier-based methods for calculating ADCIGs. The expression $|_{\xi_1,t}$ will be implicitly assumed for the remainder of the paper.

Defining subsurface stretch

One question arising from the geometric factors in equation 7 is what does the term $\frac{\partial h_{x_1}}{\partial h_{\xi_1}}$ represent? One way to evaluate this quantity is use a partial derivative expansion

$$\frac{\partial h_{x_1}}{\partial h_{\xi_1}} = \frac{\partial h_{x_1}}{\partial x_1} \frac{\partial x_1}{\partial \xi_1} \middle/ \frac{\partial h_{\xi_1}}{\partial \xi_1}, \quad (8)$$

to isolate three separate terms. I interpret each contribution in the following way:

- $\frac{\partial h_{x_1}}{\partial x_1}$ - a scale factor of the transformation between h_{x_1} and x_1 usually given by $h_{x_1} = x_1$ such that $\frac{\partial h_{x_1}}{\partial x_1} = 1$;
- $\frac{\partial x_1}{\partial \xi_1}$ - the partial derivative mapping between the two coordinate systems derivable from equations 4; and
- $\frac{\partial h_{\xi_1}}{\partial \xi_1}$ a scale factor of the transformation between h_{ξ_1} and ξ_1 normally defined as $h_{\xi_1} = \xi_1$ such that $\frac{\partial h_{\xi_1}}{\partial \xi_1} = 1$, but permitted to have a parametric form $h_{\xi_1} = A(\xi_1)$.

Using the above partial derivative expansion allows us to write a general ADCIG relationship

$$\tan\gamma = - \frac{\partial\xi_3}{\partial h_{\xi_1}} \left[\left(\frac{\partial x_3}{\partial\xi_3} \frac{\partial h_{\xi_1}}{\partial \xi_1} \right) \middle/ \left(\frac{\partial h_{x_1}}{\partial x_1} \frac{\partial x_1}{\partial \xi_1} \right) \right]. \quad (9)$$

In the examples below, unless stated otherwise, I assume regular sampling with linear wavefield shifting such that the following hold,

$$\begin{bmatrix} h_{x_1} \\ h_{\xi_1} \end{bmatrix} = \begin{bmatrix} x_1 \\ \xi_3 \end{bmatrix} \quad \text{or} \quad \frac{\partial h_{x_1}}{\partial x_1} = \frac{\partial h_{\xi_1}}{\partial \xi_1} = 1, \quad (10)$$

which reduces the complexity of the general coordinate ADCIG expressions to

$$\tan\gamma = - \frac{\partial\xi_3}{\partial h_{\xi_1}} \left[\frac{\partial x_3}{\partial\xi_3} \middle/ \frac{\partial x_1}{\partial \xi_1} \right]. \quad (11)$$

CANONICAL EXAMPLES

This section presents three canonical examples that illustrate the generalized ADCIG theory: Cartesian, sheared Cartesian and elliptic coordinate systems.

Cartesian Coordinates

A Cartesian coordinate system \mathbf{x} can be defined from a unit square ξ by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_3 \end{bmatrix}. \quad (12)$$

The partial differential transformation matrix is

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad (13)$$

leading to the following differential travel-time equations

$$\begin{bmatrix} \frac{\partial t}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial \xi_3} \end{bmatrix} = 2s \cos \alpha \begin{bmatrix} a \sin \gamma \\ b \cos \gamma \end{bmatrix}. \quad (14)$$

The Cartesian ADCIG computation is given by

$$\tan \gamma = -\frac{b}{a} \frac{\partial \xi_3}{\partial h_{\xi_1}}. \quad (15)$$

Note that where the axes are equally sampled, one recovers the correct reflection opening angle; situations where the axes are not equally sampled require an additional scaling. This stretch is usually taken into account during the Fourier transformation implicit in equation 7.

Sheared Cartesian Coordinates

A sheared Cartesian coordinate system (see Figure 2) is an instructional, though impractical, generalized coordinate system for shot-profile migration. A sheared Cartesian mesh is defined by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & \sin \theta \\ 0 & \cos \theta \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_3 \end{bmatrix}, \quad (16)$$

where θ is the shearing angle. The transformation matrix is

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} 1 & \sin \theta \\ 0 & \cos \theta \end{bmatrix}, \quad (17)$$

which leads to the following differential travel-time equations

$$\begin{bmatrix} \frac{\partial t}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial \xi_3} \end{bmatrix} = 2s \cos \alpha \begin{bmatrix} \sin \gamma \\ \cos \theta \cos \gamma \end{bmatrix}. \quad (18)$$

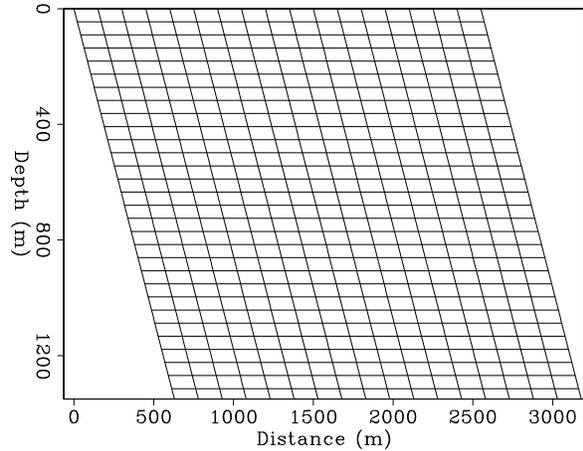
The computation for ADCIGs in sheared Cartesian coordinates is

$$\tan \gamma = -\cos \theta \frac{\partial \xi_3}{\partial h_{\xi_1}}. \quad (19)$$

From equation 19, we can see that the apparent dips must be filtered initially by $\cos \theta$ in order to recover the true reflection opening angle.

Figure 2: Example of a sheared Cartesian coordinate system with a shear angle of 25° .[\[NR\]](#)

`jeff1/. Sheared`



Elliptic Coordinates

The elliptic coordinate system (see figure 3) is defined by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \cosh \xi_3 \cos \xi_1 \\ a \sinh \xi_3 \sin \xi_1 \end{bmatrix}. \quad (20)$$

The transformation matrix is defined by

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = a \begin{bmatrix} \cosh \xi_3 \sin \xi_1 & \sinh \xi_3 \cos \xi_1 \\ -\sinh \xi_3 \cos \xi_1 & \cosh \xi_3 \sin \xi_1 \end{bmatrix}, \quad (21)$$

which leads to the following differential travel-time equations

$$\begin{bmatrix} \frac{\partial t}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial \xi_3} \end{bmatrix} = 2s \cos \alpha \begin{bmatrix} \sin \gamma (a \cosh \xi_3 \sin \xi_1) \\ \cos \gamma (a \cosh \xi_3 \sin \xi_1) \end{bmatrix}. \quad (22)$$

The computation for the ADCIG in elliptic coordinates is given by

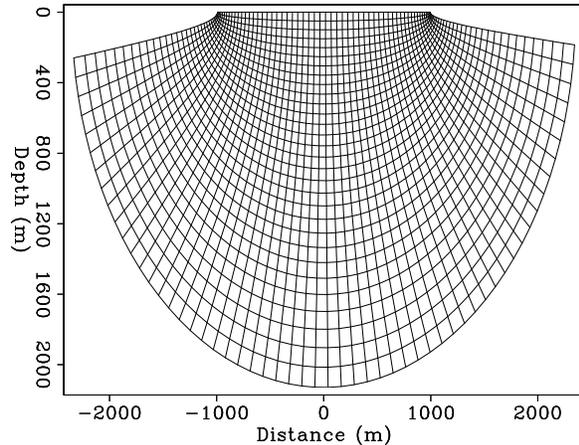
$$\tan \gamma = -\frac{\partial \xi_3}{\partial h_{\xi_1}}. \quad (23)$$

Thus, ADCIGs calculated in elliptic coordinates directly yield the reflection opening angle without any additional filtering.

NUMERICAL EXAMPLES

This section presents a numerical test of the generalized theory by comparing ADCIGs from Cartesian and elliptic coordinate systems for the BP synthetic velocity model. Shragge and Shan (2008) demonstrate that the elliptic coordinate system does not induce an anisotropic wavenumber stretch during wavefield extrapolation. I assert that equation 23 is a similar expression, and is a further argument that the Cartesian and elliptic coordinate ADCIGs should yield similar results. (The results are not necessarily equal due to the differing wavefield extrapolation accuracy.)

Figure 3: Example of an elliptic coordinate system. [NR] jeff1/. EC



The numerical examples presented herein were generated using a shot-profile migration algorithm with extrapolation operators accurate to roughly 80° (Lee and Suh, 1985). For each profile I did the following: i) computed 31 subsurface shifts at each extrapolation step; ii) calculated ADCIGs using the procedure described in Sava and Fomel (2003); and iii) interpolated the single-shot ADCIG output to the global image volume.

The top and bottom panels of figure 4 show the image volumes for the elliptic and Cartesian coordinate systems, respectively. I indicate a number of locations where the elliptic coordinate system produces superior images. Figure 5 shows the ADCIGs corresponding to figure 4 for the elliptic (top panel) and Cartesian (bottom panel) coordinate systems. The ADCIGs are spaced out every 500 meters, and have an angular bandwidth of $[-60^\circ < \gamma < 60^\circ]$. Note that the ADCIGs are flat, though with slightly different amplitudes caused by differing illumination. The similarity between these gathers indicates the validity of the general coordinate ADCIG theory.

A second test that illustrates the validity of this approach is to examine how the ADCIGs change when the velocity profile is altered. For this test, we rescale the BP synthetic velocity profile by factors from 0.92x to 1.08x in increments of 0.02x and migrate a single shot-profile. Figure 6 presents the elliptic coordinate ADCIG results for an ADCIG and shot point coincidentally located at 12000 m. As we progress from the leftmost (too slow) to the rightmost (too fast) panels, we observe that the imaged reflections go from smiling to frowning. As expected, the ADCIG is most flat and well-focused where the true velocity model is used.

ELIMINATING SPATIAL DEPENDENCY

This section presents a method for eliminating the spatial dependency of generalized ADCIG transforms through a judicious stretching of the subsurface offset axis. Revisiting how one can introduce a stretch using equation 11, one obvious restriction to make is maintaining uniform Cartesian spatial sampling. I accomplish this by enforcing $\frac{\partial h_{x_1}}{\partial x_1} = 1$ such that

$$\tan\gamma = -\frac{\partial\xi_3}{\partial h_{\xi_1}} \left[\frac{\partial h_{\xi_1}}{\partial \xi_1} \frac{\partial x_3}{\partial \xi_3} \bigg/ \frac{\partial x_1}{\partial \xi_1} \right]. \quad (24)$$

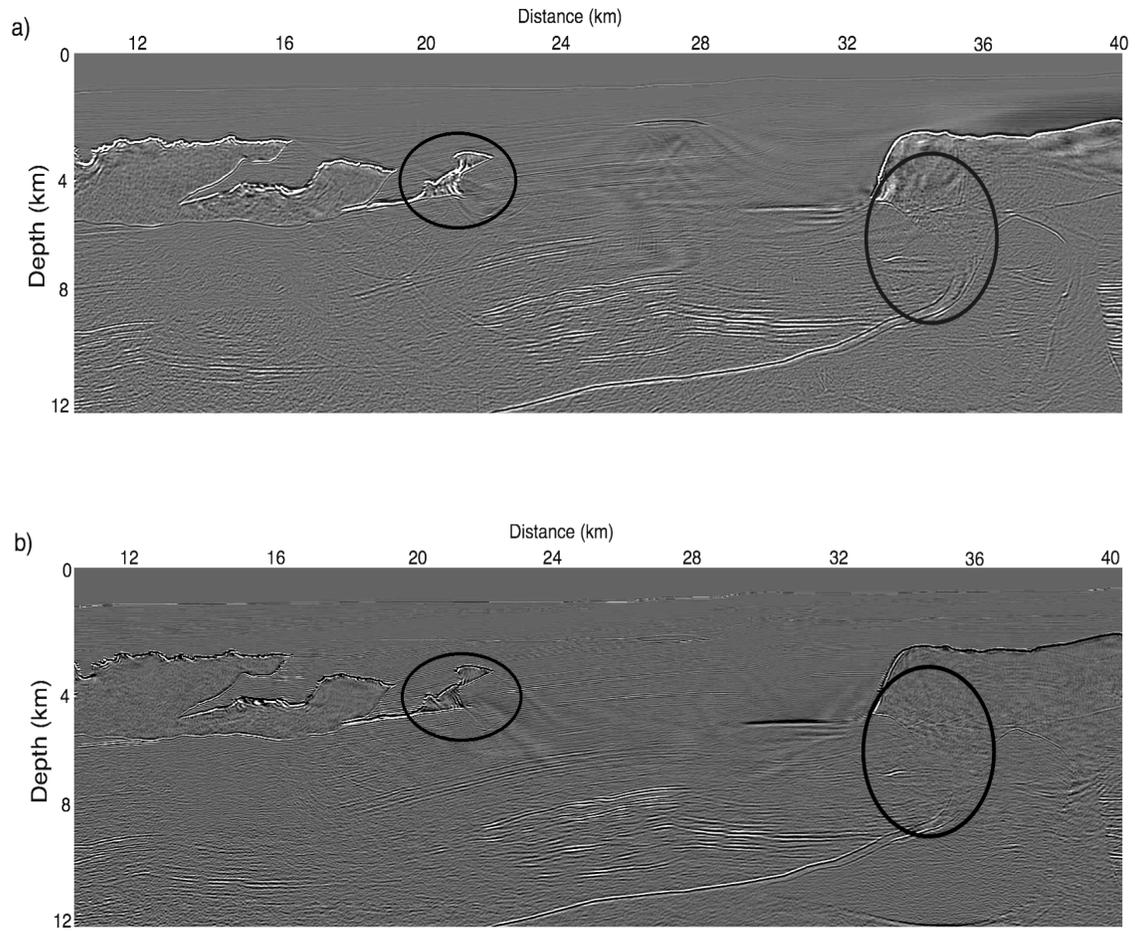


Figure 4: Comparative BP velocity model images for the elliptic (top panel) and Cartesian (bottom panel) coordinate systems. [CR] [jeff1/. Images](#)



Figure 5: ADCIGs corresponding to the images in figure 4 calculated in the elliptic coordinates. [CR] jeff1/. EllipticADCIG

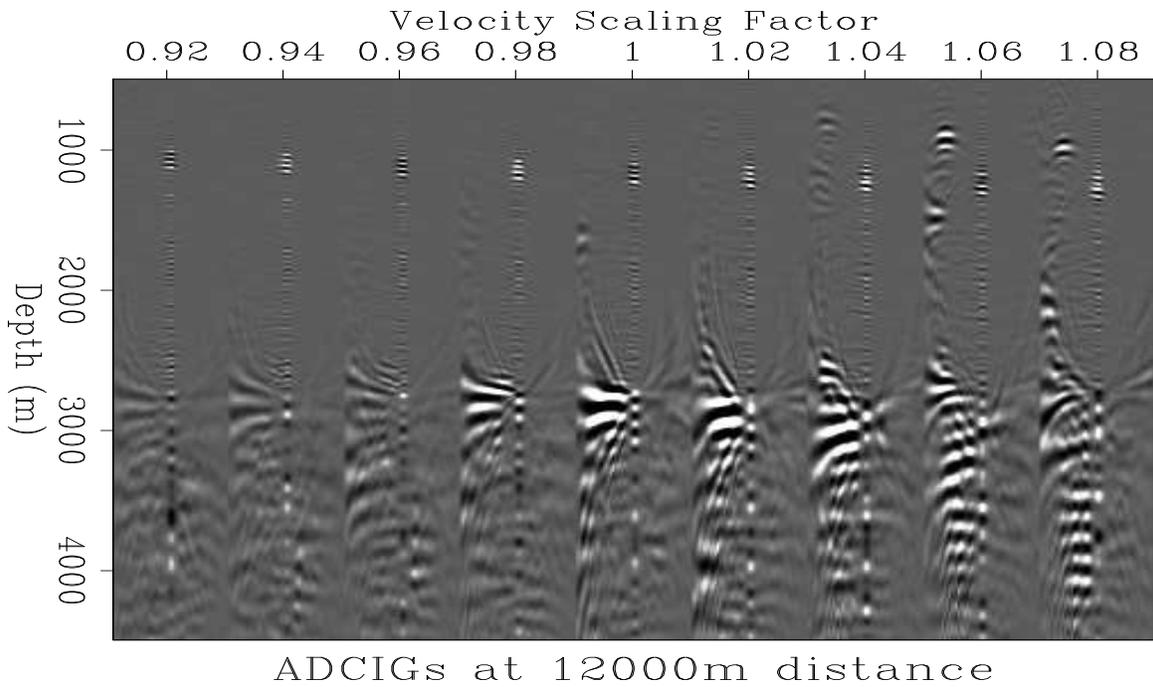


Figure 6: Single shot-profile migration ADCIGs for a coincident ADCIG and source point at 12000 m. Note that the image is best focused when the correct velocity is used, and frowns and smiles are observed when migration velocity is used. [ER] jeff1/. Nice

The next step is to specify the relationship between h_{ξ_1} and ξ_1 that enables us to calculate $\frac{\partial h_{\xi_1}}{\partial \xi_1}$. One useful *ansatz* solution is

$$\frac{\partial h_{\xi_1}}{\partial \xi_1} = \frac{\partial x_1}{\partial \xi_1} / \frac{\partial x_3}{\partial \xi_3}. \quad (25)$$

Substituting equation 25 into equation 24 generates the following ADCIG

$$\tan \gamma = -\frac{\partial \xi_3}{\partial h_{\xi_1}}. \quad (26)$$

Equation 25 implies that if we can define an appropriate coordinate system stretch for each ξ location, we may still recover the reflection opening angle using Fourier-based techniques.

Polar coordinate example

The polar coordinate system, where the extrapolation direction is oriented in the angular rather than the radial direction (see figure 7), is defined by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} a \xi_1 \cos \xi_3 \\ a \xi_1 \sin \xi_3 \end{bmatrix}. \quad (27)$$

The partial derivative transformation matrix between the two systems is

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} a \cos \xi_3 & -a \xi_1 \sin \xi_3 \\ a \sin \xi_3 & a \xi_1 \cos \xi_3 \end{bmatrix}, \quad (28)$$

leading to the following differential travel-time equations

$$\begin{bmatrix} \frac{\partial t}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial \xi_3} \end{bmatrix} = 2 s \cos \alpha \begin{bmatrix} a \cos \xi_3 \sin \gamma \\ a \xi_1 \cos \xi_3 \cos \gamma \end{bmatrix}. \quad (29)$$

Inserting equations 28 into equation 24 generates the expression for polar coordinate ADCIGs

$$\tan \gamma = \frac{\partial \xi_3}{\partial h_{\xi_1}} \left[\xi_1 \frac{\partial h_{\xi_1}}{\partial \xi_1} \right]. \quad (30)$$

Thus, one cannot calculate ADCIGs directly in a polar coordinate system unless the spatial dependency is judiciously eliminated.

The polar coordinate system provides an example where ADCIGs contain a geometric dependence on ξ . Inserting the geometric factors $\frac{\partial x_1}{\partial \xi_1}$ and $\frac{\partial x_3}{\partial \xi_3}$ from above into equation 25 leads to

$$\frac{\partial h_{\xi_1}}{\partial \xi_1} = \frac{\partial x_1}{\partial \xi_1} / \frac{\partial x_3}{\partial \xi_3} = \frac{1}{\xi_1}. \quad (31)$$

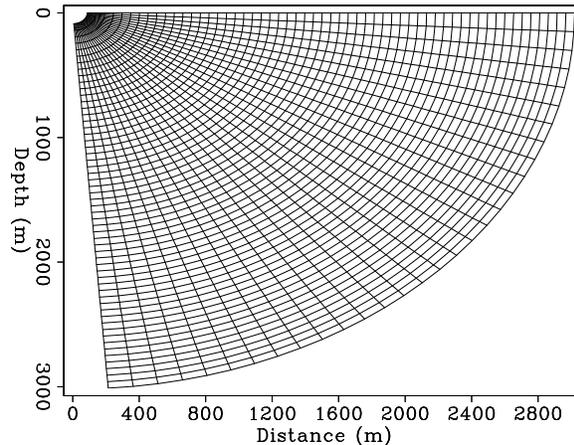
Integrating along surfaces of constant ξ_3 yields

$$h_{\xi_1} = \ln \xi_1. \quad (32)$$

Equation 32 defines the subsurface axis stretch required to directly calculate ADCIGs by Fourier-based approaches.

One question is how best to perform this stretch. One approach would be to perform linear shifting and then regrid that result to an natural log grid. However, the computational overhead renders this method less-than-ideal, especially for situations where estimating $\frac{\partial h_{\xi_1}}{\partial x_3}$ directly by slant-stack processing is more efficient. However, this remains an open research topic.

Figure 7: Example of a polar coordinate system. [NR] `jeff1/. PC`



CONCLUSIONS

This paper extends the Cartesian ADCIG theory to generalized coordinate systems. The expressions for ADCIGs contain additional coordinate factors describing mesh geometry and wavefield shifts. The method for calculating ADCIGs in an elliptic coordinate system is identical to a Cartesian ADCIG calculation as confirmed by numerical examples. An approach for eliminating the spatial dependency is given; however, this approach is not likely efficient in practice and slant stack calculations will be used as an alternative.

ACKNOWLEDGMENTS

I would like to thank Alejandro Valenciano for helpful discussions, and BP for releasing the synthetic data used in the examples.

REFERENCES

- Biondi, B., 2006, 3D Seismic Imaging: Society of Exploration Geophysicists.
- de Bruin, C., C. Wapenaar, and A. Berkhout, 1990, Angle-dependent reflectivity by means of prestack migration: *Geophysics*, **55**, 1223–1234.
- Lee, M. W. and S. Y. Suh, 1985, Optimization of one-way wave-equations (short note): *Geophysics*, **50**, 1634–1637.
- Mosher, C. and D. Foster, 2000, Common angle imaging conditions for prestack depth migration: 70th Ann. Internat. Mtg, 830–833, Soc. of Expl. Geophys.
- Prucha, M., B. Biondi, and W. Symes, 1999, Angle-domain common image gathers by wave-equation migration: 69th Ann. Internat. Mtg, 824–827, Soc. of Expl. Geophys.

- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis. I. Theory: *Geophys. Prosp.*, **52**, 593–606.
- , 2004b, Wave-equation migration velocity analysis. II. Subsalt imaging examples: *Geophys. Prosp.*, **52**, 607–623.
- Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: *Geophysics*, **68**, 1065–1074.
- Shragge, J., 2008, Riemannian wavefield extrapolation: Nonorthogonal coordinate systems: *Geophysics*, **73**, T11–T21.
- Shragge, J. and G. Shan, 2008, 2D prestack depth migration in elliptic coordinates: accepted in *Geophysics*.

3D shot-profile migration in ellipsoidal coordinates

Jeff Shragge and Guojian Shan

ABSTRACT

We present an approach for performing 3D shot-profile migration in ellipsoidal coordinate systems. Wavefields are extrapolated on confocal ellipsoidal shells that are well suited for accurately propagating steeply dipping and turning waves in all azimuthal directions. Numerical implementation of the corresponding dispersion relationship, though, is somewhat problematic due to first-order, complex-valued wavenumbers. We show that an integral transform recasts the problem in a way that eliminates first-order wavenumbers. The corresponding dispersion relationship is similar to that in elliptically anisotropic media. This similarity allows us to use existing implementations of wavefield extrapolation in elliptically anisotropic media to propagate wavefields on ellipsoidal meshes. Impulse response tests demonstrate the stability and accuracy of the approach.

INTRODUCTION

Imaging steeply dipping structure, such as salt flanks, in complex geologic environments remains an issue in 3D seismic migration. Many difficulties arise due to the related problems of incorrect velocity models and poor downgoing illumination. One strategy is to use turning-wave energy arriving from all azimuths to image the steep dip structure that remains unilluminated by downgoing energy alone. A number of studies have examined this approach with varying degrees of success (Hale et al., 1992; Zhang and McMechan, 1997; Zhang and Xu, 2006).

The practical imaging improvements afforded by imaging turning-wave energy, though, are unavoidably linked to data acquisition geometry and velocity model structure. For example, narrow-azimuth migrations of data sets containing predominantly inline turning-wave energy are usually constrained to have minimal crossline aperture. This restriction precludes imaging turning-wave energy originating from crossline structure arriving at near-zero offsets. The increasing popularity of wide-azimuth acquisition - well suited for recording turning-wave energy from all azimuths - suggests the need for improved wide-aperture wavefield propagation techniques. Determining which seismic imaging methods optimally realize these requirements, both physically and computationally, remains an open research question.

Wave-equation migration (WEM) techniques generally generate superior images relative to other approaches in complex geologic environments. Reverse-time migration, which solves the full acoustic wave equation, is one class of WEM techniques able to propagate turning waves in all directions (Baysal et al., 1983; Biondi and Shan, 2002). The computational costs associated with 3D wave propagation and imaging remain significant, though, especially in wide-azimuth contexts. A second class of WEM approaches, based on one-way

wavefield extrapolation, rapidly generates solutions to approximate one-way wave equations. The computational advantages of one-way wavefield extrapolation, relative to reverse-time migration, are less obvious when considering the lower accuracy of high-angle propagation and an inherent inability to propagate turning waves by design.

Formulating the seismic imaging problem in more generalized coordinate systems is one way to exploit the computational advantages of one-way wavefield extrapolation while reducing the steep-dip limitations. The general strategy involves extrapolating source and receiver wavefields on meshes oriented toward the wave-propagation direction, generating local images through cross-correlation in the transformed coordinate system, and interpolating the result back to the global image volume. This process is repeated for all source-receiver wavefield pairs. Coordinate transforms proved to be an effective strategy for 2D and 3D plane-wave migration when using tilted Cartesian coordinate systems oriented toward the take-off vector (Shan and Biondi, 2004, 2006). Shragge and Shan (2008) apply this strategy in developing a 2D shot-profile migration in confocal elliptic coordinates.

In this paper, we apply the coordinate transform strategy to 3D ellipsoidal meshes. We extrapolate source and receiver wavefields outward in confocal ellipsoidal shells and perform cross-correlations to form local images. In most circumstances the wave-propagation direction is conformal to the ellipsoidal shells, permitting the imaging of turning waves with one-way operators. A second advantage of ellipsoidal coordinate systems is that the inline/crossline aspect ratio is controlled by a single parameter, and meshes can be rescaled to fit either narrow or wide-azimuth geometries. Ellipsoidal coordinates also can be defined using integral transforms that leave the corresponding dispersion relationship no more complicated than that of elliptically anisotropic media. Wavefield extrapolation is thus achieved using numerical approaches similar to optimized elliptically anisotropic finite-difference extrapolation (Shan and Biondi, 2005).

The paper begins with a discussion regarding two definitions of an ellipsoidal coordinate system. We generate the corresponding extrapolation wavenumber appropriate for performing wavefield continuation in ellipsoidal coordinates. We then discuss the implicit 3D finite-difference implementation used to propagate source and receiver wavefields, and give the results of impulse response tests. The paper ends with discussions on future work and the computational overhead associated with performing shot-profile migration in ellipsoidal coordinates.

ELLIPSOIDAL GEOMETRY

A common definition of an ellipsoidal coordinate system denoted by $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3]$ relative to a Cartesian mesh given by $\boldsymbol{x} = [x_1, x_2, x_3]$ is (Morse and Feshbach, 1953):

$$\begin{aligned} x_1^2 &= \frac{(a^2 + \xi_1^2)(a^2 + \xi_2^2)(a^2 + \xi_3^2)}{(b^2 - a^2)(c^2 - a^2)}, \\ x_2^2 &= \frac{(b^2 + \xi_1^2)(b^2 + \xi_2^2)(b^2 + \xi_3^2)}{(a^2 - b^2)(c^2 - b^2)}, \\ x_3^2 &= \frac{(c^2 + \xi_1^2)(c^2 + \xi_2^2)(c^2 + \xi_3^2)}{(a^2 - c^2)(b^2 - c^2)}. \end{aligned} \tag{1}$$

Parameters a , b , and c are constants defining the coordinate system ellipticity and are constrained by three inequalities: $a^2 < \xi_1^2 < b^2$, $b^2 < \xi_2^2 < c^2$, and $c^2 < \xi_3^2 < \infty$. Surfaces of constant ξ_3 are confocal ellipsoids and represent the direction of wavefield extrapolation,

while those of constant ξ_1 and ξ_2 form two- and one-sheeted hyperboloids, respectively. We note that because equations 1 are defined by the squares of the variables, each Cartesian point \mathbf{x} is represented by eight ellipsoidal points $\boldsymbol{\xi}$, one located in each octant.

Figures 1 and 2 present two ellipsoidal coordinate examples. In each figure, we infilled the four octants with positive ξ_3 arguments to form a coordinate system appropriate for performing 3D wavefield extrapolation. The difference between the two coordinate systems is controlled by parameter b , where decreasing b leads to a more spherical mesh. Note that in this coordinate system waves can propagate in all azimuthal directions, and usually at low angles to the extrapolation direction in typical Gulf of Mexico velocity profiles.

Figure 1: Example of an ellipsoidal coordinate system conforming to narrow-azimuth acquisition geometry created with the parameters $[a, b, c] = [0, 0.995, 1]$. The figure shows five confocal shells. [NR] `jeff2/. NarrowAzimuth`

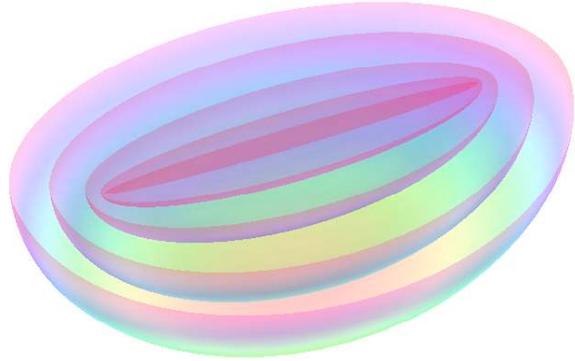
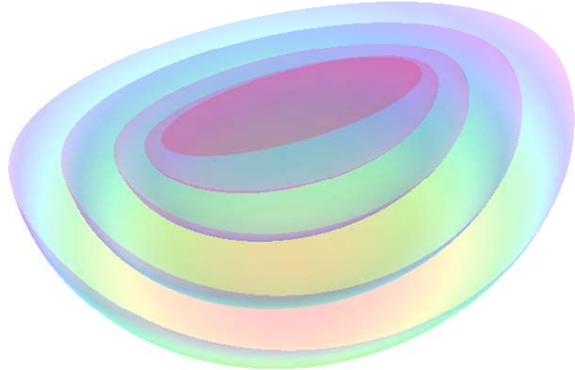


Figure 2: Example of an ellipsoidal coordinate system conforming to wide-azimuth acquisition geometry created with the parameters $[a, b, c] = [0, 0.925, 1]$. The figure shows five confocal shells. [NR] `jeff2/. WideAzimuth`



Morse and Feshbach (1953) define the elliptic-coordinate Helmholtz equation as

$$\begin{aligned} \nabla^2 U = & (\xi_3^2 - \xi_2^2) S(\xi_1) \frac{\partial}{\partial \xi_1} \left[S(\xi_1) \frac{\partial}{\partial \xi_1} \right] U + \\ & (\xi_1^2 - \xi_3^2) S(\xi_2) \frac{\partial}{\partial \xi_2} \left[S(\xi_2) \frac{\partial}{\partial \xi_2} \right] U + \\ & (\xi_1^2 - \xi_2^2) S(\xi_3) \frac{\partial}{\partial \xi_3} \left[S(\xi_3) \frac{\partial}{\partial \xi_3} \right] U = -\omega^2 s^2 U, \end{aligned} \quad (2)$$

where U is a wavefield, ∇^2 is the Laplacian operator, ω is angular frequency, s is slowness (reciprocal of velocity), and $S(\sigma)$ is a general parameter defined by

$$S(\sigma) = \sqrt{(\sigma^2 + a^2)(\sigma^2 + b^2)(\sigma^2 + c^2)}. \quad (3)$$

Equation 2 contains partial derivatives with respect to the parameter S . This leads to a dispersion relationship of the type studied in Shragge (2008),

$$k_{\xi_3} = ia_3 \pm \sqrt{a_4^2 \omega^2 s^2 - a_5^2 k_{\xi_1}^2 - a_6^2 k_{\xi_2}^2 + ia_8 k_{\xi_1} + ia_9 k_{\xi_2} - a_{10}^2}, \quad (4)$$

where a_i are geometric coefficients, $i = \sqrt{-1}$ and k_{ξ_j} is the wavenumber in the corresponding j^{th} axis.

Overall, the ellipsoidal coordinate system as defined in equation 1 is well suited to 3D shot-profile migration in a geometric sense; however, two issues make it difficult to implement accurately. First, the dispersion relationship in equation 4 does not easily lend itself to implicit finite-difference methods because of the imaginary first-order terms (e.g., $ia_8 k_{\xi_1}$). Second, the octant-based definition introduces non-uniqueness to the coordinate system variables. Overall, making an ellipsoidal coordinate system practical for wavefield extrapolation will require an alternate definition that overcomes these two issues.

Integral Confocal Ellipsoidal Equations

A second definition for confocal ellipsoidal coordinates uses auxiliary variables defined through integral transforms. Byerly (1959) defines the following Jacobi elliptic integral transforms for each coordinate system axis:

$$\begin{aligned} \beta &= \int_b^{\xi_1} \frac{c d\xi_1}{(c^2 - \xi_1^2)(\xi^2 - c^2)}, \\ \gamma &= \int_0^{\xi_2} \frac{c d\xi_2}{(b^2 - \xi_2^2)(c^2 - \xi^2)}, \\ \alpha &= \int_c^{\xi_3} \frac{c d\xi_3}{(\xi_2^2 - b^2)(\xi^2 - c^2)}, \end{aligned} \quad (5)$$

where axes $[\beta, \gamma, \alpha]$ are conformal to the $[\xi_1, \xi_2, \xi_3]$ axes, but are stretched by the integral transforms defined in equation 5. Axes $[\beta, \gamma, \alpha]$ are defined on the following ranges: $0 \leq \xi_1 \leq \infty$, $0 \leq \xi_2 \leq \infty$ and $0 \leq \xi_3 \leq \infty$. Additional information on the integral transforms can be found in Appendix A. Figure 3 illustrates the relative stretching for each axis for the wide-azimuth geometry case presented in figure 2.

The integral ellipsoidal-coordinate Helmholtz equation is (Byerly, 1959):

$$\nabla^2 U = (\xi_2^2 - \xi_3^2) \frac{\partial^2}{\partial \beta^2} U + (\xi_1^2 - \xi_3^2) \frac{\partial^2}{\partial \gamma^2} U + (\xi_1^2 - \xi_2^2) \frac{\partial^2}{\partial \alpha^2} U = -\omega^2 s^2 U. \quad (6)$$

Note that this definition effectively rescales the ξ coordinate axes to eliminate the first-order partial-differential terms in equation 2. This represents the most important theoretical result in this paper, as it removes the main implementation difficulty. In addition, integral ellipsoidal coordinates are defined globally, not just in octants, which eliminates the non-uniqueness noted above.

Obtaining a dispersion relationship from the expression in equation 6 is fairly straightforward. Replacing the partial differential terms with their Fourier-domain counterparts (i.e. $\frac{\partial}{\partial j} \leftrightarrow -ik_j$, for $j = \alpha, \beta, \gamma$) and solving for the extrapolation direction wavenumber k_α yields

$$k_\alpha = \sqrt{A^2 \omega^2 s^2 - B^2 k_\beta^2 - C^2 k_\gamma^2}, \quad (7)$$

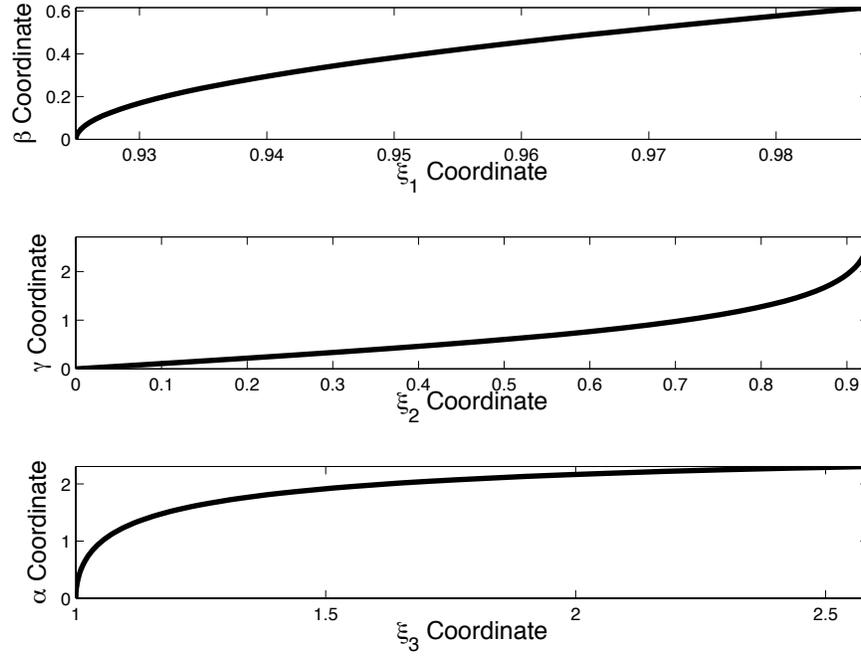


Figure 3: Integral transform stretches for the ξ axes given by equation 5. Top panel: $\beta - \xi_1$ coordinate stretch. Middle panel: $\gamma - \xi_2$ coordinate stretch. Bottom panel: $\alpha - \xi_3$ coordinate stretch. [NR] jeff2/. IntegralTransform

where

$$A = \frac{1}{\sqrt{\xi_1^2 - \xi_2^2}}, \quad B = \sqrt{\frac{\xi_2^2 - \xi_3^2}{\xi_1^2 - \xi_2^2}}, \quad \text{and} \quad C = \sqrt{\frac{\xi_1^2 - \xi_3^2}{\xi_1^2 - \xi_2^2}}. \quad (8)$$

In general, equation 7 is inexact because the A, B and C coefficients in equations 8 (and possibly slowness) vary spatially along the β and γ axes.

Relationship to elliptically anisotropic media

A naturally arising concern is whether the dispersion relationship in equation 7 can be implemented accurately and efficiently in a wavefield extrapolation scheme. We address this question by comparing the ellipsoidal and the (Cartesian-coordinate) elliptically anisotropic media dispersion relationships. By defining an effective slowness $s_A = As$ and rewriting equation 7 as

$$\frac{k_\alpha}{\omega s_A} = \sqrt{1 - B^2 \frac{k_\beta^2}{\omega^2 s_A^2} - C^2 \frac{k_\gamma^2}{\omega^2 s_A^2}}, \quad (9)$$

we observe that the ellipsoidal coordinate dispersion relationship resembles that for elliptically anisotropic media (Tsvankin, 1996). Specifically, ellipsoidal coordinates relate to the

Coeff. order i	Coeff. a_i	Coeff. b_i
1	0.040315157	0.873981642
2	0.457289566	0.222691983

Table 1: Coefficients used in 3D implicit finite-difference wavefield extrapolation.

case where Thomsen parameters (Thomsen, 1986) obey $\epsilon = \delta$

$$\left. \frac{k_{x_3}}{\omega s} \right|_{\epsilon=\delta} = \sqrt{\frac{1 - (1 + 2\epsilon) \frac{k_{x_1}^2 + k_{x_2}^2}{\omega^2 s^2}}{1 - 2(\epsilon - \delta) \frac{k_{x_1}^2 + k_{x_2}^2}{\omega^2 s^2}}} \Bigg|_{\epsilon=\delta} = \sqrt{1 - (1 + 2\epsilon) \frac{k_{x_1}^2}{\omega^2 s^2} - (1 + 2\epsilon) \frac{k_{x_2}^2}{\omega^2 s^2}}. \quad (10)$$

From equation 20 we see that equation 7 is no more complex than the dispersion relationship for propagating waves in elliptically anisotropic media, which is now routinely handled with optimized finite-difference approaches (Zhang et al., 2001; Baumstein and Anderson, 2003; Shan and Biondi, 2005).

3D IMPLICIT FINITE-DIFFERENCE PROPAGATION

A general approach to 3D implicit finite-difference propagation is to approximate the square-root by a series of rational functions (Ma, 1982)

$$S_\alpha = \sqrt{1 - B^2 S_\beta^2 - C^2 S_\gamma^2} \approx \sum_{i=1}^n \frac{a_i S_r^2}{1 - b_i S_r^2}, \quad (11)$$

where $S_i = \frac{k_i}{\omega s_A}$ for $i = \alpha, \beta, \gamma$, term $S_r^2 = B^2 S_\beta^2 + C^2 S_\gamma^2$, and n is the order of the coefficient expansion. At this point, we do not address the anisotropy generated by the B and C coefficients, as they can be implemented through additional slowness model stretches.

One procedure for finding an optimal set of coefficients is to solve the following optimization problem (Shan and Biondi, 2005),:

$$\min \int_0^{\sin \phi} \left[\sqrt{1 - S_r^2} - \sum_{i=1}^n \frac{a_i S_r^2}{1 - b_i S_r^2} \right]^2 dS_r, \quad (12)$$

where ϕ is the maximum optimization angle. We generated the following results using a 4th-order approximation and coefficients found in Table 1 (Lee and Suh, 1985).

Extrapolation Algorithm

Using the 4th-order approximation is equivalent to solving a cascade of partial differential equations (Shan and Biondi, 2005)

$$\begin{aligned} \frac{\partial}{\partial \alpha} U &= i\omega s U, \\ \frac{\partial}{\partial \alpha} U &= i\omega s \left[\frac{\frac{a_1 B^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \beta^2} + \frac{a_1 C^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \gamma^2}}{1 + \frac{b_1 B^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \beta^2} + \frac{b_1 C^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \gamma^2}} \right] U, \\ \frac{\partial}{\partial \alpha} U &= i\omega s \left[\frac{\frac{a_2 B^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \beta^2} + \frac{a_2 C^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \gamma^2}}{1 + \frac{b_2 B^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \beta^2} + \frac{b_2 C^2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \gamma^2}} \right] U. \end{aligned} \quad (13)$$

We solve these equations implicitly at each $\Delta\alpha$ extrapolation step by a finite-difference splitting method that alternately advances the wavefield in the β and γ directions. Splitting methods allow us to apply the B and C scaling factors directly by introducing rescaled effective slowness models: $s_{eff}^B = \frac{s_A}{B} = \frac{A}{B}s$ and $s_{eff}^C = \frac{s_A}{C} = \frac{A}{C}s$.

One drawback to splitting methods is that they generate numerical anisotropy. To minimize these effects, we apply a Fourier-domain phase-correction filter (Li, 1991)

$$U = U e^{i\Delta\alpha k_L}, \quad (14)$$

where

$$k_L = \sqrt{1 - \frac{k_\beta^2}{(\omega s_B^r)^2} - \frac{k_\gamma^2}{(\omega s_C^r)^2}} - \left[1 - \sum_{i=1}^2 \left(\frac{a_i (\frac{k_\beta}{\omega s_A^r})^2}{1 - b_i (\frac{k_\beta}{\omega s_A^r})^2} - \frac{a_i (\frac{k_\gamma}{\omega s_A^r})^2}{1 - b_i (\frac{k_\gamma}{\omega s_A^r})^2} \right) \right], \quad (15)$$

and s_B^r and s_C^r are reference slownesses chosen to be the mean value of s_{eff}^B and s_{eff}^C defined above.

Impulse Response Tests

We conducted impulse response tests on a 500 x 400 x 400 mesh in a homogeneous medium with slowness $s = 0.0005$ s/m. The initial wavefield consisted of three smoothed point sources at $t = [0.5, 0.75, 1]$ seconds. Using this experimental setup, we expect the impulse responses to consist of three hemispherical surfaces of radii $r = [1000, 1500, 2000]$ meters. We used the narrow-azimuth coordinate system pictured in figure 1.

Figure 3's upper and lower panels show the inline and crossline responses, respectively. To illustrate the accuracy of the approach, we overlaid three lines showing the analytical results. Note that the impulse responses are limited at high angles both by the coordinate system boundaries and by the 50 sample cosine-taper function applied at the edges. Figure 4 shows the 1300 m depth slice. The symmetric response indicates that the Li filter has accounted for the numerical anisotropy from the numerical splitting.

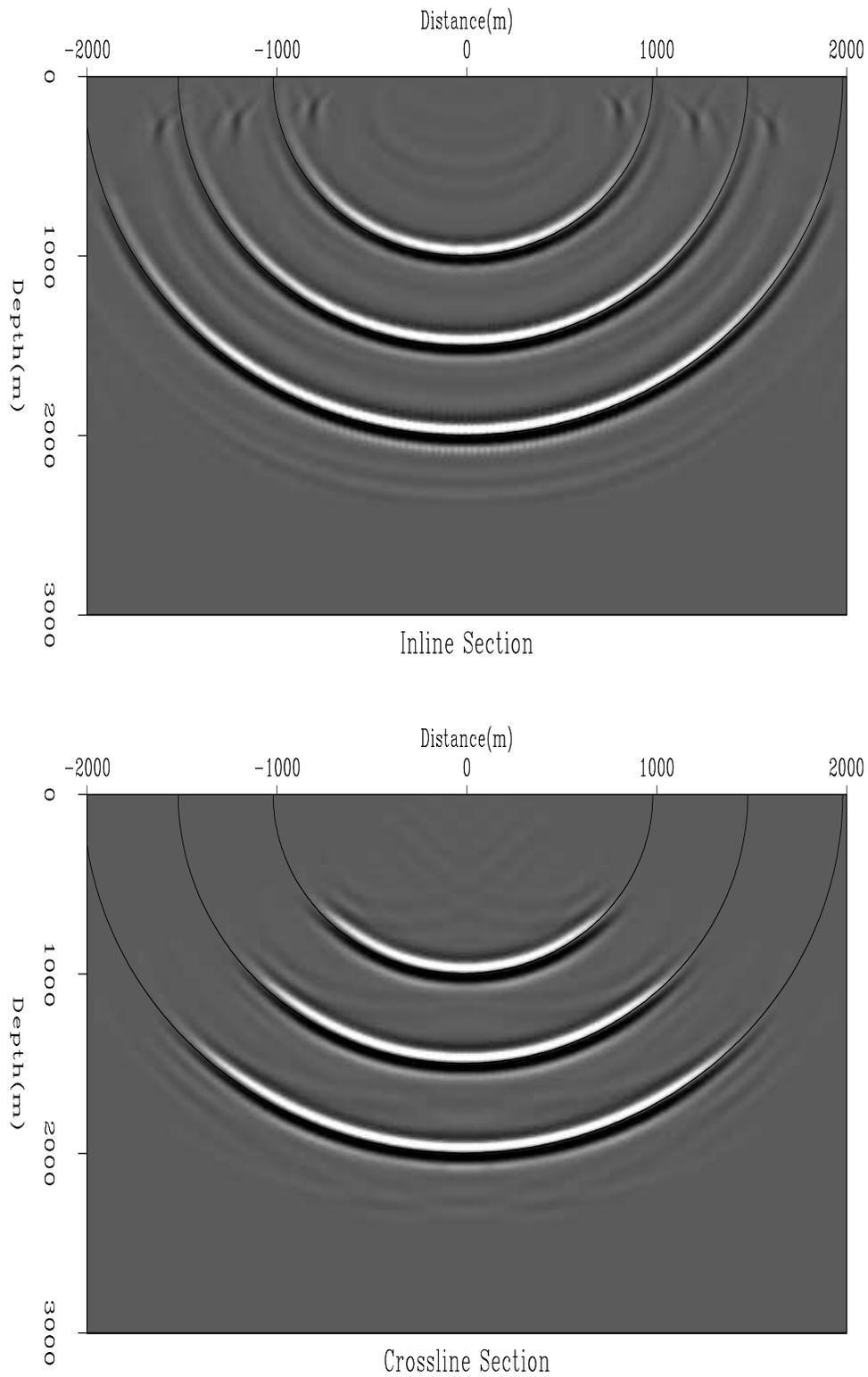
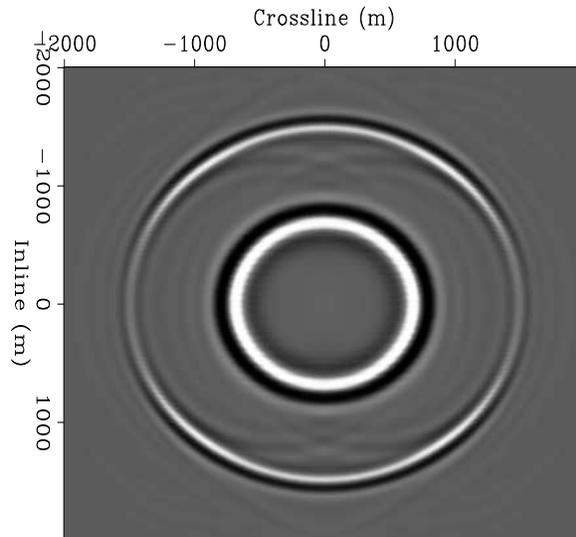


Figure 4: Ellipsoidal coordinate impulse response cross sections through the 3D image volume for the narrow-azimuth ellipsoidal coordinate system defined in figure 1. Top panel: Inline response. Bottom panel: Crossline response. [CR] jeff2/. CrossIn

Figure 5: Image volume depth slice taken at 1300m. Note the circular symmetry of the impulse response and the stronger inline amplitudes relative to the crossline. [CR]

jeff2/. Depth1300



Depth Slice @ 1300 m

FUTURE WORK

Our main motivation for exploring alternative 3D imaging geometries is to facilitate the propagation of overturning events in all directions. While this study suggests this goal is achievable, we recognize that imaging overturning events is likely to be of greater significance in situations where: i) the illumination from downgoing waves is minimal (i.e. salt body shadow zones); and ii) the seismic data acquisition enables the recording of these events (i.e. wide apertures and long recording times). Currently, SEP has neither a demonstrative wide-azimuth data set (with an isotropic velocity model) for testing purposes nor the computational power to do so. However, the ellipsoidal coordinate shot-profile algorithm will, hopefully, be tested on an appropriate data set using an industrial-sized cluster.

Another issue warranting mention is the additional computational overhead associated with ellipsoidal coordinate shot-profile migration. Below is a short list discussing 3D migration issues in ellipsoidal (ECC) and Cartesian (CCC) coordinate systems:

- Both the inline and crossline migration aperture could be reduced in an ECC because, unlike in a CCC, an ECC spreads out naturally and increases the effective aperture. (Sufficient volume sampling is still required to obtain a good image.)
- Imaging in non-Cartesian coordinate systems require two additional full-volume interpolations, one each between the corresponding velocity models and image spaces. While the former interpolation is easily parallelized, the later requires spraying each ECC image point to a neighborhood of CCC points. This remains difficult to parallelize efficiently, except by expanding the image volume by a factor equal to the number of threads used for interpolation.
- Migration costs can be lowered, perhaps by an order of magnitude, by imaging neighboring shot profiles in the same ECC system.
- Because the ECC and CCC originate on the same initial surficial plane, no wavefield

injection is required. However, the sampling of the two initial surfaces is different and requires an initial interpolation.

CONCLUSIONS

This paper presents a 3D shot-profile migration approach in ellipsoidal coordinates. The common form of ellipsoidal coordinates generates confocal ellipsoidal shells well-positioned for propagating turning waves in all azimuths. The numerical aspects of implementing the corresponding dispersion relationship, though, are somewhat problematic. We use an integral transform to recast the problem in a way that facilitates numerical implementation. We note that the similarity between the new ellipsoidal and elliptically anisotropic media dispersion relationships allows us to use existing implicit finite-difference extrapolation implementations. The impulse responses demonstrate that the numerical solution is both accurate and stable. Future work will focus on generating 3D shot-profile migration results in the ellipsoidal coordinate system.

ACKNOWLEDGMENTS

We would like to thank Ben Witten for helpful conversations.

REFERENCES

- Abramowitz, M. and I. E. Stegun, 1965, "Jacobian Elliptic Functions and Theta Functions." Ch. 17.1-17.6 in Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing: New York: Dover.
- Baumstein, A. and J. Anderson, 2003, Wavefield extrapolation in laterally varying VTI media: 73rd Annual International Meeting, SEG, Expanded Abstracts, 945–948.
- Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524.
- Biondi, B. and G. Shan, 2002, Prestack imaging of overturned reflections by reverse-time migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1284–1287.
- Byerly, E., 1959, An Elementary Treatise on Fourier's series, and Spherical, Cylindrical, and Ellipsoidal Harmonics, with Applications to Problems in Mathematical Physics: New York: Dover.
- Hale, D., N. Hill, and J. Stephani, 1992, Imaging salt with turning seismic waves: *Geophysics*, **57**, 1453–1462.
- Lee, M. W. and S. Y. Suh, 1985, Optimization of one-way wave-equations (short note): *Geophysics*, **50**, 1634–1637.
- Li, Z., 1991, Compensating finite-difference errors in 3D migration and modeling: *Geophysics*, **56**, 1650–1660.
- Ma, Z., 1982, Finite-difference migration with higher-order approximation: *Oil Geophysical Prospecting*, **1**, 6–15.
- Morse, P. and H. Feshbach, 1953, *Methods of Theoretical Physics*: Cambridge University Press.
- Shan, G. and B. Biondi, 2004, Imaging overturned waves by plane-wave migration in tilted coordinates: 74th Annual International Meeting, SEG, Expanded Abstracts, 696–672.
- , 2005, 3D wavefield extrapolation in laterally-varying tilted TI media: 75th Annual International Meeting, SEG, Expanded Abstracts, 104–107.
- , 2006, Imaging steep salt flank with plane-wave migration in tilted coordinates: 76th Annual International Meeting, SEG, Expanded Abstracts, 2372–2375.
- Shragge, J., 2008, Riemannian wavefield extrapolation: Nonorthogonal coordinate systems: *Geophysics*, **73**, T11–T21.
- Shragge, J. and G. Shan, 2008, 2D prestack depth migration in elliptic coordinates: accepted in *Geophysics*.
- Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966.
- Tsvankin, I., 1996, P-wave signatures and notation for transversely isotropic media: An overview: *Geophysics*, **61**, 467–483.
- Zhang, J. and G. McMechan, 1997, Turning wave migration by horizontal extrapolation: *Geophysics*, **62**, 291–297.
- Zhang, J., D. Verschuur, and C. A. P. Wapenaar, 2001, Depth migration of shot records in heterogeneous, transversely isotropic media using optimum explicit operators: *Geophysical Prospecting*, **49**, 287–299.
- Zhang, Y. and S. Xu, 2006, Imaging complex salt bodies with turning-wave one-way wave-equation: 76th Annual International Meeting, SEG, Expanded Abstracts, 2323–2327.

APPENDIX A

INTEGRAL ELLIPTIC REPRESENTATIONS

The integral elliptic representations of the ellipsoidal coordinate system are given by the following three equations (Abramowitz and Stegun, 1965):

$$\begin{aligned}
 \beta &= c \int_b^{\xi_1} \frac{d\xi_1}{(c^2 - \xi_1^2)(\xi_1^2 - c^2)} = F \left[\sqrt{1 - \frac{b^2}{c^2}}, \sin^{-1} \left(\sqrt{\frac{1 - \frac{b^2}{c^2}}{1 - \frac{\xi_1^2}{c^2}}} \right) \right], \\
 \gamma &= c \int_0^{\xi_1} \frac{d\xi_1}{(b^2 - \xi_1^2)(c^2 - \xi_1^2)} = F \left[\frac{b}{c}, \sin^{-1} \left(\frac{\xi_2}{b} \right) \right], \\
 \alpha &= c \int_c^{\xi_2} \frac{d\xi_2}{(\xi_2^2 - b^2)(\xi_2^2 - c^2)} = F \left[\frac{b}{c}, \frac{\pi}{2} \right] - F \left[\frac{b}{c}, \sin^{-1} \left(\frac{c}{\xi_3} \right) \right],
 \end{aligned} \tag{A-1}$$

where $F[\cdot]$ are elliptic integrals of the first kind defined by

$$u = F[\phi, k] = \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}, \tag{A-2}$$

where the elliptic modulus k satisfies $0 < k^2 < 1$ and ϕ is the Jacobi amplitude. Solutions to equation A-2 are calculated using the method of arithmetic-geometric mean and descending Landen Transformation described in Abramowitz and Stegun (1965).

The integral transforms are invertible and can be represented in terms of Jacobi elliptic functions $\text{dc}(u, k)$, $\text{nd}(u, k)$ and $\text{sn}(u, k)$ (Abramowitz and Stegun, 1965):

$$\begin{aligned}
 \xi_1 &= b \text{nd} \left[\beta, \sqrt{1 - \frac{b^2}{c^2}} \right], \\
 \xi_2 &= b \text{sn} \left[\gamma, \frac{b}{c} \right], \\
 \xi_3 &= c \text{dc} \left[\alpha, \frac{b}{c} \right],
 \end{aligned} \tag{A-3}$$

where k is again elliptic modulus and u is defined by equation A-2. The Jacobi elliptic functions are calculated using the method of the arithmetic-geometric mean described in Abramowitz and Stegun (1965).

Toward 3D conical-wave migration in tilted elliptic cylindrical coordinates

Jeff Shragge and Guojian Shan

ABSTRACT

We extend conical-wave migration to tilted elliptic cylindrical (TEC) coordinate systems. When inline coordinate tilt angles are well-matched to the inline plane-wave ray parameters, the TEC coordinate extension affords accurate propagation of steep-dip and turning-wave components of conical wavefields in both the in- and crossline directions. We show that wavefield extrapolation in TEC coordinates is no more complicated than propagation in elliptically anisotropic media. Impulse response tests illustrate the accuracy of the approach. Future work will apply the conical-wave migration approach to field data sets.

INTRODUCTION

Wave-equation migration (WEM) methods routinely are now used to generate seismic images in areas of complex geology. One common class of WEM approaches is shot-profile migration based on one-way wavefield extrapolation. The first step is to specify source and receiver wavefields that consist of modeled point sources and the recorded shot-profile data, respectively. The second step is to propagate these two wavefields through the velocity model and correlate them to form an image. A drawback of this approach is that individual profiles routinely cover only a small portion of the total survey area and many shot-profile migrations are required to infill the image volume. One way to make this type of WEM more efficient is to migrate a reduced number of composite source and receiver profiles each covering a broader area of the image space; however, this requires mixing the information from different shots and can lead to image crosstalk.

Plane-wave migration (PWM) is one technique for reducing total migration cost by generating a lower number of composite wavefields (Rietveld et al., 1992; Whitmore, 1995; Mosher and Foster, 1998; Duquet et al., 2001; Zhang et al., 2003; Liu et al., 2004, 2006). The key idea is to synthesize a set of composite receiver wavefields that would have been recorded were a planar source wavefield used. One then generates the PWM image by propagating the modeled planar source wavefield and the composite receiver wavefield through the velocity and computing a (weighted) correlation. Liu et al. (2006) and Duquet and Lailly (2006) demonstrate that PWM is equivalent to shot-profile migration in situations where one uses large numbers of plane waves with well-sampled plane-wave dip spectra. Liu et al. (2006) also prove that 3D PWM is equivalent to conical-wave migration of individual sail lines formed as inline composite wavefields. The approach is termed conical-wave because the source wavefronts form cones for non-zero inline plane-wave ray parameters. More generally, these PWM approaches are examples of phase-encoding class of WEM (Morton and Ober, 1998; Jing et al., 2000; Romero et al., 2000; Sun et al., 2002).

The migration of plane-wave sections - though more efficient than shot-profile migration - is similarly restricted by one-way wavefield extrapolation assumptions. The most common limitation is a difficulty in propagating waves at steep angles and turning waves by design, both of which are important for accurate imaging in complex geologic areas. Shan and Biondi (2004) circumvent this problem by proposing fully 3D PWM in a suite of tilted Cartesian meshes. This effectively orients the wavefield extrapolation axis toward the plane-wave take-off vector, enabling more accurate bulk propagation of plane-wave energy.

One logistical complication of performing fully 3D PWM is it requires propagating image-space-sized data volumes on a number of meshes tilting in both the in- and crossline directions. This leads to a number of computational issues associated with the significant memory footprint. In this paper, we follow a similar alternative coordinate system approach for performing PWM. However, we chose to concentrate on using a coordinate system well-suited for migrating individual sail lines with a reduced memory footprint. We term this approach conical-wave migration on tilted elliptic cylindrical coordinates.

A 3D elliptic cylindrical coordinate system is formed by concatenating a set of confocal 2D elliptic surfaces along the remaining axis. We orient the invariant and elliptic surfaces in the in- and crossline direction, respectively. We extrapolate the conical wavefields outward on a series of confocal elliptic cylindrical shells. This allows source and receiver wavefields with zero inline dip to overturn, if necessary, in the crossline direction. We also introduce an additional degree of freedom to tilt the coordinate system in the direction of the invariant inline axis. Consequently, we propagate plane-wave sources with non-zero inline dips accurately to steep, and even turning, angles.

We begin by revisiting 3D plane-wave and conical-wave migration theory. We introduce the tilted elliptic cylindrical coordinates and then develop the corresponding wavenumber used to form the wavefield extrapolation operator. We then discuss numerical issues regarding implementing the finite-difference extrapolation scheme, and present the point-source responses.

PLANE-WAVE MIGRATION

The 3D plane-wave and conical-wave migration theory discussed in this paper draws heavily on that in Liu et al. (2006). We restate a number of significant and related points for completeness, though with a slightly different notation.

Performing 3D plane-wave migration is similar in many respects to 3D shot-profile migration. The main difference is rooted in how the complete source and receiver wavefield volumes, \bar{S} and \bar{R} , are specified from individual source and receiver records, S_{ij} and R_{kl} , prior to imaging. The complete wavefields are generated by filtering the source and receiver profiles by a function dependent on the inline and cross-line plane-wave slownesses, $\mathbf{p}\boldsymbol{\xi} = [p_{\xi_1}, p_{\xi_2}]$. The phase-encoded wavefields are then propagated through the migration domain (herein defined by $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3]$) to generate the source and receiver wavefield

volumes

$$\overline{S(\boldsymbol{\xi}|\omega)} = \sum_{i=1}^M \sum_{j=1}^N S_{ij}(\boldsymbol{\xi}|\omega) f(\omega) e^{i\omega[p_{\xi_1} \Delta\xi_1(i-p) + p_{\xi_2} \Delta\xi_2(j-q)]}, \quad (1)$$

$$\overline{R(\boldsymbol{\xi}|\omega)} = \sum_{k=1}^M \sum_{l=1}^N R_{kl}(\boldsymbol{\xi}|\omega) f(\omega) e^{i\omega[p_{\xi_1} \Delta\xi_1(k-p) + p_{\xi_2} \Delta\xi_2(l-q)]}, \quad (2)$$

where $f(\omega)$ is a frequency (ω) filter to be discussed below, $\Delta\xi_1$ and $\Delta\xi_2$ are the inline and crossline sampling intervals, p and q are reference spatial indices in the inline and crossline directions, i and j are the inline and crossline components fixing the source position, k and l are the inline and crossline components fixing the receiver position, and M and N are the number of inline and cross-line records. Note that the phase-encoding is performed at the surface and is unaffected by the subsequent wavefield extrapolation. Thus, this formulation is valid for any coordinate system, including TEC computational meshes.

The image volume $I(\boldsymbol{\xi})$ is formed from a series of plane-wave migration images, $I^{pw}(\boldsymbol{\xi}|\mathbf{p}_\xi)$, by correlating the composite plane-wave source and receiver wavefields and summing the result over each frequency. The plane-wave migration kernel mixes energy from a source, $S_{ij}(\boldsymbol{\xi}|\omega)$, and a receiver, $R_{kl}(\boldsymbol{\xi}|\omega)$, wavefield according to

$$\begin{aligned} I(\boldsymbol{\xi}) &= \sum_{p_{\xi_1}} \sum_{p_{\xi_2}} \sum_{i,k=1}^M \sum_{j,l=1}^N I_{ijkl}^{pw}(\boldsymbol{\xi}|\mathbf{p}_\xi) \\ &= \sum_{p_{\xi_1}} \sum_{p_{\xi_2}} \sum_{i,k=1}^M \sum_{j,l=1}^N \sum_{\omega} |f(\omega)|^2 S_{ij}^*(\boldsymbol{\xi}|\omega) R_{kl}(\boldsymbol{\xi}|\omega) e^{i\omega[p_{\xi_1} \Delta\xi_1(i-k) + p_{\xi_2} \Delta\xi_2(j-l)]}, \end{aligned} \quad (3)$$

where * indicates complex conjugate.

Generally, mixing wavefields of differing S_{ij} and R_{kl} indices introduces crosstalk into the image volume. However, plane-wave migration will be crosstalk-free when the following hold:

$$\begin{aligned} \lim_{N_{p_{\xi_1}} \rightarrow \infty} \sum_{\alpha=-N_{p_{\xi_1}}}^{N_{p_{\xi_1}}} e^{i\omega\alpha\Delta p_{\xi_1} \Delta\xi_1(i-k)} &= |\omega|^{-1} \delta_{ik}, \\ \lim_{N_{p_{\xi_2}} \rightarrow \infty} \sum_{\alpha=-N_{p_{\xi_2}}}^{N_{p_{\xi_2}}} e^{i\omega\alpha\Delta p_{\xi_2} \Delta\xi_2(j-l)} &= |\omega|^{-1} \delta_{jl}. \end{aligned} \quad (4)$$

Assuming that equation 4 approximately is valid (i.e., for large values of $N_{p_{\xi_1}}$ and $N_{p_{\xi_2}}$), we may rewrite equation 3 as

$$I(\boldsymbol{\xi}) \approx \sum_{i=1}^M \sum_{j=1}^N \sum_{\omega} |f(\omega)|^2 |\omega|^{-2} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{ij}(\boldsymbol{\xi}|\omega), \quad (5)$$

or by defining $|f(\omega)|^2 = |\omega|^2$ we obtain

$$I(\boldsymbol{\xi}) \approx \sum_{i=1}^M \sum_{j=1}^N \sum_{\omega} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{ij}(\boldsymbol{\xi}|\omega), \quad (6)$$

which demonstrates the equivalence between plane-wave and shot-profile migration (Liu et al., 2006).

Conical-wave migration

A second plane-wave migration formulation is to treat each 3D source plane-wave migration as if it were independent of the cross-line ray parameter. Note that this framework only is true in certain circumstances: straight sail lines and no flip-flop sources (Liu et al., 2006). Mathematically, this requires satisfying the condition that $j - l = 0$ in the exponential function in equation 3. If this relation holds, then we may write the following for the conical-wave image, I_j^{cw} , for the j^{th} sail line:

$$\begin{aligned}
 I_j^{cw}(\boldsymbol{\xi}) &= \sum_{p_{\xi_1}} I_j(\boldsymbol{\xi}|p_{\xi_1}) \\
 &= \sum_{p_{\xi_1}} \sum_{i,k=1}^M \sum_{\omega} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{kj}(\boldsymbol{\xi}|\omega) |f(\omega)|^2 e^{i\omega p_{\xi_1} \Delta \xi_1 (i-k)} \\
 &\approx \sum_{i=1}^M \sum_{\omega} |\omega|^{-1} |f(\omega)|^2 S_{ij}^*(\boldsymbol{\xi}|\omega) R_{ij}(\boldsymbol{\xi}|\omega),
 \end{aligned} \tag{7}$$

or, by defining $|f(\omega)|^2 = |\omega|$, we may rewrite

$$I_j^{cw}(\boldsymbol{\xi}) = \sum_{i=1}^M \sum_{\omega} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{ij}(\boldsymbol{\xi}|\omega). \tag{8}$$

Stacking the individual sail-line conical-wave migration results yields the full image volume,

$$I(\boldsymbol{\xi}) = \sum_{j=1}^N I_j^{cw}(\boldsymbol{\xi}) = \sum_{j=1}^N \sum_{i=1}^M \sum_{\omega} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{ij}(\boldsymbol{\xi}|\omega). \tag{9}$$

This completes the proof of the equivalence of conical-wave and shot-profile migration (Liu et al., 2006).

TILTED ELLIPTIC CYLINDRICAL COORDINATES

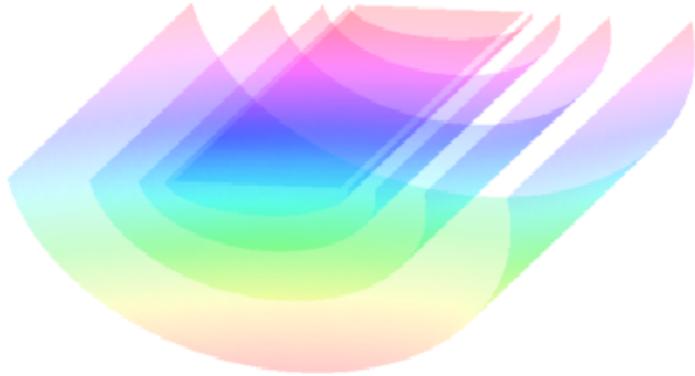
One question to be addressed is why did we choose TEC coordinates for conical-wave migration? We assert that the geometry of a TEC coordinate system, shown in figures 1 and 2, are well-suited to conical-wave migration. In particular, the width of the first extrapolation step allows us to position individual, as well as multiple, streamers of a single sail line directly on a single mesh. The singular points of the coordinate system also are located along the side edges, which facilitates numerical stability. We note that the other natural coordinate system, polar cylindrical coordinates, is not a judicious choice for conical-wave migration because the geometry: i) permits only single-streamer data to be placed on the mesh; and ii) has singular points located on the surface at the first extrapolation step.

We begin by defining the transformation between the Cartesian and elliptic cylindrical meshes. We denote a Cartesian coordinate system by $\boldsymbol{x} = [x_1, x_2, x_3]$ and specify the tilted

Figure 1: Four extrapolation steps in ξ_3 of an TEC coordinate system with a 0° tilt angle. The ξ_1 coordinate axis is oriented in the inline direction, while the ξ_2 coordinate axis is directed toward the crossline direction. [NR] `jeff3/. NoTiltTEC`



Figure 2: As in figure 1, but for a 25° tilt angle. [NR] `jeff3/. TiltTEC`



elliptic cylindrical mesh by $\boldsymbol{\xi} = [\xi_1, \xi_2, \xi_3]$. We set up our elliptic cylindrical geometry as follows:

- $\xi_1 \in [-\infty, \infty]$ is the inline direction, where surfaces of constant ξ_1 are 2D elliptic coordinate meshes;
- $\xi_2 \in [0, 2\pi)$ is the cross-line direction, where surfaces of constant ξ_2 are folded hyperbolic planes; and
- $\xi_3 \in [0, \infty]$ is the extrapolation direction, where surfaces of constant ξ_3 form concentric elliptic cylinders.

The mapping relationship between the two coordinate systems is given by (Arfken, 1970)

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \xi_1 \cos \gamma - a \sinh \xi_3 \sin \xi_2 \sin \gamma \\ a \cosh \xi_3 \cos \xi_2 \cos \gamma \\ \xi_1 \sin \gamma + a \sinh \xi_3 \sin \xi_2 \cos \gamma \end{bmatrix}, \quad (10)$$

where γ is the inline tilt angle of the coordinate system and parameter a controls the coordinate system breadth. Figure 1 shows a TEC coordinate system for a 0° tilt angle, while figure 2 shows the same coordinate system, but for a 25° tilt angle.

TEC extrapolation wavenumber

The metric tensor ($g_{ij} = \frac{\partial x_k}{\partial \xi_i} \frac{\partial x_k}{\partial \xi_j}$ with an implicit summation over k) describing the geometry of the elliptic coordinate system is given by,

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & A^2 & 0 \\ 0 & 0 & A^2 \end{bmatrix}, \quad (11)$$

where $A = a\sqrt{\sinh^2 \xi_3 + \sin^2 \xi_2}$. The determinant of the metric tensor is: $\sqrt{|g|} = A^4$. The associated (inverse) metric tensor is given by,

$$[g^{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & A^{-2} & 0 \\ 0 & 0 & A^{-2} \end{bmatrix}, \quad (12)$$

and weighted metric tensor ($m^{ij} = \sqrt{|g|} g^{ij}$) is given by,

$$[m^{ij}] = \begin{bmatrix} A^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (13)$$

Note that even though the TEC coordinate system varies spatially, the local curvature parameters ($n^i = \frac{\partial m^{ij}}{\partial \xi_j}$) remain constant: $n^1 = n^2 = n^3 = 0$. The corresponding extrapolation wavenumber, k_{ξ_3} , can be generated by inputting tensors g^{ij} , m^{ij} and n^j into the general wavenumber expression for 3D non-orthogonal coordinate systems (Shragge, 2008).

$$k_{\xi_3} = \pm \sqrt{A^2 s^2 \omega^2 - k_{\xi_1}^2 - A^2 k_{\xi_2}^2}, \quad (14)$$

where s is the slowness (reciprocal of velocity).

The wavenumber developed in equation 14 is central to developing a shot-profile migration algorithm. This procedure consists of two parts. First, one extrapolates the source S_{ij} and receiver R_{kl} wavefields according to

$$S_{ij}(\xi_1, \xi_2, \xi_3 + \Delta\xi_3|\omega) = \mathbf{E}[S_{ij}(\xi_1, \xi_2, \xi_3|\omega)], \quad (15)$$

$$R_{kl}(\xi_1, \xi_2, \xi_3 + \Delta\xi_3|\omega) = \mathbf{E}[R_{kl}(k_{\xi_1}, k_{\xi_2}, \xi_3|\omega)], \quad (16)$$

where $\mathbf{E}[\cdot]$ is an extrapolation operator. The results in this paper were computed using frequency-space, $\omega - \mathbf{x}$, finite-difference extrapolators (see below).

The second step involves summing the individual shot images contributions, $I_{ij}^{sp}(\boldsymbol{\xi})$, into the total image volume, $I(\boldsymbol{\xi})$. Individual shot images are generated by correlating the S_{ij} and R_{kl} profiles:

$$I(\boldsymbol{\xi}) = \sum_{i=1}^M \sum_{j=1}^N I_{ij}^{sp}(\boldsymbol{\xi}) \quad (17)$$

$$= \sum_{i,k=1}^M \sum_{j,l=1}^N \sum_{\omega} S_{ij}^*(\boldsymbol{\xi}|\omega) R_{kl}(\boldsymbol{\xi}|\omega) \delta_{ik} \delta_{jl} \quad (18)$$

where symbol $*$ denotes complex conjugation, M and N are the maximum number of inline and crossline sources, and $\delta_{ik} \delta_{jl}$ are delta functions forcing the source and receiver profiles to be from the same inline and cross-line shot location.

3D IMPLICIT FINITE-DIFFERENCE EXTRAPOLATION

A naturally arising concern is whether the dispersion relationship in equation 14 can be implemented accurately and efficiently in a wavefield extrapolation scheme. We address this question by comparing the elliptic cylindrical and the Cartesian elliptically anisotropic media dispersion relationships. By defining an effective slowness $s_A = As$ and rewriting equation 14 as

$$\frac{k_{\xi_3}}{\omega s_A} = \sqrt{1 - \frac{k_{\xi_1}^2}{\omega^2 s_A^2} - A^2 \frac{k_{\xi_2}^2}{\omega^2 s_A^2}}, \quad (19)$$

we observe that the TEC coordinate dispersion relationship resembles that for elliptically anisotropic media (Tsvankin, 1996). More specifically, TEC coordinates relate to a special case where the Thomsen parameters (Thomsen, 1986) obey $\epsilon = \delta$

$$\left. \frac{k_{x_3}}{\omega s} \right|_{\epsilon=\delta} = \sqrt{\frac{1 - (1 + 2\epsilon) \frac{k_{x_1}^2 + k_{x_2}^2}{\omega^2 s^2}}{1 - 2(\epsilon - \delta) \frac{k_{x_1}^2 + k_{x_2}^2}{\omega^2 s^2}}} \Bigg|_{\epsilon=\delta} = \sqrt{1 - (1 + 2\epsilon) \frac{k_{x_1}^2}{\omega^2 s^2} - (1 + 2\epsilon) \frac{k_{x_2}^2}{\omega^2 s^2}}. \quad (20)$$

From equation 20 we see that equation 14 is no more complex than the dispersion relationship for propagating waves in elliptically anisotropic media, which is now routinely handled with optimized finite-difference approaches (Zhang et al., 2001; Baumstein and Anderson, 2003; Shan and Biondi, 2005).

Rational Approximations

A general approach to 3D implicit finite-difference propagation is to approximate the square-root by a series of rational functions (Ma, 1982)

$$S_{\xi_3} = \sqrt{1 - S_{\xi_1}^2 - A^2 S_{\xi_2}^2} \approx \sum_{i=1}^n \frac{a_i S_r^2}{1 - b_i S_r^2}, \quad (21)$$

where $S_{\xi_i} = \frac{k_{\xi_i}}{\omega s_A}$ and $S_r^2 = S_{\xi_1}^2 + A^2 S_{\xi_2}^2$, for $i = 1, 2, 3$, and n is the order of the coefficient expansion. At this point, we do not address the anisotropy generated by the A coefficient, as it can be implemented through an additional slowness model stretch.

One procedure for finding an optimal set of coefficients is to solve the following optimization problem (Shan and Biondi, 2005),

$$\min \int_0^{\sin \phi} \left[\sqrt{1 - S_r^2} - \sum_{i=1}^n \frac{a_i S_r^2}{1 - b_i S_r^2} \right]^2 dS_r, \quad (22)$$

where ϕ is the maximum optimization angle. We generated the following results using a 4th-order approximation and coefficients found in Table 1 (Lee and Suh, 1985).

Coeff. order i	Coeff. a_i	Coeff. b_i
1	0.040315157	0.873981642
2	0.457289566	0.222691983

Table 1: Coefficients used in 3D implicit finite-difference wavefield extrapolation.

Extrapolation Algorithm

Using the 4th-order approximation is equivalent to solving a cascade of partial differential equations (Shan and Biondi, 2005)

$$\begin{aligned}
\frac{\partial}{\partial \xi_3} U &= i\omega s U, \\
\frac{\partial}{\partial \xi_3} U &= i\omega s \left[\frac{\frac{a_1}{\omega^2 s_A^2} \frac{\partial^2}{\partial \xi_1^2} + \frac{a_1}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}}{1 + \frac{b_1}{\omega^2 s_A^2} \frac{\partial^2}{\partial \xi_1^2}} + \frac{\frac{a_1}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}}{1 + \frac{b_1}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}} \right] U, \\
\frac{\partial}{\partial \xi_3} U &= i\omega s \left[\frac{\frac{a_2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \xi_1^2} + \frac{a_2}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}}{1 + \frac{b_2}{\omega^2 s_A^2} \frac{\partial^2}{\partial \xi_1^2}} + \frac{\frac{a_2}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}}{1 + \frac{b_2}{\omega^2 s^2} \frac{\partial^2}{\partial \xi_2^2}} \right] U.
\end{aligned} \tag{23}$$

We solve these equations implicitly at each extrapolation step by a finite-difference splitting method that alternatively advances the wavefield in the ξ_1 and ξ_2 directions. Splitting methods allow us to apply the A scaling factor directly by introducing the original slowness model: $\frac{s_A}{A} = s$.

One drawback to splitting methods is that they generate numerical anisotropy. To minimize these effects, we apply a Fourier-domain phase-correction filter (Li, 1991)

$$U = U e^{i\Delta \xi_3 k_L}, \tag{24}$$

where

$$k_L = \sqrt{1 - \frac{k_{\xi_1}^2}{(\omega s_1^r)^2} - \frac{k_{\xi_2}^2}{(\omega s_2^r)^2}} - \left[1 - \sum_{i=1}^2 \left(\frac{a_i \left(\frac{k_{\xi_1}}{\omega s_A^r}\right)^2}{1 - b_i \left(\frac{k_{\xi_1}}{\omega s_A^r}\right)^2} - \frac{a_i \left(\frac{k_{\xi_2}}{\omega s_A^r}\right)^2}{1 - b_i \left(\frac{k_{\xi_2}}{\omega s_A^r}\right)^2} \right) \right], \tag{25}$$

and s_1^r and s_2^r are reference slownesses chosen to be the mean value of s_{eff}^A and s defined above.

NUMERICAL TESTS

We conducted impulse response tests on a 500 x 400 x 400 mesh in a homogeneous medium with slowness $s = 0.0005$ s/m. The initial wavefield consisted of three smoothed point sources at $t = [0.5, 0.75, 1]$ s. Using this experimental setup, we expect the impulse responses to consist of three hemispherical surfaces of radii $r = [1000, 1500, 2000]$ m.

Figure 3 shows the inline and crossline responses in the upper and lower panels, respectively. To illustrate the accuracy of the approach, we overlaid three lines showing that the analytical answers. Note that the impulse responses are limited at high angles both

by the coordinate system boundaries and by the 50 sample cosine-taper function applied at the edges. Figure 4 shows the 1300m depth slice. The symmetric response indicates that the numerical anisotropy from the numerical splitting has been accounted for by the Li phase-correction filter.

CONCLUSIONS

This paper introduces the tilted elliptic cylindrical coordinate system to conical-wave migration. We demonstrate that corresponding extrapolation wavenumber is no more complicated than that of elliptically anisotropic media. This allows us to implement an accurate finite-difference extrapolation approach that can handle the effective anisotropy. Future work will apply the algorithm to 3D data.

ACKNOWLEDGMENTS

We would like to thank Ben Witten for helpful conversations.

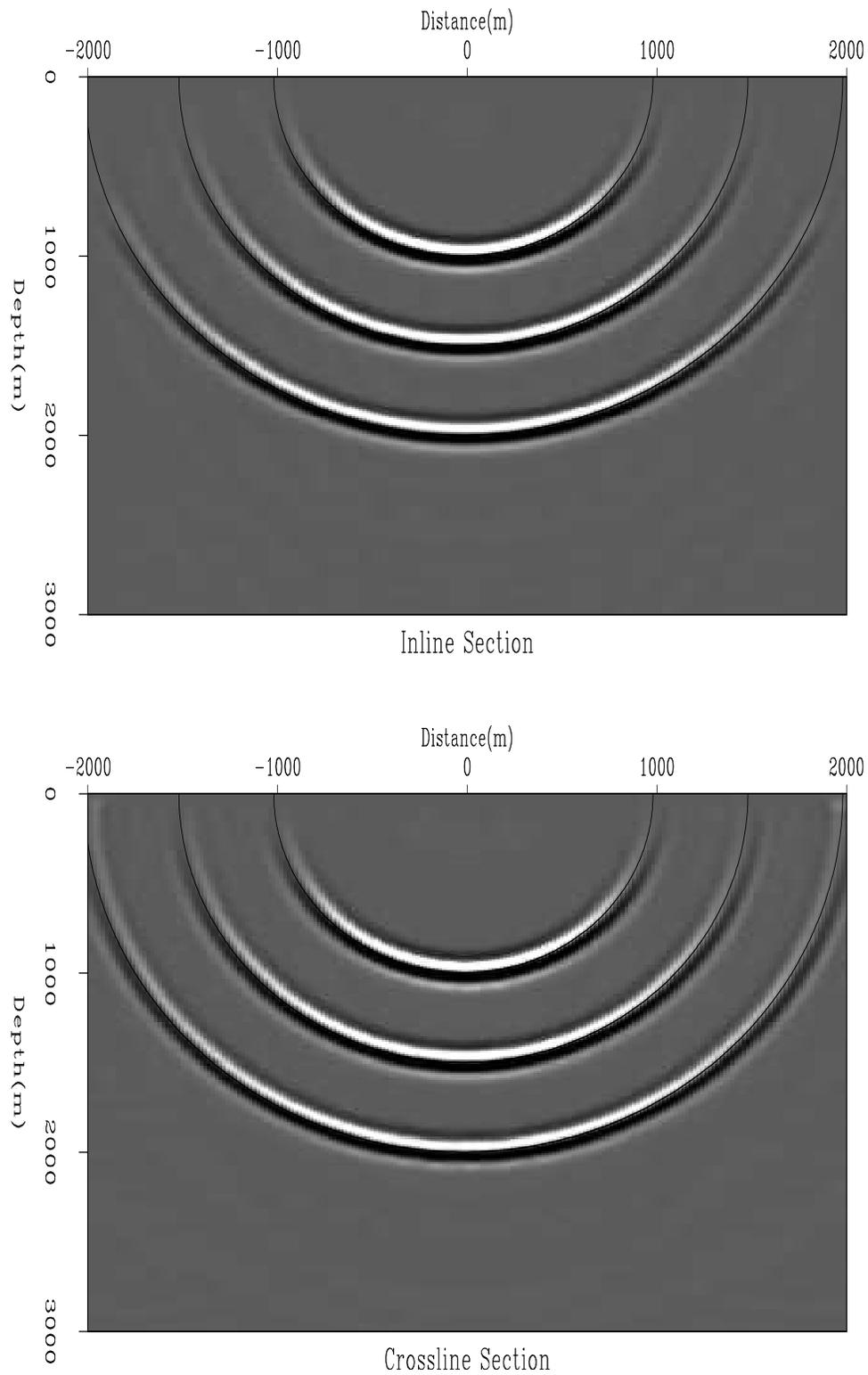
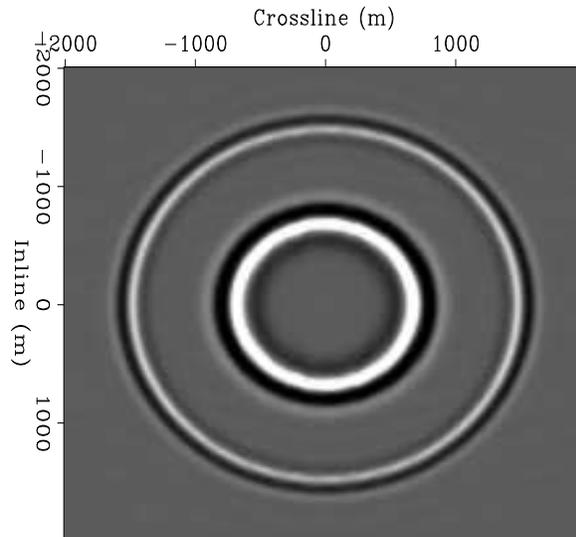


Figure 3: Elliptic cylindrical coordinate impulse response cross sections through the 3D image volume. Top panel: Inline response. Bottom panel: Crossline response.[CR]

jeff3/. CrossIn

Figure 4: Image volume depth slice taken at 1300m. Note the circular symmetry of the impulse response.

[CR] `jeff3/. Depth1300`



REFERENCES

- Arfken, G., 1970, "Confocal Ellipsoidal Coordinates" Ch 2.15 in *Mathematical Methods for Physicists*, 2nd ed.: New York: Academic Press.
- Baumstein, A. and J. Anderson, 2003, Wavefield extrapolation in laterally varying VTI media: 73rd Annual International Meeting, SEG, Expanded Abstracts, 945–948.
- Duquet, B. and P. Lailly, 2006, Efficient 3D wave-equation migration using virtual planar sources: *Geophysics*, **71**, S185–S197.
- Duquet, B., P. Lailly, and A. Ehinger, 2001, 3D plane wave migration of streamer data: 71st Annual International Meeting, SEG, Expanded Abstracts, 1033–1036, Soc. of Expl. Geophys.
- Jing, X., C. Finn, T. Dicken, and D. Willen, 2000, Encoding multiple shot gathers in prestack migration: 70th Annual International Meeting, SEG, Expanded Abstracts, 786–789.
- Lee, M. W. and S. Y. Suh, 1985, Optimization of one-way wave-equations (short note): *Geophysics*, **50**, 1634–1637.
- Li, Z., 1991, Compensating finite-difference errors in 3D migration and modeling: *Geophysics*, **56**, 1650–1660.
- Liu, F., D. Hanson, N. Whitmore, R. Day, and R. Stolt, 2006, Toward a unified analysis for source plane-wave migration: *Geophysics*, **71**, S129–S139.
- Liu, F., H. D. Whitmore, N.D., and R. Day, 2004, The impact of reciprocity on prestack source plane wave migration: 68th Annual International Meeting, SEG, Expanded Abstracts, 1131–1134.
- Ma, Z., 1982, Finite-difference migration with higher-order approximation: *Oil Geophysical Prospecting*, **1**, 6–15.
- Morton, S. and C. Ober, 1998, Faster shot record depth migrations using depth encoding: 68th Annual International Meeting, SEG, Expanded Abstracts, 1131–1134.
- Mosher, C. and D. Foster, 1998, Offset plane wave propagation in laterally varying media: *Proceedings of SPIE*.
- Rietveld, W., A. Berkhout, and C. Wapenaar, 1992, Optimum seismic illumination hydro-

- carbon reservoirs: *Geophysics*, **52**, 1334–1345.
- Romero, L., D. Ghiglian, C. Ober, and S. Morton, 2000, Phase Encoding of Shot Records in Prestack Migration: *Geophysics*, **65**, 426–436.
- Shan, G. and B. Biondi, 2004, Imaging overturned waves by plane-wave migration in tilted coordinates: 74th Annual International Meeting, SEG, Expanded Abstracts, 969–972.
- , 2005, 3D wavefield extrapolation in laterally-varying tilted TI media: 75th Annual International Meeting, SEG, Expanded Abstracts, 104–107.
- Shragge, J., 2008, Riemannian wavefield extrapolation: Nonorthogonal coordinate systems: *Geophysics*, **73**, T11–T21.
- Sun, P., S. Zhang, and F. Liu, 2002, Prestack migration of areal shot records with phase encoding: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1172–1175.
- Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966.
- Tsvankin, I., 1996, P-wave signatures and notation for transversely isotropic media: An overview: *Geophysics*, **61**, 467–483.
- Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismograms, *in* Ph.D. thesis, University of Tulsa.
- Zhang, J., D. Verschuur, and C. A. P. Wapenaar, 2001, Depth migration of shot records in heterogeneous, transversely isotropic media using optimum explicit operators: *Geophysical Prospecting*, **49**, 287–299.
- Zhang, Y., C. Notfors, S. Gray, L. Chernis, and J. Young, 2003, Delayed-shot 3D prestack depth migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1027–1030.

Spectral analysis of the non-proliferation experiment

Sjoerd de Ridder

ABSTRACT

To monitor a UN Test Ban Treaty, the US Department of Energy conducted a 1.5 kiloton chemical explosion at the Nevada Test Site, named the Non-Proliferation Experiment (NPE). I study a rarely known recording of the seismic waves of the NPE on a dense, transversely oriented array; the recording contains an extraordinary coda. The spectral information in this wavefield is analyzed to infer information about the subsurface. To first order, the energy arrives isotropically at the array. This allows for interferometric reconstruction of the direct wave between the receivers in the array.

INTRODUCTION

The US Department of Energy detonated a 1.5 kiloton chemical explosive charge at the Nevada Test Site at 12:01am on the 22nd of September 1993. This experiment, named the Non-Proliferation Experiment (NPE) was conducted in anticipation of the Comprehensive Test Ban Treaty approved by the United Nations in September 1996. The measurement was recorded by over 50 broadband seismic stations on the western United States (Tinker and Wallace, 1997). Scientists used these recordings and other measurements to learn to distinguish between nuclear and chemical explosions (Carr, 1994).

A less known and never published data set was recorded by the Subsurface Exploration Company of Pasadena, CA, which operated a 610-channel petroleum-exploration seismic array approximately 200 km distant in Railroad Valley (1). This sign-bit recording equipment was activated at midnight and recorded an extraordinary coda of over 10 minutes. This study reports an analysis of the first continuous 131 seconds of the recording, containing the first break and most of the coda energy.

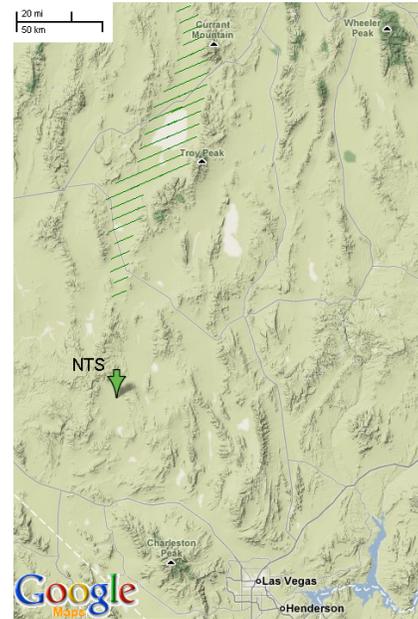
The incoming waves are spatially coherent at early times, when all the energy comes in as one plane wave; while the incoming waves in the coda display a progressively more diverse range of angles. We will investigate how to use the information buried in a recording of such a distant source, by using processing techniques like cross-correlation.

DATA SET

The array was located approximately transverse to the direction of propagation. The station separation was 45 feet, and the array spans a length of over 7 km. The first 56 channels were cross-line to the other 553 stations oriented at N75E, and one channel was dead. There are 5 sign-bit recordings of 2^{14} samples at a sampling rate of 125Hz. There is less than a one second gap between each of the 5 recordings. The tape file was converted to an SEPLib file, and the above information was used to sort the headers. Two additional traces of

the first recording were not found in the tape file, three traces were muted because their signals showed errors. Figures 2 to 5 contain time windows of the first recording for 512 stations (receivers 67 to 568). In a layered earth, all 609 channels would exhibit almost the same signal, because the array is approximately transverse to the direction of propagation. Clearly, from the difference in the observed signals along the array, the data set contains information about the lateral velocity variations in the earth.

Figure 1: Map showing the location of Nevada Test Site (green arrow) and Railroad Valley (green shaded area). [NR] [sjjoerd1/. nte-rrv-map](https://www.researchgate.net/publication/312111111)



SPECTRAL FILTERING OF THE RAW DATA

We first investigate the data in the frequency-wavenumber domain. The first break appears as a narrow line of energy through the origin, see left panel in Figure 6. It represents the spatial coherency and frequency content of the first arrivals that arrive directly from the source with limited multiple scattering in the medium. At later times scattered energy starts arriving from a compact range of angles, seen in Figures 6 and 7. These form a triangle containing all apparent slownesses of plane waves traveling along the array. The highest slownesses are of the plane waves that travel purely along the receiver array. This angle is measured to be associated with an apparent velocity of $\frac{\omega}{k} = 770$ m/s. We apply a Butterworth low-pass filter on the angles in the frequency-wavenumber domain to filter for apparent velocities lower than $\frac{\omega}{k} = 770$ m/s. In addition, we apply a Butterworth low-pass filter to filter energy higher than 15 Hz. This process interpolates the zeroed traces and improved the overall appearance of the coda.

INTERFEROMETRIC GREEN'S FUNCTION RETRIEVAL

Seismic interferometry is a novel geophysical tool to generate Green's functions by cross-correlations of the recorded background coda. Claerbout (1968) showed how the reflec-

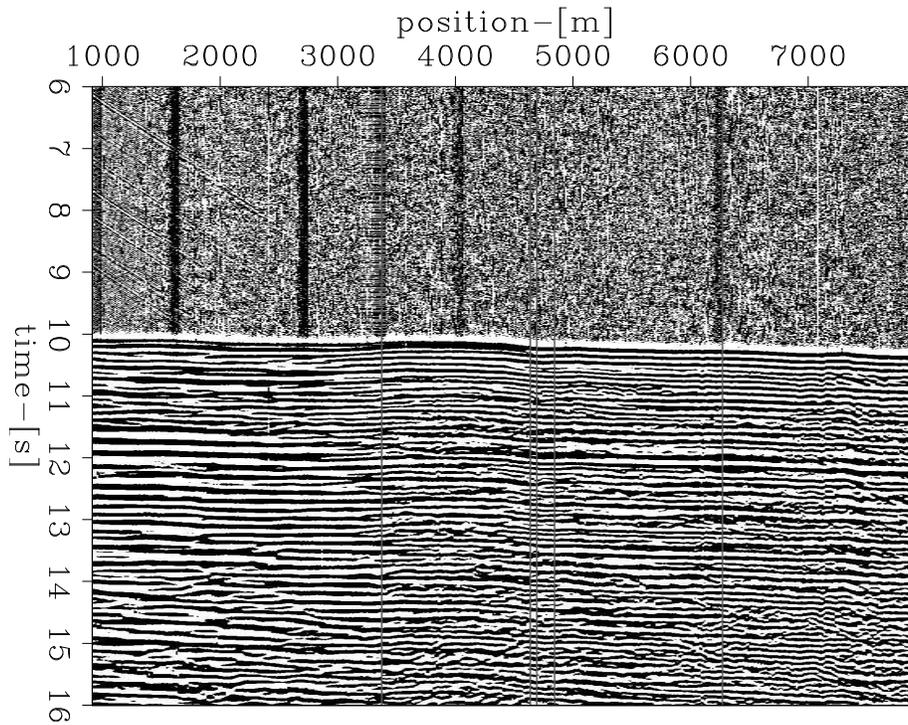


Figure 2: Black-white image of 10 seconds containing the first arrival. The arrivals are very coherent along the array. [ER] sjoerd1/. shot1a

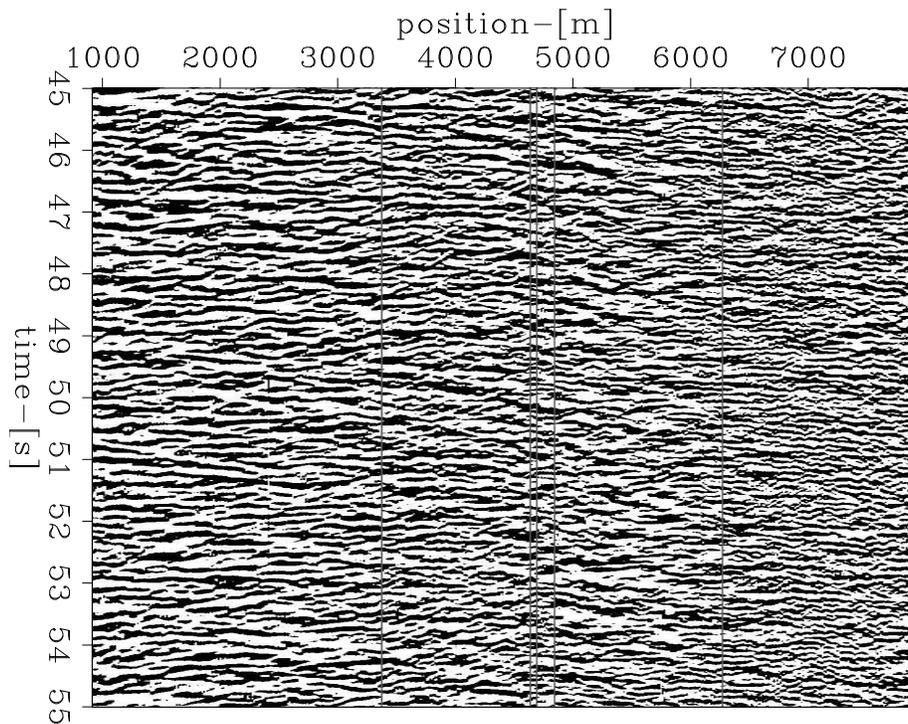


Figure 3: Black-white image of 10 seconds, 35 seconds after the first arrival. Events are incident from more angles than from purely transverse to the array. [ER] sjoerd1/. shot1b

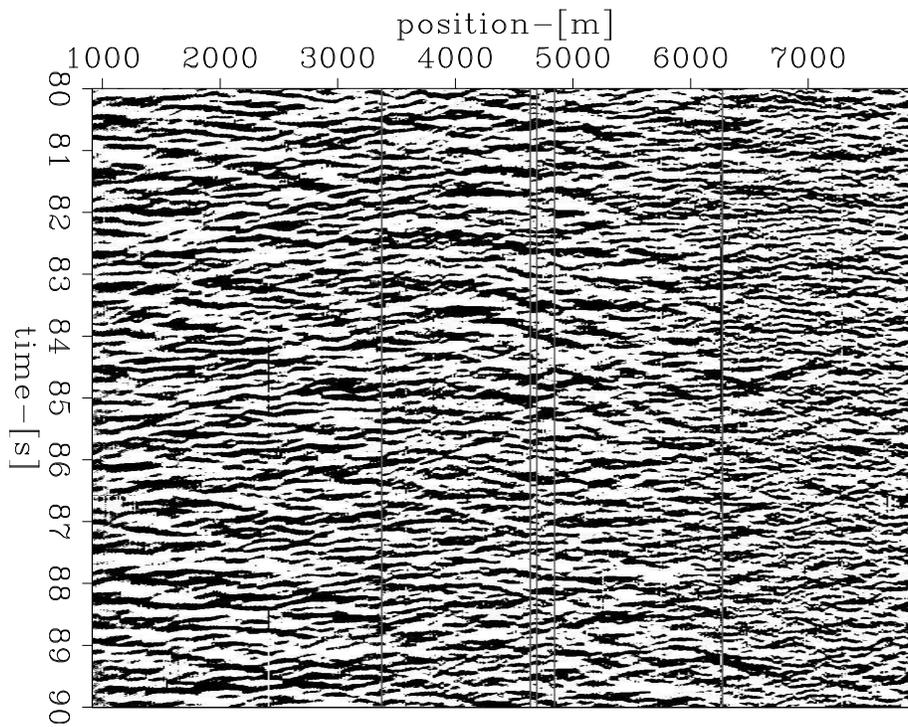


Figure 4: Black-white image of 10 seconds, 70 seconds after the first arrival. The high-frequency content becomes weaker than in Figure 2. [ER] sjoerd1/. shot1c

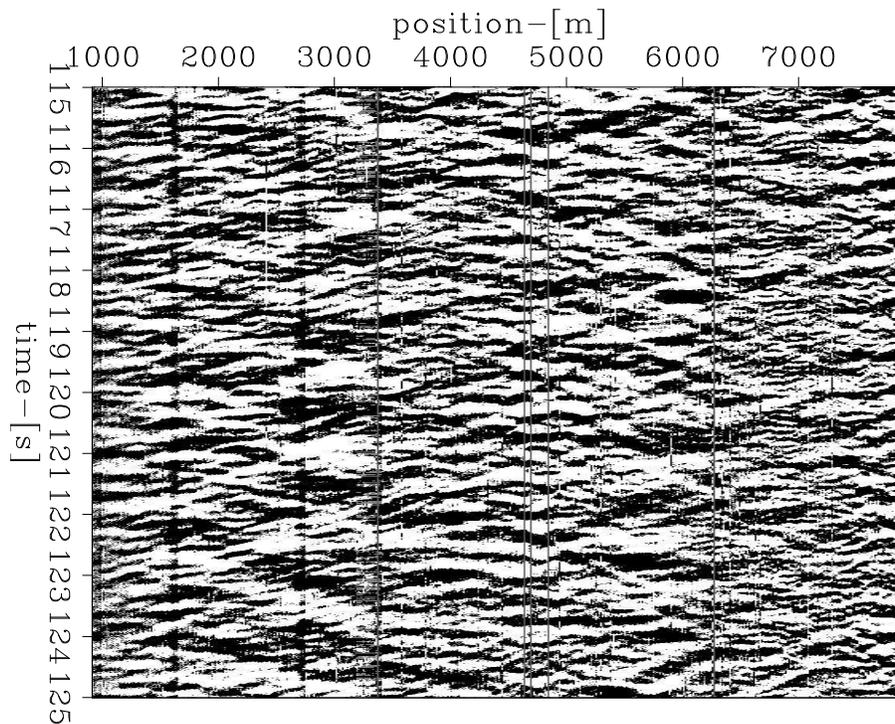


Figure 5: Black-white image of 10 seconds, 105 seconds after the first arrival. The high-frequency content becomes much weaker than in Figure 2. [ER] sjoerd1/. shot1d

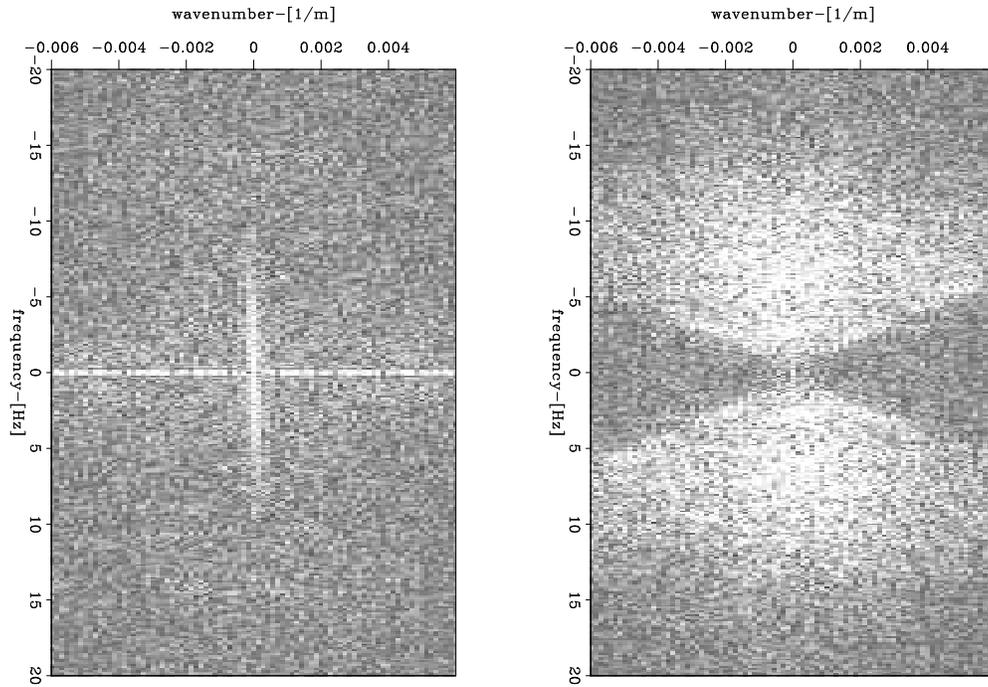


Figure 6: Frequency-wavenumber domain energy spectra of 4 second windows from the frames in Figures 2 and 3. Left panel: The spatial coherence along the array of the first arrival forms a delta function in space. Right panel: Already after 35 seconds a clearly defined cone shape is formed in slowness space. [ER] sjoerd1/. rec1wk1

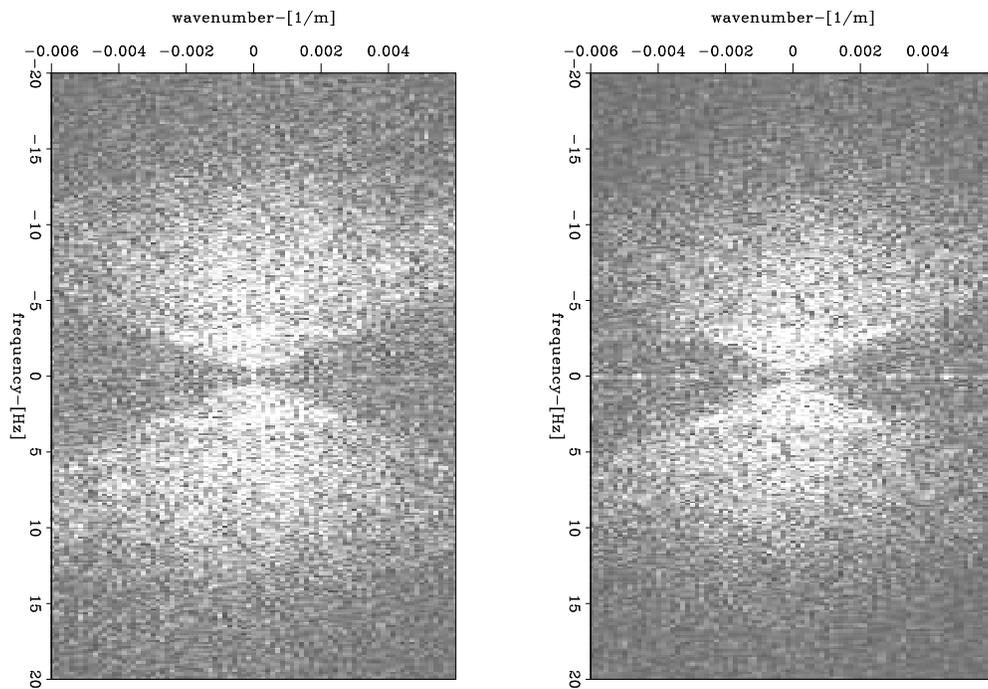


Figure 7: Frequency-wavenumber domain energy spectra of 4 second windows from the frames in Figures 4 and 5. Left panel: The cone persists for later times. Right panel: The power of the high-frequency content decreases. [ER] sjoerd1/. rec1wk2

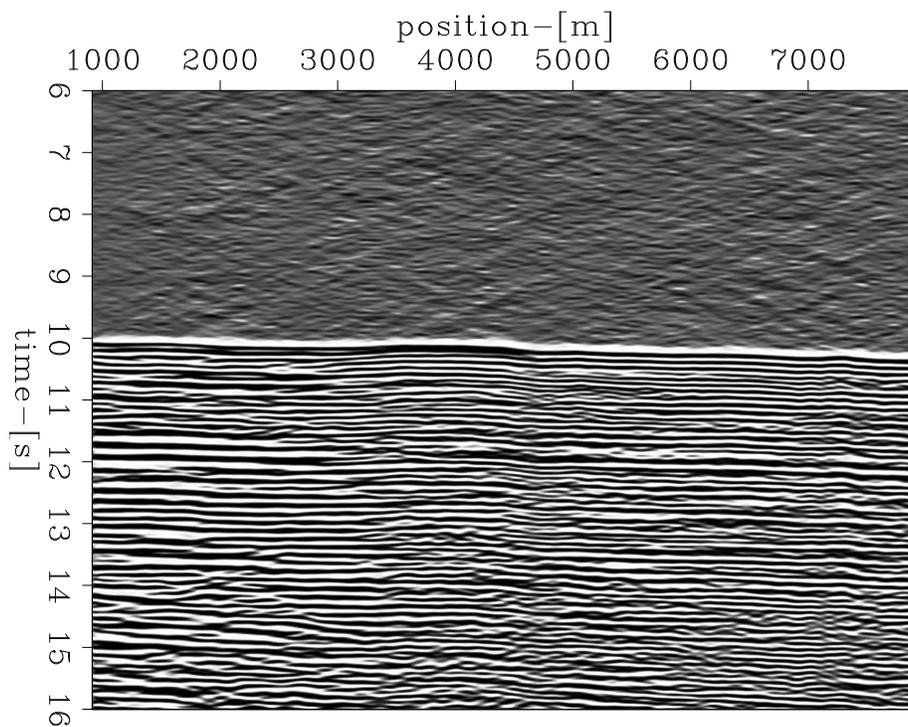


Figure 8: Gray-scale image of 10 seconds of the frequency-wavenumber filtered data containing the first arrival, compare to Figure 2. [ER] sjoerd1/. shot1e

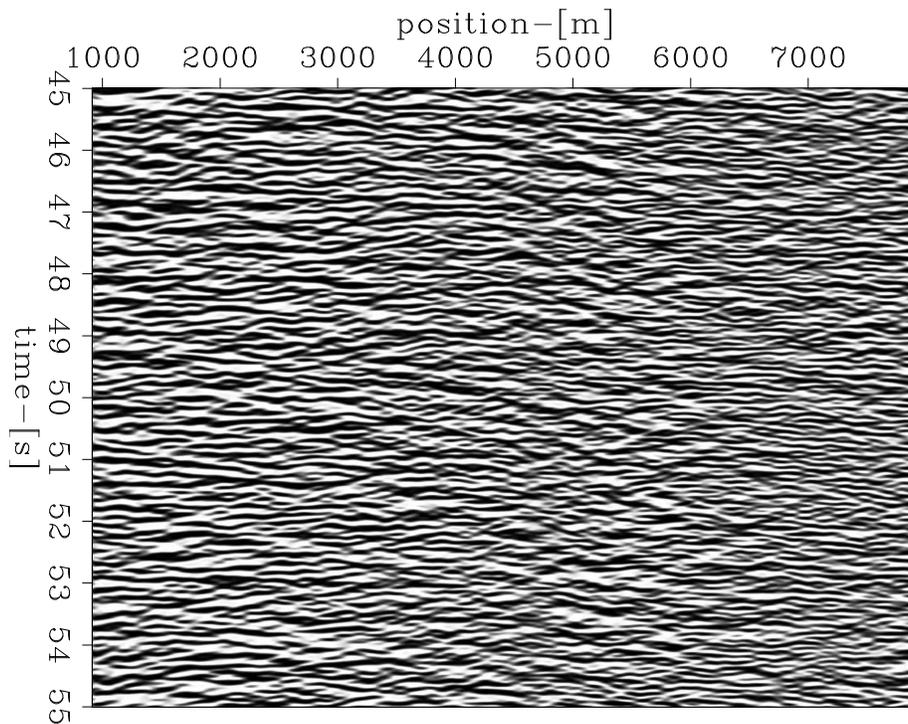


Figure 9: Gray-scale image of 10 seconds of the frequency-wavenumber filtered data, 35 seconds after the first arrival, compare to Figure 3. [ER] sjoerd1/. shot1f

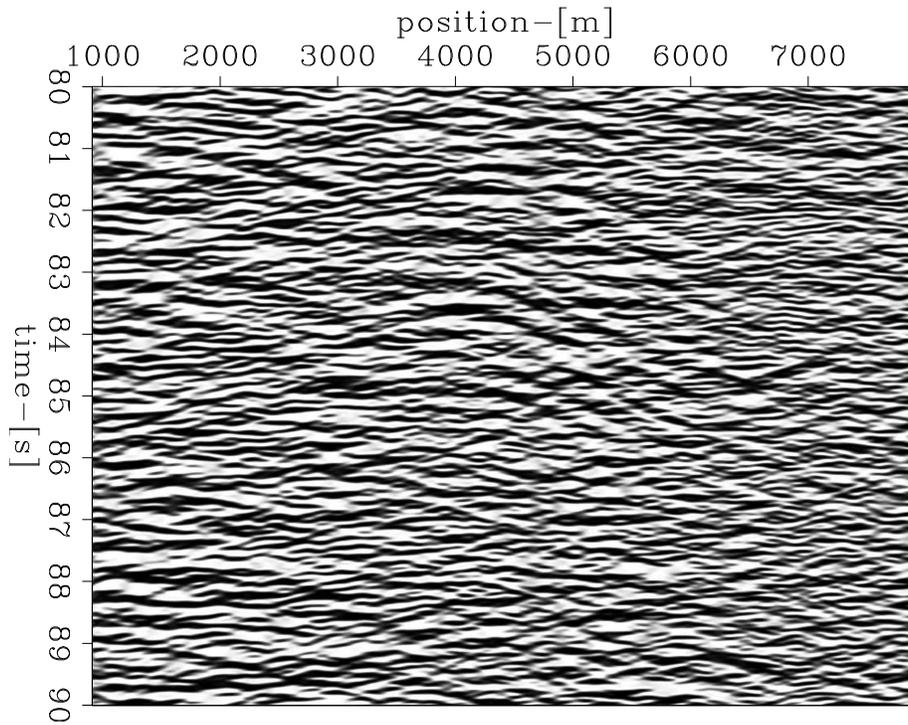


Figure 10: Gray-scale image of 10 seconds of the frequency-wavenumber filtered data, 70 seconds after the first arrival, compare to Figure 4. [ER] sjoerd1/. shot1g

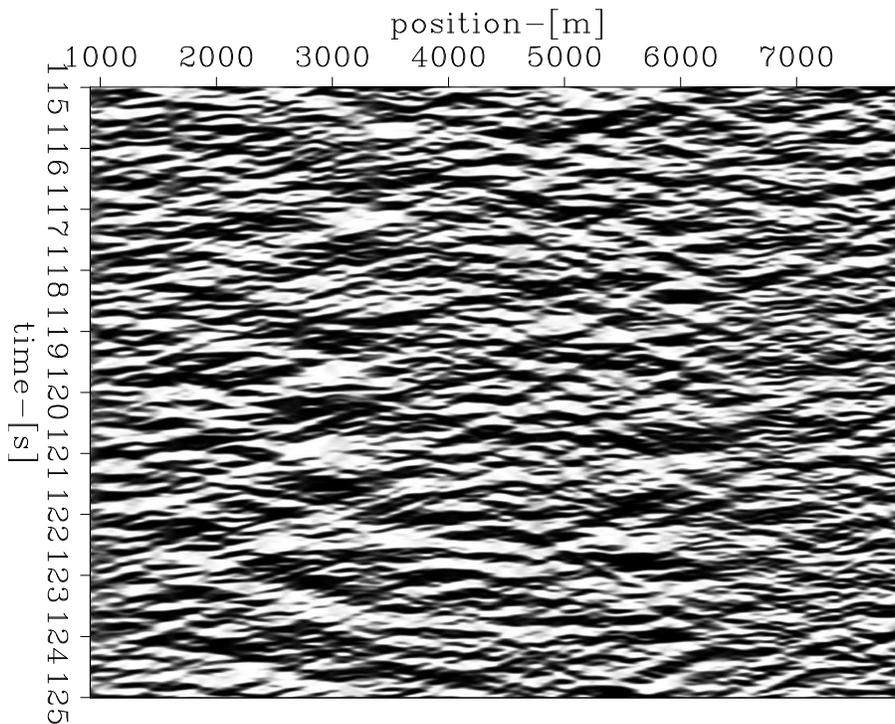


Figure 11: Gray-scale image of 10 seconds of the frequency-wavenumber filtered data, 105 seconds after the first arrival, compare to Figure 5. [ER] sjoerd1/. shot1h

tion response of a 1D medium can be retrieved from its transmission response. A general derivation for arbitrarily anisotropic and inhomogeneous media can be made through the reciprocity theorem of the time-correlation type (Wapenaar, 2003, 2004; Wapenaar and Fokkema, 2006). A reciprocity theorem interrelates two independent excitation states of a physical system in one and the same domain \mathbb{D} (de Hoop, 1966; Fokkema and van den Berg, 1993). The key is to apply source-receiver reciprocity relations for the fields in the reciprocity theorem of the time-correlation type. Other derivations are based upon the diffusive character of the wavefield, (Weaver and Lobkis, 2001, 2002; Shapiro and Campillo, 2004; Shapiro et al., 2005; Roux et al., 2005). The 3D equivalent of the 1D relation of Claerbout (1968), between the transmission coefficient T and the reflection coefficient R , proven by Wapenaar (2003), is

$$\hat{R}(\mathbf{x}_r, \mathbf{x}_s, \omega) + \hat{R}^*(\mathbf{x}_s, \mathbf{x}_r, \omega) = - \oint_{\partial\mathbb{D}} \hat{T}(\mathbf{x}_r, x, \omega) \hat{T}^*(\mathbf{x}_s, x, \omega) d^2x + \delta(\mathbf{x}_s - \mathbf{x}_r) \quad (1)$$

$$\approx \langle \hat{T}(\mathbf{x}_r, \dots, \omega) \hat{T}^*(\mathbf{x}_s, \dots, \omega) \rangle \quad (2)$$

written in the frequency domain, where ω denotes angular frequency, $*$ denotes complex conjugation and $\langle \cdot \rangle$ denotes a time-average. The receiver station locations at \mathbf{x}_r and \mathbf{x}_s acquire the meaning of receiver and virtual source position in the retrieved reflection coefficients. The integral on the right-hand side of Equation 1 require individual recordings of sources positioned on a domain boundary $\partial\mathbb{D}$, enclosing the receiver stations and all heterogeneity. In ideal conditions this could be met by a time-averaging of recordings of sources, acting uncorrelated and positioned on the domain boundary.

The left-hand side of Equation 1 will be referred to as the retrieved Green's functions. We retrieve the superposition of a causal Green's function between \mathbf{x}_s and \mathbf{x}_r and the time reversed reciprocal Green's function between \mathbf{x}_r and \mathbf{x}_s . In practice, we would recover these Green's functions convolved with the auto-correlation of an average source signal emitted by the source distribution (Wapenaar and Fokkema, 2006). The requirement of plane waves arriving isotropically from all directions is satisfied, to the first order. In practice not all contributions will be equally important: sources at stationary angles contribute dominantly to the retrieved result (Snieder, 2004; Snieder et al., 2006). For the recording of the NPE, we do not satisfy the conditions outlined above. Thus the retrieved Green's functions will be infested with spurious events with arrival times of non-existing interfaces (Dragonov et al., 2004; Snieder et al., 2006).

SYNTHESIS OF AN INTERFEROMETRIC SHOT GATHER

Although the NPE probably emits a spike-like source, the incident source spectrum is incident at the array from a wider range of angles, after multiple scattering, is unknown. After cross-correlation the receiver signals are deconvolved with an estimated source signal. The power spectra of the received signals are estimated using multitaper spectrum analysis codes from a library by Prieto et al. (2008). The procedure is as follows. The data is multiplied with a chosen number of Slepian tapers of increasing order; then a weighted average of the Fourier transformation is computed for each of the tapered data copies, resulting in a low variance estimate, while maintaining high resolution (Thomson, 1982; Prieto et al., 2007). One receiver is chosen as the 'master' receiver, and all other receivers are cross-correlated with this master, effectively turning this master receiver into a synthetic

source. The source spectrum is estimated on the trace of the 'master' receiver using the multitaper estimation technique.

An example, choosing receiver number 400 as 'master' of a computed interferometric synthetic shot gather, is shown in the left panel of Figure 12. The first break arrives almost simultaneously at all receivers, and after cross-correlation appears as a strong spike with zero time lag. High frequencies are introduced because of the division with the estimated source spectrum. I apply a Butterworth low-pass filter to remove frequencies higher than 12 Hz, and the result is shown in the right panel of Figure 12. There are some interesting flat events at early and later times. These are however not appearing in the shot gathers synthesized for receivers further away from receiver number 400. We also retrieve a little bit of energy in the direct wave.

Most of the coherent energy between receiver stations along the array is associated with the direct wave from the synthetic shot to the receivers. We cannot identify any subsurface reflections. We see a spurious source at the right side of the synthetic shot gather. This could be caused by the correlation of the first break with the reflections from the side of the valley. We approximate our medium to have a 1D velocity structure, and we stack all the interferometric shot gathers that can be computed, by in turn selecting all 512 receivers as a 'master', with constant offset. All events from non-flat boundaries in the subsurface will destructively interfere. Thus we obtain a synthetic super-gather, in which only the direct arrival can be distinguished with certainty. The moveout of this event is estimated as approximately 1000 m/s, this is a reasonable velocity for ground roll.

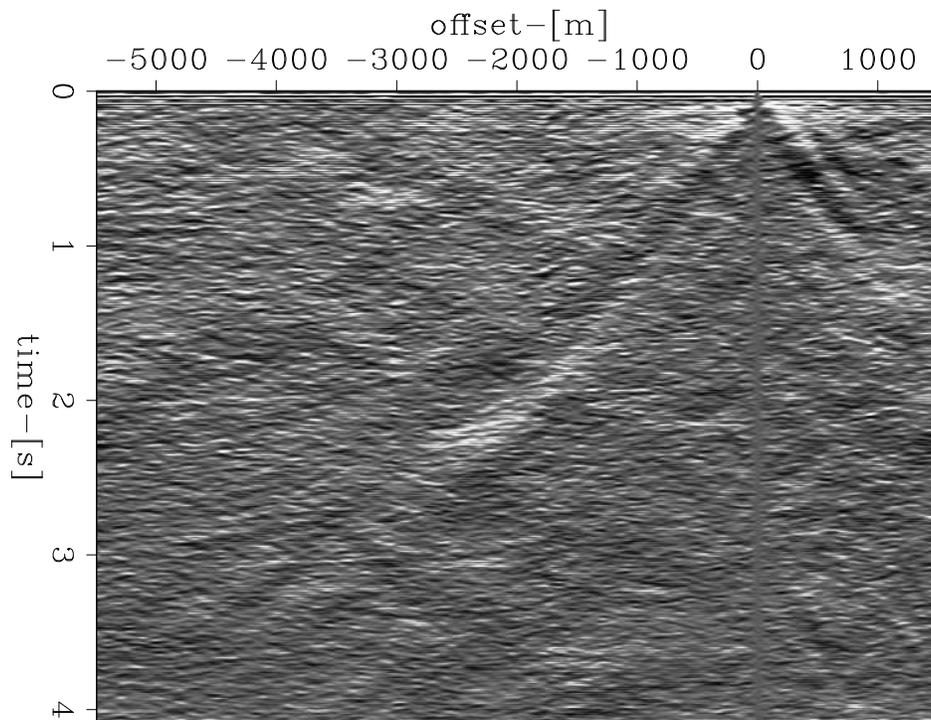


Figure 12: Interferometric shot gather with receiver number 400 as the master receiver.

[CR] `sjoerd1/. mtdeco400`

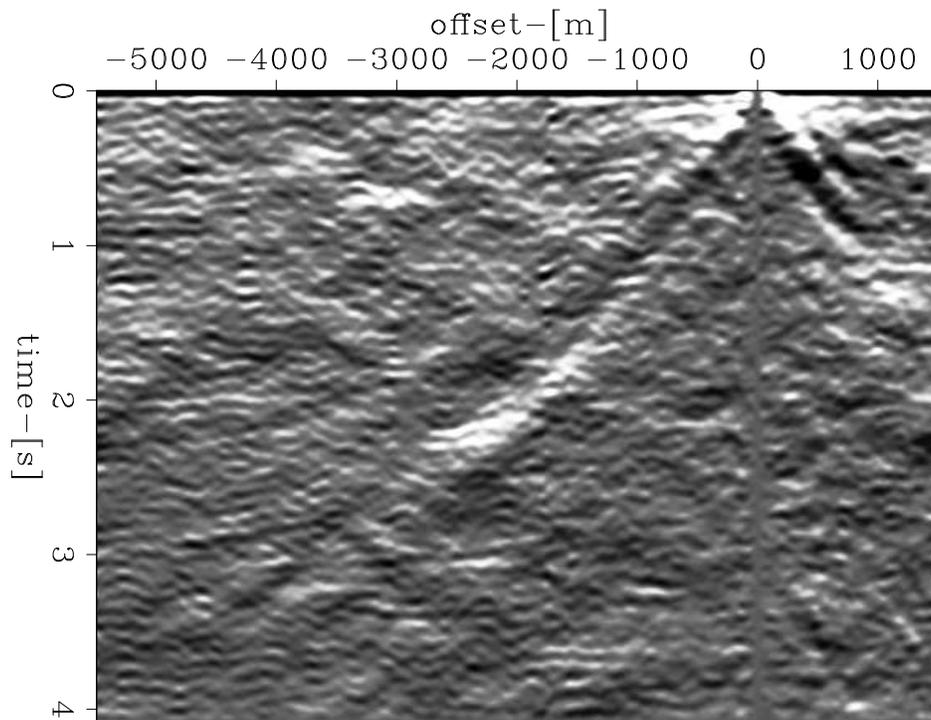


Figure 13: Interferometric shot gather with receiver number 400 as master receiver, low-pass Butterworth filtered for 12 Hz. [CR] sjoerd1/. mtdeco400F

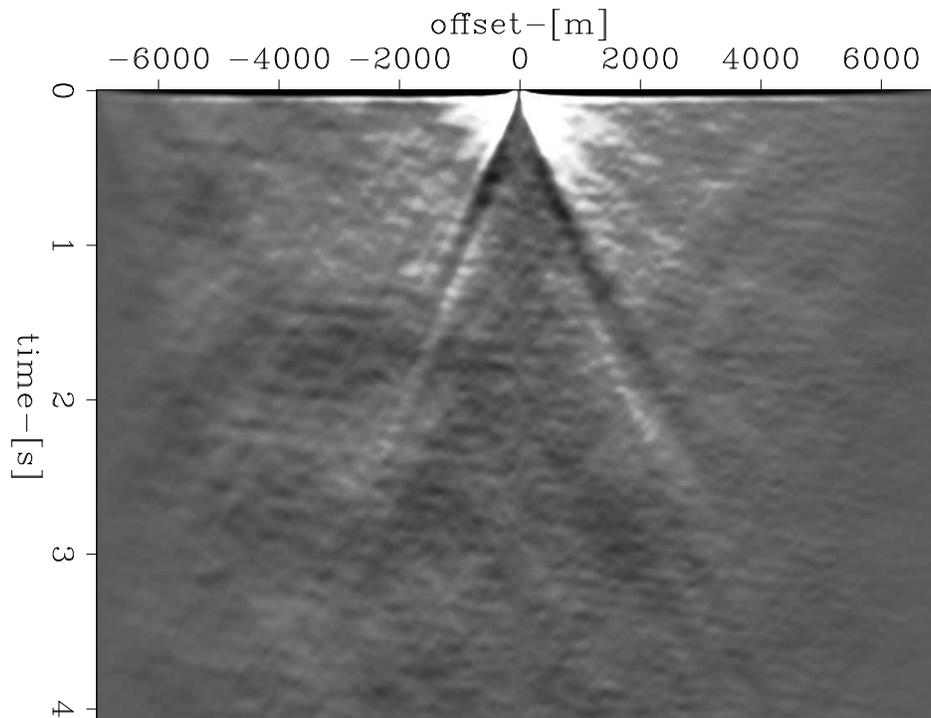


Figure 14: Interferometric super-gather, obtained by stacking cross-correlations of all possible receiver couples over common offsets. [CR] sjoerd1/. mtdeco-stack

CONCLUSIONS

We use the coda of a distant recording of the NPE containing multiple scattered waves incident at the array from a wide range of angles. We deconvolve for an estimated source spectrum. We discover that to first order, the waves are incident transverse isotropically to the array. The coherent energy between receivers is dominantly associated with the direct wave between receivers. To retrieve subsurface reflections, we need to sample the stationary phases of subsurface interfaces. For the direction of the array, these would lie approximately inline with and beneath the array, this is not in the dominant direction of propagation by the waves from the NPE.

ACKNOWLEDGMENTS

The author thanks Germán Prieto for many discussions on Green's function retrieval and for providing support for his multitaper spectrum analysis Fortran library, Jon Claerbout for discussions and encouragement to investigate this data set, Bob Clapp for writing the C routines that read the tape file, John Vidale for supplying the data tape file.

REFERENCES

- Carr, D. B., 1994, Non-proliferation experiment recorded at the pinedale seismic research facility: Sandia National Laboratories.
- Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: *Geophysics*, **33**, 264–269.
- de Hoop, A. T., 1966, An elastodynamic reciprocity theorem for linear, viscoelastic media: *Appl. Sci. Res.*, **16**, 39–45.
- Dragonov, D., C. P. A. Wapenaar, and J. Thorbecke, 2004, Passive seismic imaging in the presence of white noise sources: *The Leading Edge*, **23**, 889–892.
- Fokkema, J. T. and P. M. van den Berg, 1993, *Seismic applications of acoustic reciprocity*: Elsevier.
- Prieto, G. A., R. L. Parker, and F. L. Vernon, 2008, A fortran 90 library for multitaper spectrum analysis. Submitted to *Computers and Geosciences*.
- Prieto, G. A., D. J. Thomson, F. L. Vernon, P. M. Shearer, and R. L. Parker, 2007, Confidence intervals of earthquake source parameters.
- Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux, 2005, Ambient noise cross correlation in free space: Theoretical approach: *JASA*, **117**, 79–84.
- Shapiro, N. M. and M. Campillo, 2004, Emergence of broadband rayleigh waves from correlations of the ambient seismic noise: *Geophys. Res. Lett.*, **32**, L07614–1–L07614–4.
- Shapiro, N. M., M. Campillo, L. Stehly, and M. R. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: *Science*, **307**, 1615–1618.
- Snieder, R., K. Wapenaar, and K. Larner, 2006, Spurious multiples in seismic interferometry of primaries: *Geophysics*, **71**, SI111–SI124.
- Snieder, R. K., 2004, Extracting the green's function from the correlation of coda waves: A derivation based on stationary phase: *Phys. Rev. E*, **69**, 046610–1 – 046620–8.
- Thomson, D. J., 1982, *Spectrum estimation and harmonic analysis*.
- Tinker, M. A. and T. C. Wallace, 1997, Regional phase development of the non-proliferation experiment within the western united states.

- Wapenaar, K., 2003, Synthesis of an inhomogeneous medium from its acoustic transmission response: *Geophysics*, **68**, 1756–1759.
- , 2004, Retrieving the elastodynamic green's function of an arbitrary inhomogeneous medium by cross correlation: *Phys. Rev. Lett.*, **93**, 254301–1 – 254301–4.
- Wapenaar, K. and J. Fokkema, 2006, Green's function representations for seismic interferometry: *Geophysics*, **71**, SI33–SI46.
- Weaver, R. L. and O. I. Lobkis, 2001, Ultrasonics without a source: thermal fluctuation correlations at mhz frequencies.: *Phys. Rev. Lett.*, **87**, 134301–1 – 134301–4.
- , 2002, On the emergence of the green's function in the correlations of a diffuse field: pulse-echo using thermal phonons: *Ultrasonics*, **40**, 435–439.

Image segmentation for velocity model construction and updating

Adam Halpert and Robert G. Clapp

ABSTRACT

Image segmentation can automatically delineate salt bodies in seismic data, an otherwise human-intensive and time-consuming task. In many instances, current segmentation algorithms successfully pick salt boundaries; a logical extension of such work is to apply these methods to the task of building and updating seismic velocity models. We apply image segmentation tools in conjunction with sediment- and salt-flood velocity estimation techniques to identify the top and base of a salt body. Furthermore, previously existing velocity models may be updated based on the results of segmentation and automated boundary picking. By using the existing model as a priori information for the picking algorithm in areas where the segmentation is ambiguous, we calculate an optimized boundary path across a seismic image. For both synthetic and real seismic data, migrations with velocity models derived from this method produce greatly improved images.

INTRODUCTION

The purpose of image segmentation is to automatically divide an image into sections based on specific attributes. Because of its global optimization properties, Normalized Cuts Image Segmentation (Shi and Malik, 2000), or NCIS, is one algorithm with a variety of potential applications to seismic interpretation. The NCIS method was first applied to seismic data by Hale and Emanuel (2002, 2003) to paint 3D atomic meshes of seismic images. Recent work (Lomask, 2007; Lomask et al., 2007) presents an image segmentation algorithm for automatic picking of salt boundaries on migrated seismic images. Such a scheme offers many potential benefits for the seismic velocity model building process. In this paper, we adapt Lomask's methods for use in iterative velocity model construction and updating.

A reasonably accurate velocity model is an essential component of the seismic imaging process. Much of today's seismic data is acquired in regions with complex salt bodies; in such cases, clearly delineating salt interfaces is often one of the most human-intensive, time consuming and ambiguous aspects of velocity estimation. Accurate salt interface interpretation becomes especially important when the imaging target is located beneath a salt body, as is often the case for modern surveys. The method we propose here is designed to function as a tool for either velocity model construction or updating. We show our method to be highly effective when combined with widely used sediment- and salt-flooding migration techniques to make original salt interface interpretations. A current limitation of the segmentation algorithm is its ability to handle uncertainties along the boundary path; the same picking rules are imposed throughout the image, even if some areas would benefit from using different rules. Therefore, we also propose a means of producing

an improved boundary path by solving a global, non-linear optimization problem. We employ this optimized boundary to update an existing velocity model, so the optimization scheme incorporates the original model as prior information about the boundary in areas of uncertainty. We show both real and synthetic data examples of this method producing an improved velocity model.

THEORY

The segmentation algorithm we use (Lomask, 2007; Lomask et al., 2007) divides a seismic image into two segments, salt and sediment; the boundary separating them is the salt interface. The segmentation is based on a specific seismic attribute, in most cases amplitude of the envelope, that can clearly differentiate between a salt body and the surrounding sediments. The first step is to create a weight matrix, \mathbf{W} , relating each pixel in a migrated seismic image to a collection of neighboring pixels. Thus, \mathbf{W} is necessarily many times larger than the input image, although Lomask (2007) describes several methods to improve computational efficiency. One such method is to assign weights for only a random sampling of neighbors for each pixel, rather than all pixels within a given distance. The weight associated with a pair of pixels is inversely proportional to the maximum value of the envelope along the shortest path connecting those two pixels, multiplied by a distance weighting term. Thus, low weights are assigned to pixel pairs most likely to be separated by a salt boundary. Subject to constraints, the path across the image minimizing the sum of the weights through which it passes is the salt boundary.

Following the NCIS algorithm of Shi and Malik (2000), Lomask showed that the determination of a boundary path across a seismic image may be set up as an eigenvector problem via the Rayleigh quotient

$$\min_{\mathbf{y}} \frac{\mathbf{y}^T(\mathbf{D} - \mathbf{W})\mathbf{y}}{\mathbf{y}^T\mathbf{D}\mathbf{y}}, \quad (1)$$

where \mathbf{y} is the eigenvector and \mathbf{D} is a diagonal matrix whose elements are the sum of each column of \mathbf{W} . Because of constraints introduced on the Rayleigh quotient, it will be minimized by the eigenvector corresponding to the second smallest eigenvalue of the eigensystem

$$(\mathbf{D} - \mathbf{W})\mathbf{y} = \lambda\mathbf{D}\mathbf{y}, \quad (2)$$

where λ is the eigenvalue. The eigenvector \mathbf{y} will have values ranging from -1 to 1 across the boundary; in most cases, following the “zero contour” across the image yields the most appropriate salt interface pick. When the zero contour is not clearly defined (see Figure 1), it may be necessary to use another value of the eigenvector to pick the boundary. The determination of this value in different parts of the image can be posed as an optimization problem, and is discussed later.

A note about migration

Because segmentation operates on migrated data, the success of this method is highly dependent on the migration algorithms used and the quality of the resulting images. For the examples in this paper, we elected to use plane wave migration in tilted Cartesian

coordinates (Shan and Biondi, 2007; Shan, 2008). Figure 1 shows eigenvectors calculated for the same section of a Gulf of Mexico dataset. The eigenvector in (a), taken from Lomask (2007), is derived from an image obtained with a one-way wave-equation method; for (b), the image used is obtained through Shan’s plane wave migration. The eigenvector in (b) is clearly superior, as the transition from positive to negative values (white to black) is sharp for more of the image than on the left. This implies reduced uncertainty in the boundary pick.

Migration in tilted coordinates also achieves better results than plane wave migration in conventional coordinates. Figure 2 displays a portion of the Sigsbee dataset that will be used throughout this paper. Both images in the top row are obtained through plane wave migration with a perfect velocity model, either in a regular coordinate system (a), or with tilted coordinates (b). The right-hand image is superior, most noticeably along the steeply dipping walls of the salt canyon. The corresponding eigenvectors, shown in (c) and (d), again feature less path uncertainty for the eigenvector derived from the superior image.

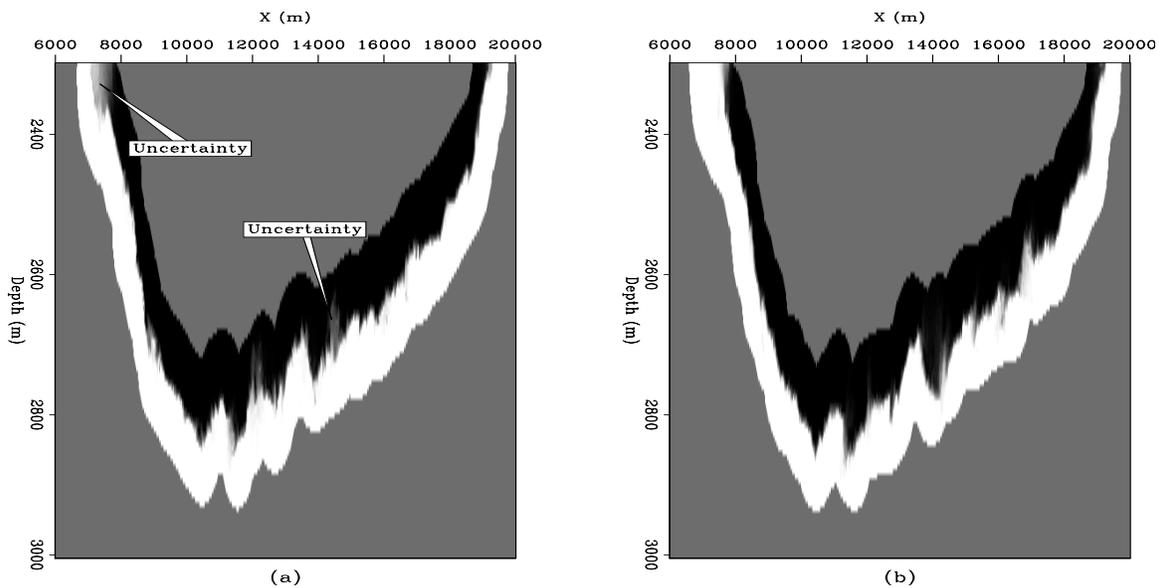


Figure 1: Eigenvectors derived from images obtained through two different migration algorithms: a conventional downward continuation method (a), and plane wave migration in tilted coordinates (b). [CR] adam1/. eigs

VELOCITY MODEL CONSTRUCTION

A typical procedure for locating salt boundaries in seismic data is to perform migrations using “flood” velocity models. First, migrate strictly with sediment velocities throughout the entire section; theoretically, the top salt interface will be well resolved. A second migration with salt velocities flooded below the top boundary pick resolves the base of the salt body. In both cases, manually picking the salt boundaries can be time-consuming and inexact, especially for salt bodies with complex geometries. Here, we show that image segmentation can greatly expedite this process.

For this example we use a portion of the Sigsbee 2a synthetic dataset, shown in Figure 2

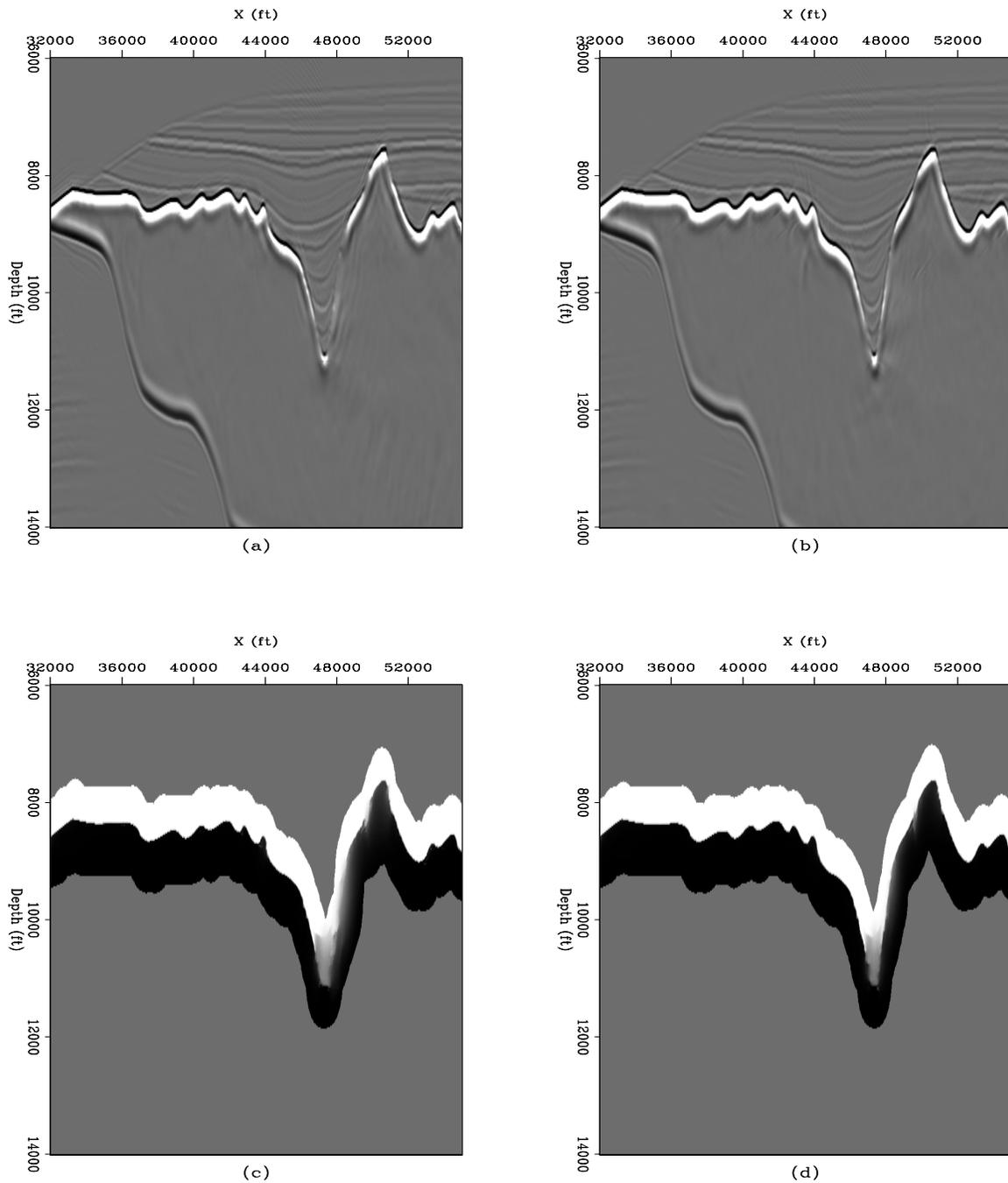
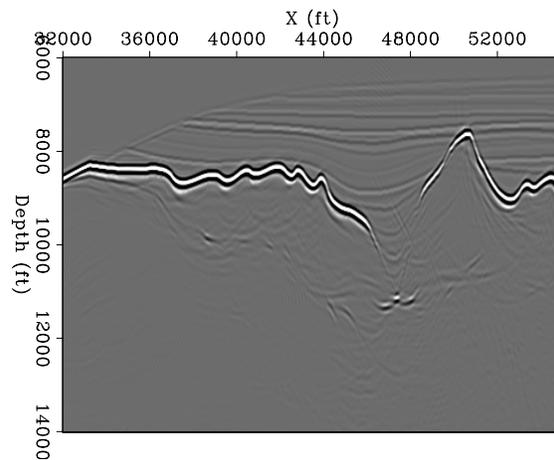


Figure 2: Top: A section of the Sigsbee synthetic model imaged using plane wave migration in (a) a regular coordinate system and (b) tilted Cartesian coordinates. Bottom: Eigenvectors corresponding to the images directly above. Note both the improved imaging of the salt canyon walls, as well as the reduced eigenvector path uncertainty when tilted coordinates are used. [CR] `adam1/. tilt4`

after migration with a perfect velocity model. Figure 3 shows the result of migration with a sediment-flooded velocity model, created by infilling the salt portions of the Sigsbee model with sediment velocities. Much of the top salt interface is extremely well resolved and relatively easy to pick; however, the boundary disappears inside the salt canyon, and would be difficult to pick manually. The top row of Figure 4 illustrates the image segmentation process: (a) is the eigenvector calculated for the image, and (b) is the boundary pick corresponding to the zero contour of the eigenvector. The method accurately picks the complex geometry along the top of the salt, and correctly interprets the presence of the canyon.

Figure 3: A sediment flood migration of the Sigsbee data. Most of the top salt boundary is well resolved, but the details of the salt canyon are highly ambiguous. [CR]

adam1/. mig-sedflood



One way to measure the quality of the method's top boundary pick is to observe the quality of a salt flood migration using a velocity model derived from the top boundary pick. Figure 4c shows the eigenvector resulting from an image migrated with salt velocities flooded below the top boundary pick. There is little to no uncertainty along the zero contour. As expected, the base salt reflection in image (d) is strong and clear throughout the image; the overlain boundary confirms that the segmentation algorithm tracks the interface well.

VELOCITY MODEL UPDATES

Image segmentation may also be used as a tool in an iterative boundary-picking process to update a previously existing velocity model or salt body interpretation. For this task, such a preexisting model offers a potential advantage over picking a boundary without any prior information, as was the case for the sediment- and salt-flood models. Namely, the existing model can act either as a guide or a penalty for the boundary-picking algorithm in areas where the appropriate boundary choice is not obvious.

Uncertainty in the method's salt interface pick arises when the eigenvector transitions smoothly from negative to positive values. These areas of uncertainty appear gray on depictions of the eigenvector such as Figure 1, rather than sharp transitions from black to white indicative of a relatively certain boundary pick. In a given area of uncertainty, the boundary pick may be improved by manually selecting a non-zero eigenvector contour value to follow throughout the image. However, Lomask (2007) notes that any such improvement in one part of the image may be "matched by a reduction of picking quality in other uncertain areas." Thus, a natural extension of Lomask's work is to pose the boundary picking task as

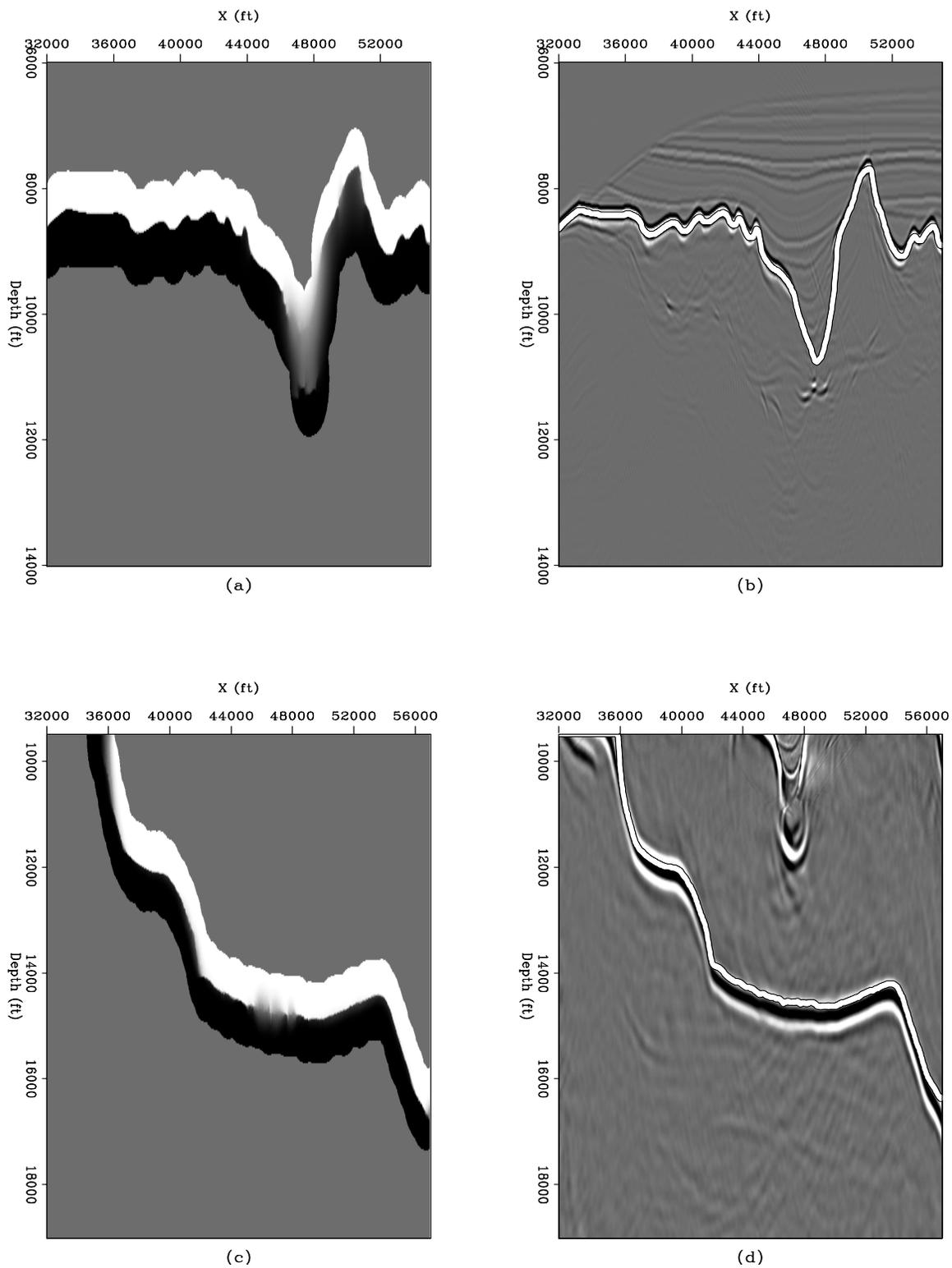


Figure 4: Top: Eigenvector (a) and picked boundary (b) from a sediment-flood migration of the Sigsbee data. Bottom: Eigenvector (c) and picked boundary (d) for the base of the Sigsbee salt body after a salt-flood migration. [CR] adam1/. floods

a global optimization problem, in which the boundary is allowed to follow different contour values throughout the image.

Optimizing the boundary

We set up a non-linear inverse problem that attempts to find the optimal depth for the salt-sediment interface at each \mathbf{x} position in the image. Currently, a limitation of this method is the inability to handle boundaries that are not single-valued functions of \mathbf{x} .

1. In most cases, the eigenvector's zero contour is the most appropriate path. Therefore, a primary fitting goal should seek to follow the zero contour wherever possible. This goal results in a non-linear system; we cannot create a linear system that maps between reflector depth and the eigenvector. As a result, we must create a linear operator, \mathbf{G} , by linearizing the problem around the zero-contour boundary \mathbf{m}_0 . First, we calculate the depth gradient at each point for every trace in the eigenvector. The operator \mathbf{G} is then formed by extracting the gradient value at each point along the zero-contour boundary \mathbf{m}_0 . In equation 3 below, the deviation $\Delta\mathbf{m}$ from the zero-contour boundary is very small when \mathbf{G} is large (high certainty), and larger when \mathbf{G} is small (uncertainty).
2. The zero contour may be inappropriate in areas of great uncertainty. Here, it is beneficial to rely more upon the a priori information - a previous boundary (\mathbf{m}_{orig}) manually suggested by an experienced interpreter, or the results of the flood migration procedure detailed above. In the former case, the prior boundary can act as a guide; the optimized boundary will tend to follow the previous one in uncertain areas. In the latter case, however, it may be obvious that the prior boundary is placed either too deep or too shallow in the model. In this circumstance, the previous boundary may be penalized so that the optimized boundary will move *away* from the previous one rather than toward it. To implement this goal, we construct a diagonal weighting matrix \mathbf{P} such that areas of uncertainty are given greater weights. The weights in \mathbf{P} may have positive or negative values, depending on the need to either penalize or reward the prior boundary.
3. Finally, to avoid unwanted fluctuations in the boundary pick, a smoothing constraint is imposed on the boundary in the form of a 1D gradient roughening operator (\mathbf{A}).

Mathematically, the above goals may be expressed as a series of minimization equations:

$$0 \approx \mathbf{G} \Delta\mathbf{m} \quad (3)$$

$$0 \approx \mathbf{P}(\mathbf{m} - \mathbf{m}_{\text{orig}}) \quad (4)$$

$$0 \approx \mathbf{A}\mathbf{m} . \quad (5)$$

Because the problem posed here is non-linear, we solve it iteratively using the Newton method to implement the above equations and calculate a new boundary model \mathbf{m}_i . Defining the previous boundary model as \mathbf{m}_{i-1} , and our current model as a function of non-linear iteration j as \mathbf{m}_i^j , and two relative weighting parameters ϵ_1 and ϵ_2 we begin with the zero-contour boundary and iterate to convergence using the following scheme:

```

Iterate over j
{
Construct  $\mathbf{L} = [\mathbf{G}\mathbf{m}_i^j \quad \epsilon_1\mathbf{P} \quad \epsilon_2\mathbf{A}]^T$ 
Calculate current residual  $\mathbf{r} = [0 \quad -\mathbf{P}\mathbf{m}_{i-1} \quad 0]^T - \mathbf{L}$ 
Solve the linear system
 $\mathbf{r} = \mathbf{L}\Delta\mathbf{m}_i^j$ 
 $\mathbf{m}_i^j = \mathbf{m}_i^{j-1} + \Delta\mathbf{m}_i^j$ 
}

```

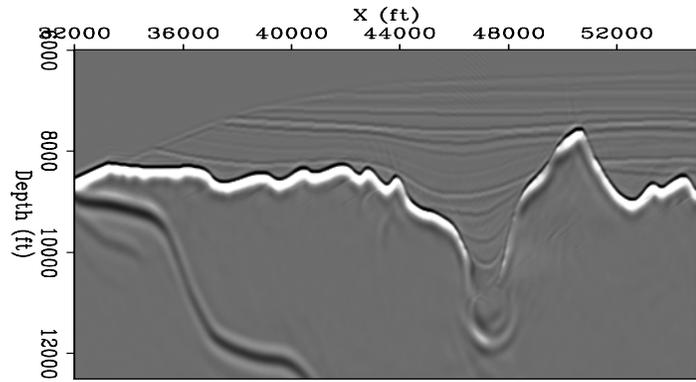
For 2D cases on the scale of the examples shown here, the computational expense for this scheme is virtually negligible. A possible future enhancement to the algorithm presented here will be to incorporate residual map migration to remap the original boundary. This would help eliminate bias introduced by movements of the migrated reflectors with respect to the image obtained with the original velocity model.

Velocity model update: Examples

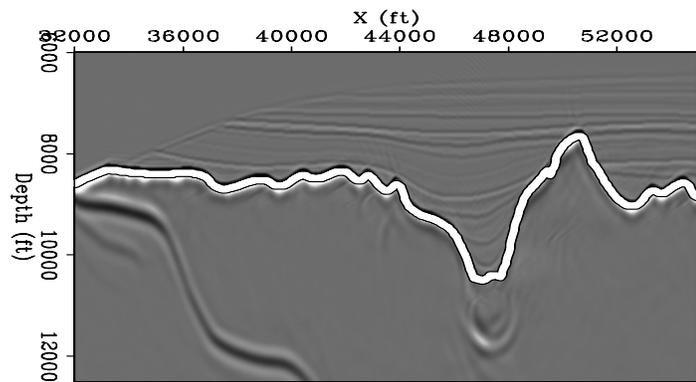
Here, we present updated velocity models for both the section of the Sigsbee synthetic data used previously and a section of the 2D Gulf of Mexico dataset that Lomask (2007) used to demonstrate his segmentation algorithm. Once we calculate an optimum boundary using the procedure detailed above, we generate an updated velocity model based on a comparison with the original velocity model. Any sediment velocities below the picked upper salt boundary are filled with salt velocities, and salt velocities above the boundary pick are replaced with nearby sediment velocities. By remigrating with the updated velocity models, improved images are obtained.

Figure 5 displays a typical sequence for producing an improved image by updating a preexisting velocity model. In this case, the prior boundary model is the salt interface picked after the sediment-flood migration of the Sigsbee data, shown in Figure 4b. It is clear that this boundary places the bottom of the canyon too shallow, as the canyon bottom image is pushed down by the presence of excess salt velocities. Figure 5b confirms that the zero contour boundary would still be picked too shallow at the bottom of the canyon. Therefore, the prior boundary is penalized in the optimization scheme; the optimized boundary will move away from the previous one in uncertain areas (in this case, only the canyon bottom). Figure 5c shows the result of the boundary optimization process - as expected, the bottom of the canyon is now placed deeper in the section. After remigration with a velocity model derived from the new boundary, the resulting image in Figure 5(d) is vastly improved.

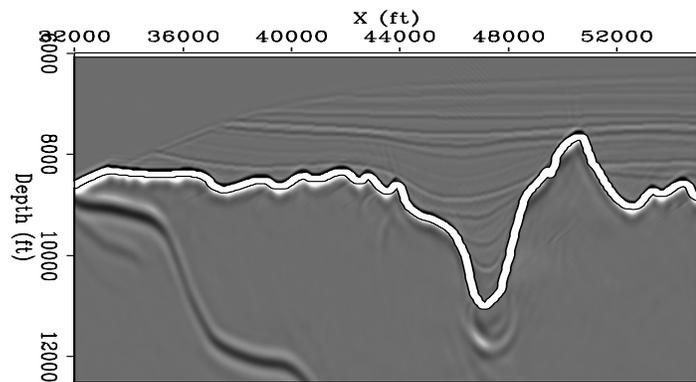
In his thesis work, Lomask used a real Gulf of Mexico dataset provided by WesternGeco to demonstrate his segmentation algorithm. Here, we use the optimized boundary/updated velocity model method to produce an improved image. Figure 7 compares the original and updated velocity models used for migration, and Figure 6 shows the results of boundary optimization and remigration. The boundary on the original image is highly discontinuous, and would be difficult to pick either manually or with most horizon tracking algorithms. In this case, it is much more difficult to tell whether the originally interpreted boundary is picked too deep or too shallow; therefore, it is advantageous to use the prior boundary as a guide rather than a negative example. The optimized boundary (Figure 6b) and resulting updated velocity model (Figure 7b) are smoother and feature fewer dramatic changes in



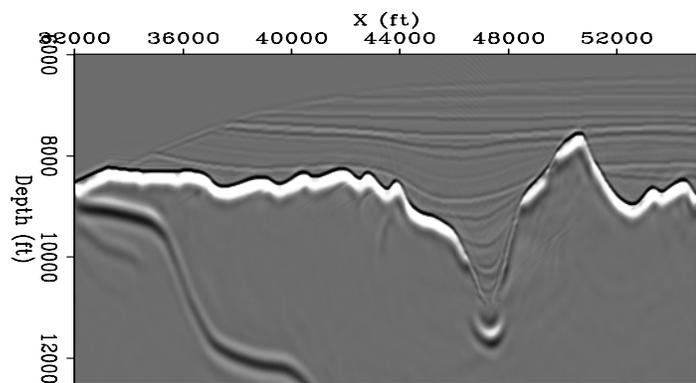
(a)



(b)



(c)



(d)

Figure 5: Sequence for using an updated velocity model to produce an improved image. (a) Salt canyon image after migration with the salt-flood velocity model. The original velocity model places the canyon bottom too shallow, pushing the image down. (b) The zero-contour boundary is still clearly too shallow, so the original boundary will be “penalized.” (c) Optimized, deeper boundary pick used to create an updated velocity model. (d) A much improved image of the canyon after remigration with the updated velocity. [CR] adam1/. update

the salt interface geometry. After migrating with the updated velocity model, the imaged boundary (Figure 6c) is smoother and more continuous than in the original image.

CONCLUSIONS

Lomask's work on salt body delineation via image segmentation is used in conjunction with plane wave migration in tilted Cartesian coordinates to build and update seismic velocity models. When applied to sediment- and salt-flood migrations of the Sigsbee synthetic dataset, this method provides relatively accurate picks of the top and bottom salt boundaries. These boundaries are used to construct a reasonably accurate seismic velocity model. The method is also used to update preexisting velocity models. Since Lomask's method is prone to error in cases where the "best" boundary pick can be produced using different eigenvector contour values in different parts of an image, the boundary choice is posed as an optimization problem. By allowing different contour values to be used throughout the same image, an optimized boundary is calculated. Velocity models derived from the optimized boundary pick produce improved migrated images, both for synthetic and real data.

ACKNOWLEDGMENTS

We would like to thank SMAART JV for the Sigsbee synthetic dataset, and WesternGeco for providing the Gulf of Mexico data.

REFERENCES

- Hale, D. and J. U. Emanuel, 2002, Atomic meshing of seismic images, *in* Expanded Abstracts, 2126–2129, SEG.
- , 2003, Seismic interpretation using global image segmentation, *in* Expanded Abstracts, 2410–2413, SEG.
- 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.
- Shan, G., 2008, Imaging of steep reflectors in anisotropic media by wavefield extrapolation: PhD thesis, Stanford University.
- Shan, G. and B. Biondi, 2007, Plane-wave migration in tilted coordinates: SEP-Report, **131**.
- 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.

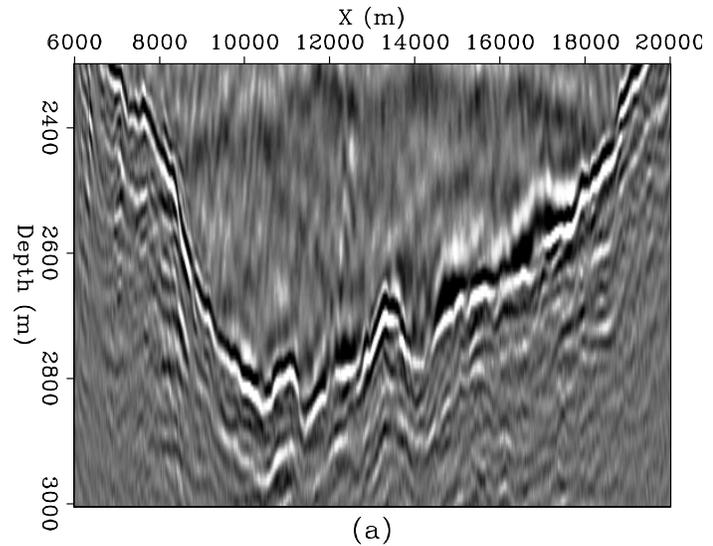
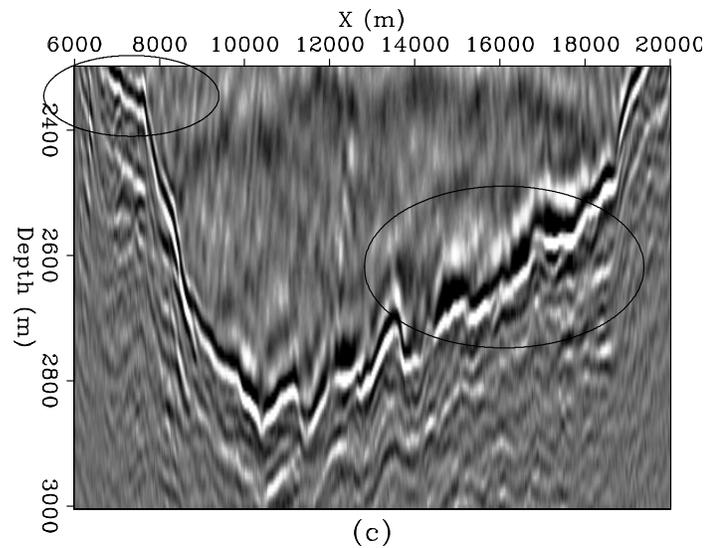
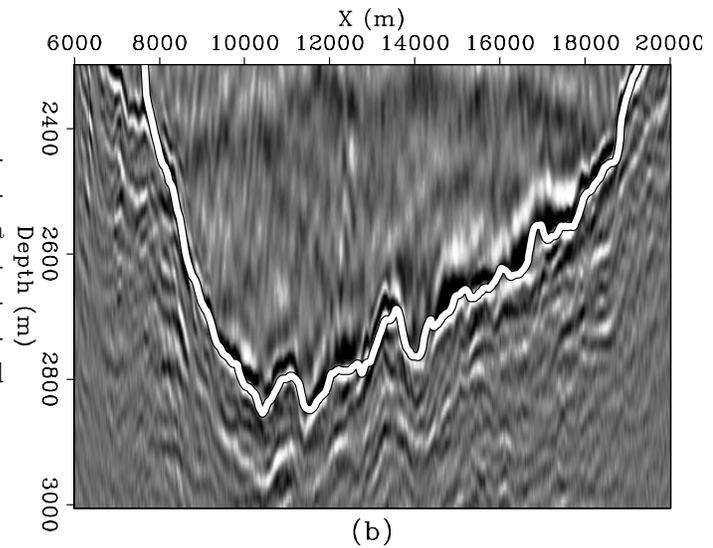


Figure 6: Original image (a), optimized boundary path (b), and updated image (c) for a portion of a Gulf of Mexico dataset. The updated image features a more continuous salt interface than in the original, especially within the indicated areas. [CR] `adam1/. gom-update`



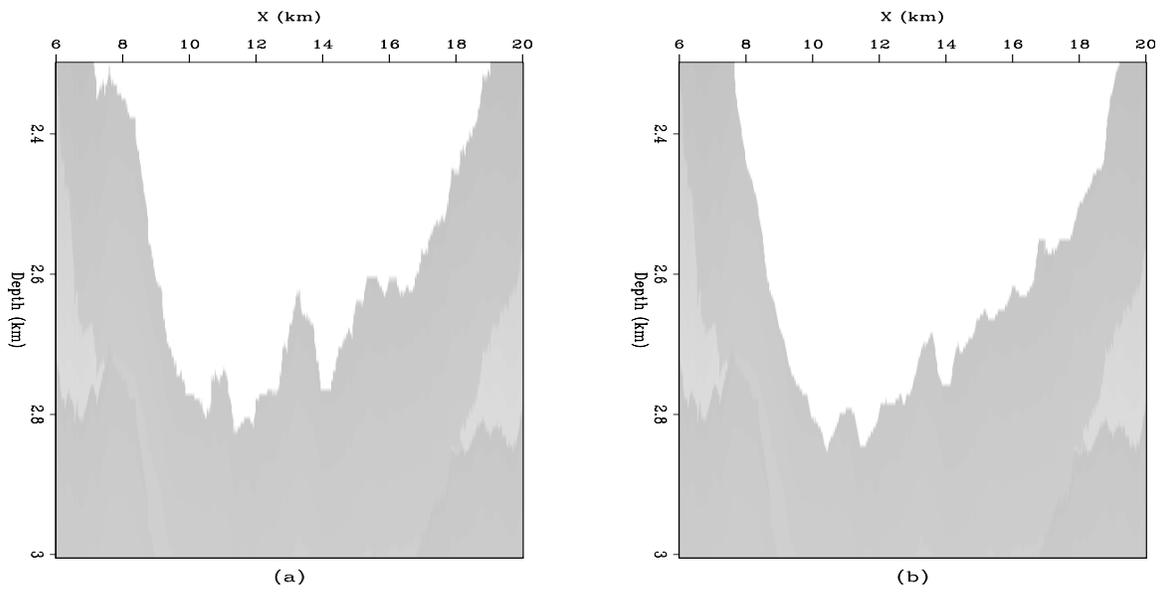


Figure 7: Original (a) and updated (b) velocity models for the section shown in Figure 6.

[ER] adam1/. gom-vel

Lloyd and Viterbi for QC and auto-picking

Robert G. Clapp

ABSTRACT

Automatic picking and the QCing of these picks are crucial step in the velocity analysis loop. In this paper I show that a modified version of Viterbi's algorithm can be an effective auto-picker when used interactively. In addition I show that Lloyd's algorithm can reduce densely auto-picked information to a representative subset that simplifies the QCing process.

INTRODUCTION

To speed up the velocity analysis loop it is important to allow the computer to do as much as work as possible, particularly in the human intensive picking portion of the loop. As essential is an easy way to evaluate and modify (QC) these automatic picks.

Dynamic programming is an effective tool for finding a solution for certain types of relatively small, non-linear problems such as semblance picking. In biology, dynamic programming is used for pairwise alignment of amino acid sequences (Needleman and Wunsch, 1970). In electrical engineering, it is used for error correction in wireless communication and speech recognition (Hosom et al., 1999) among many other things. We can also find examples of its use in geophysics. Kruse (1988) used dynamic programming for signal correlation and trace interpolation. Kruse (1988) calculates an error function based on the difference in instantaneous frequency between all points along two signals. Dynamic programming is then used to find the error path with the least energy. Liner and Clapp (2004) used dynamic programming for alignment. Zhang (1991) used it for a starting solution when doing event picking.

Quantization is an important field in both electrical engineering and computer graphics. In speech compression, it is important to accurately describe a signal in as few bytes as possible. In image processing, it is often important to reduce the number of colors in image with as little loss in image quality as possible. One family of method often employed in quantization is based on Lloyd's method (Lloyd, 1982), an iterative technique that allows for variable rate quantization. For QCing automatic picks, the ability to reduce densely picked functions to a smaller set of representative points simplifies substantially the QCing process.

In this paper I use a dynamic programming technique to automatically pick semblance gathers and reflectors. In addition, I show that quantization can be used to represent a function by a subset of representative points.

VITERBI

Dynamic programming, and specifically the Viterbi algorithm, offers a way to solve certain classes of non-linear problems. It is useful for problems that can be thought of as making one decision after another. To understand how it works, it is easiest to start with our final decision. Our goal is to maximize our ‘score’ S of a series of decision $1\dots n$. Each decision has several potential outcomes ($1\dots m$). Our final score is going to be based on some score we have calculated for all of our possible options (called ‘states’) at $j = n - 1$, and the best score we can get from moving from all of the possible states at $n - 1$ to all possible states at n . We can write the score as

$$S(i, j) = \max_{k=1\dots m}[S(k, j - 1) + v(i, j, k)], \quad (1)$$

where i is the given state and $v(i, j, k)$ is the value obtained from moving from state k to state i at decision j . The best series of decisions is then found by going backwards, taking the state at each decision i that corresponding to the highest score.

In this most general form, the algorithm is quite expensive. The cost on the order of $n * m * m$. For a large number of states the problem quickly becomes impractical. The easiest way to reduce the cost is to limit the number of states that we must search when moving from one decision to another.

For the interactive picking problem we are looking for the best path through a series of points in 2-D. In this example I am requiring that the solution is single valued along one axis (e.g. only one velocity at each time sample). I am looking for the best path between my first and last picks along the single value axis. I form an initial path by linear interpolation between the selected points. I am then going to limit my search space so it is no more than x points away from the linear path (limit the possible states of k). Figure 1 demonstrates this concept. The left panel shows a semblance scan along with three selected points. The right panel shows the semblance extracted along the path represented by the three points.

In term of equation (1) we have n decisions (the number of time samples between the first selected point and the last selected point) with m states ($2 * d + 1$ in size where d is the search distance from the initial linear path.) Generally v takes the form

$$v(i, j, k) = m(i, k) - p(j, k), \quad (2)$$

where m is the semblance at a given location and p is a scalar that punishes jumps in the selected semblance path. What is used for equation 2 can have dramatic effects on the solution. For the applications discussed in this paper I defined p so that a single sample jump was mildly punished with rapidly increasing penalty with larger gaps. This has the effect of tending to create smooth solutions.

To further create a smooth path I only search in the range $abs(j - k) < 5$. Finally to force the user selected points to be honored I modify m in a manner similar to Harlan (2001). I add large values to m which has the effect of forcing the solution through these points. Figure 6a shows the generated score matrix from the data shown in Figure 1b overlain by the traced back path through these points. The path is calculated by finding the maximum in the first row and then searching for the maximum within some range in the next row, proceeding to the bottom of the score matrix. Figure 6b shows the path on the original data

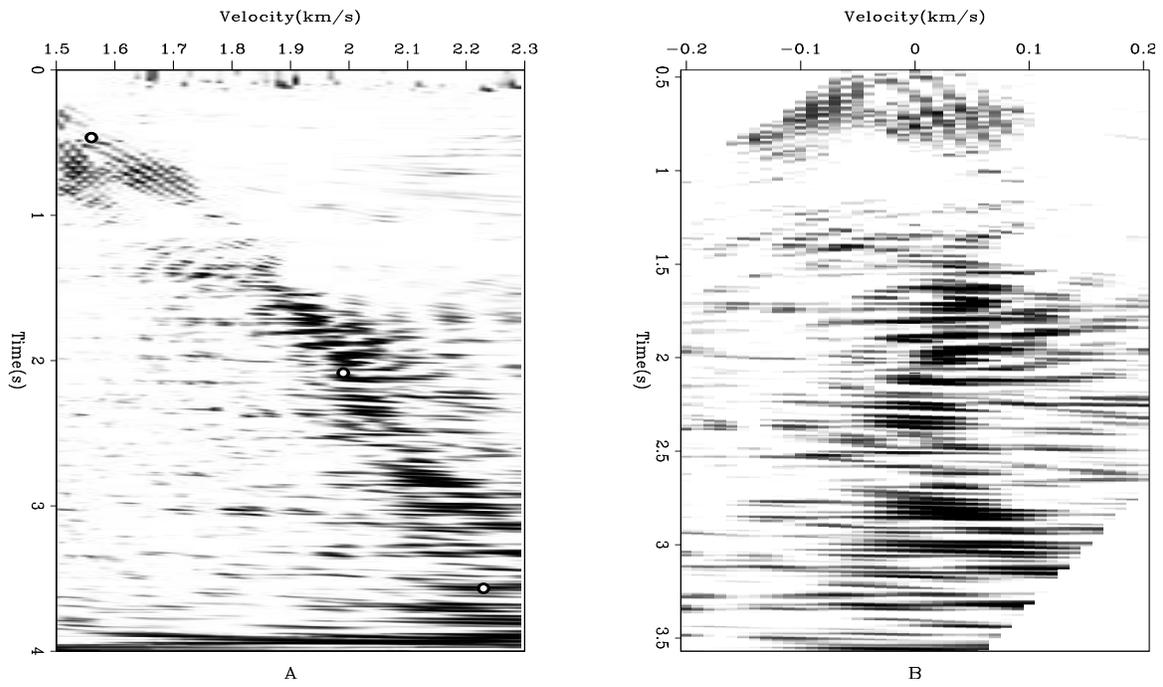


Figure 1: The left panel shows a semblance gather overlain by three picks. The right panel is the result of linearly interpolating between the picks and extracting semblance along the resulting line. [ER] bob1/. path

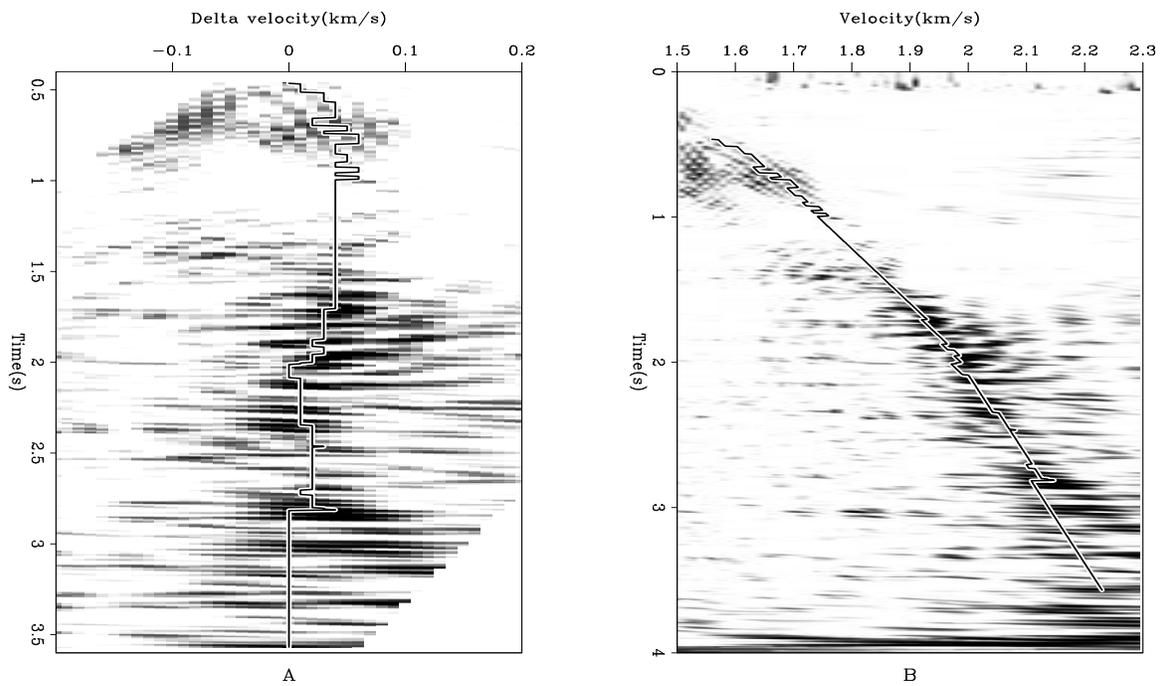


Figure 2: The left panel (A) shows the score matrix calculated using equation 1 overlain by the maximum tracked path. The right panel, B, shows the path overlain on the original semblance display. [ER] bob1/. semb

Reflectors can also be auto-picked by this method. In the case of picking reflectors, m becomes the correlation of the data along with an initial linear path. Figure 3 demonstrates the concept. The left panel of Figure 3 show the data with a set of four points selected. The right panel of Figure 3 shows the path picked by the algorithm.

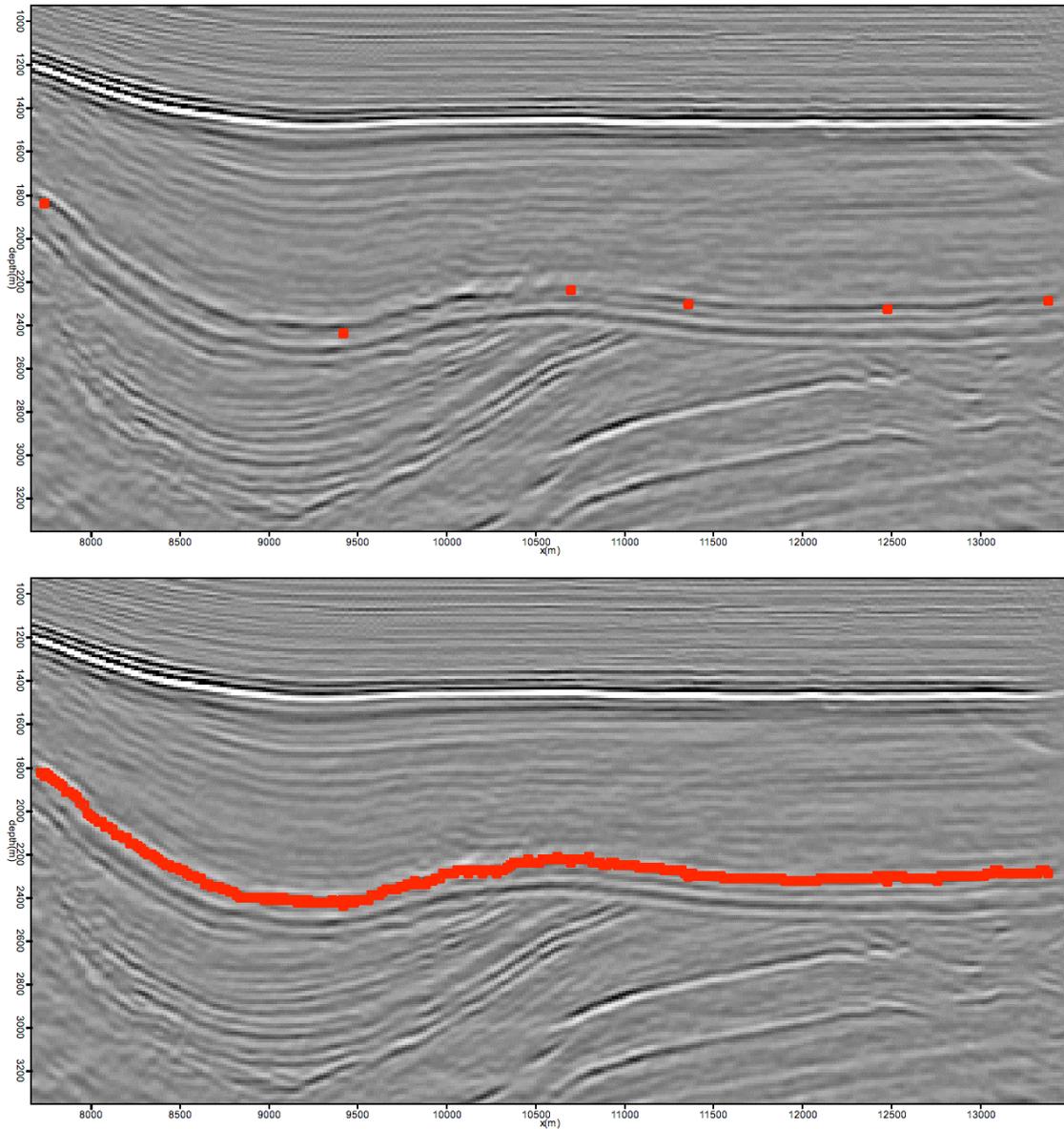


Figure 3: The left panel shows the original with four seed points selected. The right panel shows the result of using the Viterbi algorithm to pick the reflector. [NR] bob1/. refs

LLOYD

The concept of quantization originates in the field of electrical engineering. The basic idea behind quantization is to describe a continuous function, or one with a large number of samples, by a few representative values. Let x denote the input signal and $\hat{x} = Q(x)$ denote quantized values, where $Q(\cdot)$ is the quantizer mapping function. There will certainly be a distortion if we use \hat{x} to represent x . In the least-square sense, the distortion can be measured by

$$D = \sum_i^n (x - Q(x))^2. \quad (3)$$

Consider the situation with L quantizes $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_L)$. Let the corresponding quantization intervals be

$$T_i = (a_{i-1}, a_i), i = 1, 2, \dots, L, \quad (4)$$

where $a_0 = \min(x)$ and $a_L = \max(x)$. The distortion function then becomes

$$D = \sum_{i=1}^L \sum_{x=a_{i-1}}^{a_i} P(x)(x - \hat{x}_i)^2, \quad (5)$$

where $P(x)$ is the discrete version of the probability density function, or normalized histogram ($\sum_x P(x) = 1$). To minimize the distortion function D , we take derivatives of equation (5) with respect to \hat{x}_i , a_i and set them equal to zero, leading to the following conditions for the optimum quantizers \hat{x}_i and quantization interval boundaries \hat{a}_i :

$$\hat{a}_i = \frac{\hat{x}_i + \hat{x}_{i+1}}{2}, \quad (6)$$

$$\hat{x}_i = \frac{\sum_{x=\hat{a}_{i-1}}^{\hat{a}_i} P(x)x}{\sum_{x=\hat{a}_{i-1}}^{\hat{a}_i} P(x)}. \quad (7)$$

A way to solve this coupled set of nonlinear equations is to first generate an initial set $\{x_1, x_2, \dots, x_L\}$, then apply equations (6) and (7) alternately until convergence is obtained. This iteration is well known as the Lloyd-Max quantization algorithm (LMQ). A common modification is to form

$$D(i) = \sum_{x=a_{i-1}}^{a_i} P(x)(x - \hat{x}_i)^2, \quad (8)$$

and to remove a_i where the distortion is small and possibly add a s in regions where the distortion is large. The resulting a locations is often much smaller than the initial set of values.

The LMQ scheme is designed to find the best representation of a distribution, which is not what I am trying to do in this instance. Instead I am trying to achieve the representation of $y(x)$ with as few x_i, y_i points as possible. The twist on the standard LMQ scheme is the replacement of $P(x)$ in equation 5. Instead of being the probability density function I construct an error from a background piece-wise linear function. I first construct $z(x)$ by linear interpolating between x_i, y_i samples. I then calculate $d(x) = y(x) - z(x) + \min(y(x))$, the error from the piecewise linear background. Figure 4 demonstrates the methodology. Figure 4a shows a curve with ‘*’ the initial x_i, y_i points and the resulting

$z(x)$ function. Figure 4b shows the $d(x)$ function constructed from $y(x)$ and $z(x)$. We now have something that is approximating the shape of a probability density function except that it can be positive or negative. To get around this problem I first

$$sn = \sum_{i=x_{i-1}}^{x_i} d(i), \quad (9)$$

then if sn is positive I define $P(x)$,

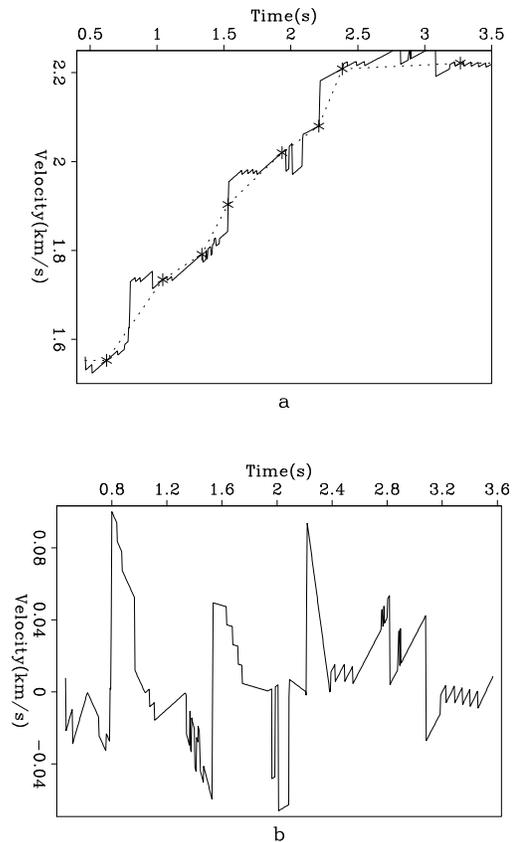
$$P(x_{i-1}..x_i) = d(x_{i-1}..x_i) + \min(d(x_{i-1}..x_i)). \quad (10)$$

If $sn < 0$ I define

$$P(x_{i-1}..x_i) = -d(x_{i-1}..x_i) - \max(d(x_{i-1}..x_i)). \quad (11)$$

Figure 4: Panel (a) shows the original curve (solid line); an initial set of a values, asterisks; and the background, dashed, curve z . Panel (b) shows the deviation d from the piece-wise linear background. [ER]

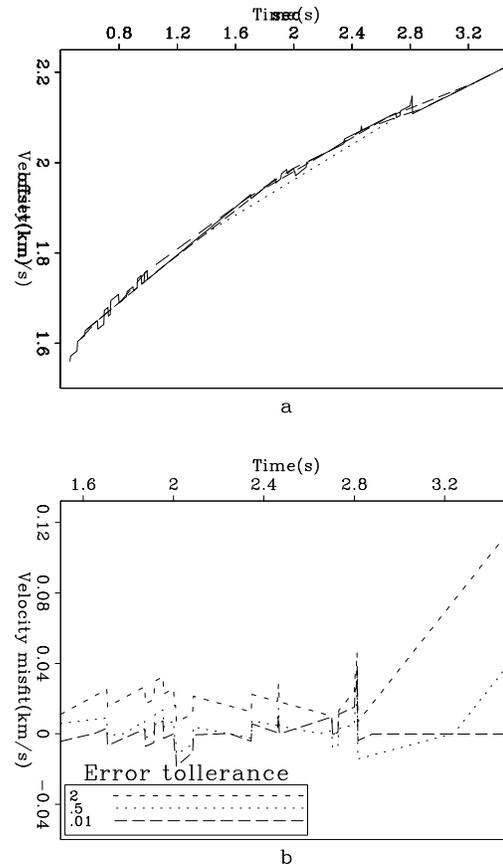
bob1/.lloyd1



As a result $P(x_{i-1}..x_i)$ is always positive. Flipping the signs does not violate the LMQ concept. What equation 7 is attempting to do is a *local* center of mass calculation. By applying equation 10 or 11 we are transforming our coordinate system to obtain an accurate center of mass calculation. How accurate the curve is represented is determined by the number of a_i terms. In practice it is best to start with a dense representation of a_i to avoid local minima and then use the fitting criteria of equation 8 to eliminate points in regions with small deviations. Figure 5 demonstrates this concept. The solid curve in Figure 5 is the original function. The three dashed curves show different deviation criteria. With increasing accuracy an increasing number of points are needed to represent the curve. In this example 2, 9, 28 and points are used.

Figure 5: The effect of modifying the deviation criteria. In panel (a) the solid curve in Figure 5 is the original function. The three dashed curves show three different deviation criteria. The closer the fit to the original curve the more points that are needed for an accurate representation. Panel (b) shows the error in the fitting functions. [ER]

bob1/. lloyd



Extensions

As demonstrated in Clapp (2006); Tang and Clapp (2006) Lloyd's algorithm is easily extended to multi-dimensions. A dense two or three dimensional volume of picks could be reduced to a few hundred or thousand points enabling a relatively easy QC process.

CONCLUSION

Dynamic programming can be effective in automatically picking surfaces and semblance. The method, in current form, is limited to 2-D, limits its applicability to the surface picking problem. Lloyd's algorithm offers a way to QC automatically picked volumes. By replacing the probability distribution function with an error misfit function, Lloyd's algorithm can be used to effectively characterize a dense series to a sparser set of picks while maintaining its essential character.

REFERENCES

- Clapp, R. G., 2006, A modified lloyd algorithm for characterizing vector fields: SEP-Report, **124**, 249–256.
- Harlan, W., 2001, Automatic moveout picking: WWW. (<http://billharlan.com/pub/paper/autopick.pdf>).

- Hosom, J.-P., R. Cole, and M. Fanty, 1999, Speech recognition using neural networks at the center for spoken language understanding: http://cslu.cse.ogi.edu/tutordemos/nnet_recog/recog.html.
- Kruse, C., 1988, A new method of nonlinear signal correlation using the instantaneous spectrum: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1321–1323.
- Liner, C. L. and R. G. Clapp, 2004, Nonlinear pairwise alignment of seismic traces: *Geophysics*, **69**, 1552–1559.
- Lloyd, S. P., 1982, Least squares quantization in pcm: *IEEE Transactions on Information Theory*, 127–135.
- Needleman, S. and C. Wunsch, 1970, A general method applicable to the search for similarities in the amino acid sequence of two proteins: *J. Mol. Biol.*, **48**, 443–453.
- Tang, Y. and R. G. Clapp, 2006, Lloyd's algorithm for selecting reference anisotropic parameters during wavefield extrapolation: *SEP-Report*, **124**, 257–270.
- Zhang, L., 1991, Automatic picking and its applications: *SEP-Report*, **70**, 275–292.

Hypercube viewer

Robert G. Clapp, David M. Chen and Simon Luo

ABSTRACT

Efficient viewing and interacting with multi-dimensional data volumes is an essential part of many scientific fields. This interaction ranges from simple visualization to steering computationally demanding tasks. The mixing of computation and interpretation requires a library that allows user inputs and generated results to easily be transferred. We wrote **Hyperview** in C++ using the QT library to facilitate this interaction. We describe the graphical user interface to the library and the basic design principles. We demonstrate the flexibility of the underlying libraries through a simple semblance picking application.

INTRODUCTION

Viewing and interacting with multi-dimensional volumes is necessary when working with 3-D data. SEP wrote its first movie program 28 years ago and has continually expanded on this initial idea (Claerbout, 1981; Sword, 1981; Ottolini, 1982, 1983, 1988, 1990). These movie programs have progressed from simply showing a series of frames to allowing greater and greater levels of interactivity.

Interactivity can take several forms. Numerous attempts have been made at SEP to take human input to geophysical algorithms. (Claerbout, 1987, 1991) built interactive filtering tools. Several authors van Trier (1988); Berlioux (1994); Clapp et al. (1994); Mora et al. (1995) have built interactive tools for velocity analysis. Cole and Nichols (1992, 1993) built a generic X11 based toolkit for interactivity.

More recent efforts have been focused on expanding Rick Ottolini's viewing program **Ricksep**. Clapp (2001) added the ability to view multiple datasets simultaneously and built hooks to allow processes like interactive NMO analysis. Chen and Clapp (2006) expanded on this work by adding the ability to overlay datasets and including the capability to display well logs and other data types. Two problems more further expanding **Ricksep** challenging. First, **Ricksep** was written in C rather than an object oriented language more appropriate to graphics. Second³⁴, it uses the Motif graphics library, whose future is uncertain.

In this paper, we present a new slice viewing program **Hyperview**. **Hyperview** is written in C++ using the QT * library. It preserves almost all of the capabilities of **Ricksep** while adding numerous additional interactive features. This paper is broadly broken into three parts. The first part gives an overview of **Hyperview**. This is followed by a description the various menu and keyboard options available. The last portion of the paper is devoted to the design of the viewer and ideas for how it can be extended.

*<http://www.trolltech.com/products/qt>

OVERVIEW

Hyperview can be more properly thought of as the front end to a library that specializes in viewing and interacting with hypercubes. **Hyperview** is a python script that reads in the files to display, guesses their type based on their suffixes and then calls the underlying program **Hyperview**. Appendix A describes the command line options to **Hyperview**. The python script calls `qt_cube`, a C++ program that takes as input one or more hypercubes of the same size. The library allows significant additional flexibility that is described in the design section of this paper. The remaining portion of this section will be limited to describing the default behavior, rather than the potential flexibility allowed by the library.

Hyperview displays one or more regularly sampled datasets that have between two and five dimensions. For **Hyperview**'s default behavior, each dataset have the same number of dimensions and number of elements along each axis. The datasets are read into memory and stored as 8-bit unsigned char. Two different windows, status and display, are brought up when **Hyperview** is invoked.

Windows

The display window (Figure 1) contains one or more views of the dataset(s) that have been loaded into memory. The number of views is controlled by the `nviews` command line argument. Each `view` can choose to see any of the loaded datasets. Using the mouse, the user can navigate through up to three dimension of the hypercube. When working with four or five dimensional datasets the user has the ability to select which three dimensions to work with. All of the views are linked together, by moving to a new location in one view you will be taken to the same location in all other views.

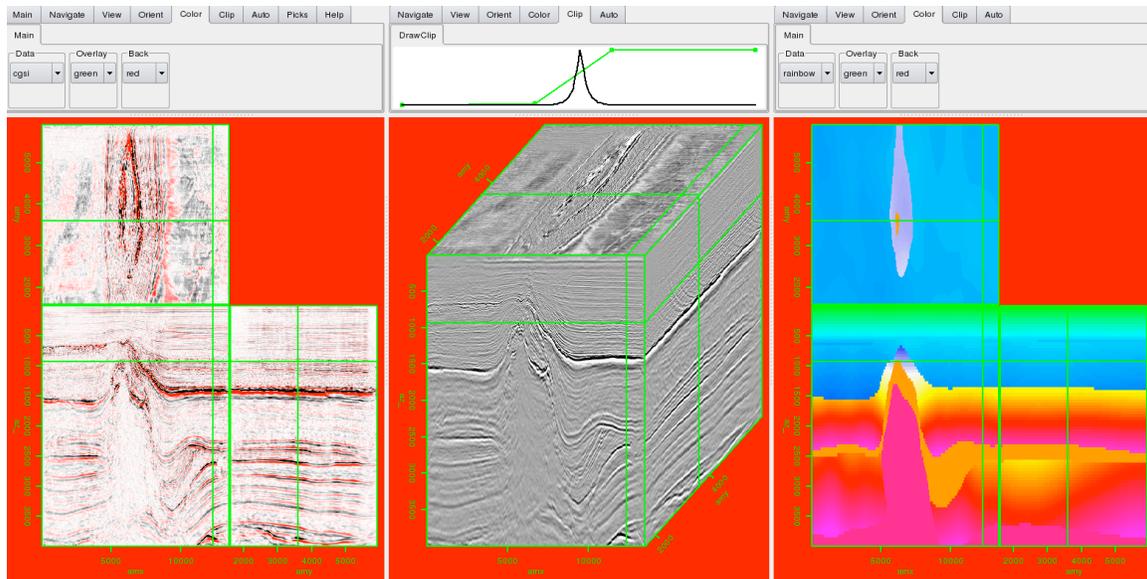
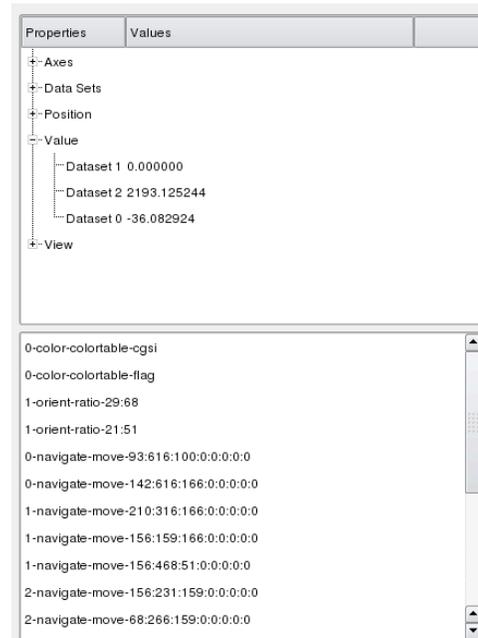


Figure 1: The display window with three different views. In this case three different datasets are being displayed. [NR] bob2/. display

The status window (Figure 2) is broken into two parts: information and history. The top portion of the window information about the datasets currently shown. You can view: axes information (origin, sampling, number of samples, and axis label), the names of the datasets that were read in, which dataset you are currently viewing in each view, the position in the hypercube, and the data value in each dataset at the current position. The history window records every mouse and menu action taken and is more fully described later.

Figure 2: The control window. The top of half provides information about the dataset(s). The bottom shows the last ten actions performed by the user. [NR] bob2/. control



Data

Hyperview has significant flexibility in the type of data it can read. Its default behavior requires that all data volumes be the same size (e.g. same number of dimensions and same number of samples per axis). The object oriented nature does not make this a requirement; the velocity viewing example presented later demonstrates this flexibility.

Currently **Hyperview** supports five of the more common seismic data formats but can easily be expanded to read almost any other format. It reads SEPlib (and RSF) regular cubes in both byte and float format. It can read float formats of SEG-Y, SU, SeisSpace and the scaled integer format of SeisSpace. By default when reading float data **Hyperview** reads the first 5 MBs to find clip parameters and stores the entire cube in byte format based on the clip information. Adding `float_format=1` to the command line it will store the data as floats rather than bytes. This feature is useful for both clipping and for actions that require more precision than bytes.

When reading in float data, the program looks for a series of command clipping options. It first looks for `bpclip` and `epclip`, corresponding to a beginning and ending clip percentile. It next looks for `pclip` which corresponds to a percentile clip based on the absolute value of the data. It then looks for minimum and maximum clip values `bclip` and `eclip`. Finally, it looks for `clip` which corresponds to $bclip = -clip$ and $eclip = clip$. If none of these parameters are found, it defaults to $bpclip = .5$ and $epclip = 99.5$.

MODES

A standard mouse has three buttons and three potential actions (click, double-click, click-move-release). As a result, only 9 unique actions are possible. *Ricksep* gets around this limitation by using keyboard modifiers as the behavior of the mouse changes based on a key being simultaneously pressed on the keyboard. This approach is effective but somewhat cumbersome. In *Hyperview* we take an alternate approach which we call ‘modes.’ Currently two modes are defined, navigation (the default) and picking. To switch to the picking mode the user hits **Ctrl-p** on the keyboard and to switch back **Ctrl-n**.

The navigation mode duplicates much of *Ricksep*’s functionality. Pressing the left mouse button and moving the cursor selects a region to zoom in on. Single-clicking the middle mouse button allows you to navigate through the cube. Double-clicking the left button unzooms. When viewing three faces of the cube (view modes: cube, cut, and three-face) you can change the relative sizes of the three faces by selecting with the left mouse button the shared corner point, moving it, and releasing the mouse button. Zooming in one axis in one display will cause they same zoom to occur in every other view where the axis is displayed.

When in picking mode the left mouse button actions are still enabled, but the center and right buttons’ functions change. Selecting the right button you can add a pick to the cube, and the center button delete the closest point.

MENUS

The menu is arranged in two levels of tabbed sub-menus. There are nine primary menus: **Main**, **Navigate**, **View**, **Orient**, **Color**, **Clip**, **Auto**, **Picks**, and **Help**. When using multiple views, the first view will contain all nine menus while subsequent views will have six of the menus, as **Main**, **Picks**, and **Help** contain global rather than view-specific options. In the following section I will describe the various options in these menus.

Main

There are three functions in the main view. The first is only relevant when running with multiple view windows. The user can choose to ‘lock’, or synchronize, the various views. When views are locked you guarantee that each view will have the same color table, perspective, orientation, etc. The main view also allows you to quit the application and to save the history of all the actions performed during the session.

Navigate

The navigation menu has two sub-menus, **Movie** and **Direction**. The **Movie** sub-menu relates to displaying a series of slices along one axis of the hypercube. The **Movie** sub-menu allows you to start and stop the movie (**Go** and **No**) and advance the move one frame. You also can control the speed of the movie with a slider.

The **Direction** sub-menu allows you to control both which axis to loop over and in which direction. You are limited to selecting one of the three axes currently displayed in

the view. To describe the effect of the various **Direction** options imagine a 3-D cube with the axes depth, X position, and Y Position. You can change the depth slice (+Z and -Z), different Y-Z slices (+X and -X) or X-Z slices (+Y and -Y) being displayed. Changing the movie direction you also automatically advances one frame.

Views

The **View** menu contains two to four sub-menus depending on whether you are working with a single dataset or multiple datasets. You always are presented with a **Main** and **Save** sub-menus and will have an additional **Data** and **Overlay** sub-menu when multiple datasets are loaded.

The **Main** sub-menu allows you to set the perspective, the font size, and whether or not to draw a colorbar. You have six perspective options, three that view a single slice of the data, and three that view multiple slices of the data. The three single slice options are **FRONT**, **SIDE**, and **TOP**; these correspond to viewing slices containing the 1-2, 1-3 and 2-3 axes, respectively. In addition, you can see a plan view **THREE**, a cube-view **CUBE**, and cut **CUT** into the data. The font menu allows you to control the font used to draw the axes, and the colorbar allows you to view a colorbar with a superimposed histogram.

The **Save** sub-menu allows you to save the current view as a ppm file or a large or small postscript file. The small postscript option is appropriate for inclusion in presentations while the large option's 1200 DPI is appropriate for papers. The large option is quite time consuming because it redraws the current view at approximately 64 times the resolution of a standard screen.

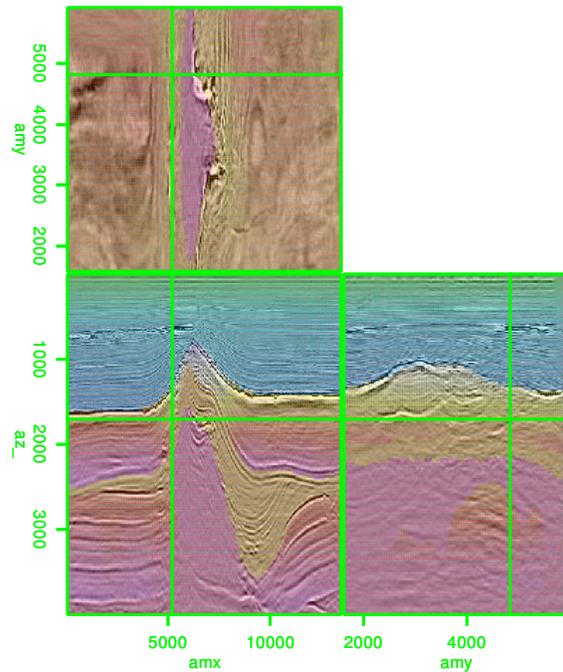
The **Data** and **Overlay** sub-menus are only available when using multiple datasets. The **Data** sub-menu allows you to choose which dataset, or combination of datasets you wish to see in the current view. You have the option of selecting any of the current datasets or overlaying two datasets. The overlay is done by modifying the opacity channel of the overlying dataset. Figure 3 shows an example of overlaying a velocity model and a migrated image. The figure itself is generated through **Save** menu. You can control the level of opacity through the **Opacity** sub-menu. In addition, you can change the colormap of the overlying dataset in this menu. With the **Cycle** button in the **View** sub-menu you can cycle through all of the datasets.

Orient

Generally, every dataset that **Hyperview** displays must be of the same size. Each view has its own axis order and direction. By default each views' first axis corresponds to the first axis in the dataset, the second the second, etc. The **Orient** menu allows you to manipulate the axis order and direction through its two sub-menus, **Main** and **Transpose**.

The **Main** sub-menu allows you to flip (reverse) any of the three currently viewed axes. This amounts to changing the direction in the view's map. With the **Main** sub-menu you can move to the center of the currently viewed cube or to an edge of the cube. In addition, you can reset the cube to its default orientation. This option will undo any flips, transposes, and/or movements you have done. The final option in the **Main** sub-menu is the ability to

Figure 3: An example of overlaying a velocity model on top of a migrated image. [NR] bob2/. overlay



view the cube in true proportions. This option will make sure that the length of each axis (the number of samples times the sampling) is proportional to the amount of space it takes on the computer screen.

The **Transpose** sub-menu allows you manipulate the view's axis order. For example, transposing the first and second axis, flips the front and top panel's positions. Most of **Hyperview's** actions work on only the view's first three axes. With the transpose options you can view and manipulate the fourth or fifth axis of a dataset.

Color

The **Color** menu controls color map of a dataset. Currently there are five colormaps: gray scale, flag, CBI, CGSI, and rainbow. Flag maps positive value to red and negative value to blue, zero is white. CGSI behaves similarly except negative values are mapped to black. CBI maps positive to blue, zeros to black, and negatives to red. The rainbow maps negative values to green transitioning to blue, white for zero, yellow, red, and finally purple for large positive values. The **Color** menu also allows you to change the color used for the text and the background color.

Picking

The picking capabilities of **Hyperview** are substantially improved over **Ricksep's**. There are nine different pick colors available. Each color represents a different pickset. Picks are stored as integers corresponding to the sample selected. When the slice you are viewing corresponds to a pick's coordinate, you see the pick in its primary color. You can also see picks when you are close to their position with the color slightly changed.

There are four sub-menus under picking: **File**, **Size/Distance**, **Color**, and **Draw**. The **File** menu allows you load and save all picks. The picks are saved in ASCII format with the pick location and color stored in the file. The **Size/Distance** sub-menu allows you to change the range in which you can see a pick and the size of the square corresponding to a picked location.

The **Color** sub-menu allows you to change the active pickset. It also allows you to view a specific pickset or all the picksets. The **Draw** sub-menu gives you the option to display the picks as either lines or points. In the simple logic of the program, in order to draw a line it must have an axis that is single valued. For example, in velocity analysis you will not select more than one velocity per time sample. In the **Draw** sub-menu you select the multi-valued axis. This indicates that any slice that contains the multi-valued axis also contains a single valued axis. A line can then be drawn connecting points along the single valued axis. By default, the multi-valued axis is set to the first axis (depth). Figure 4 shows an example of picking. Note the three different colors used and how lines appear in the front and side views while points are displayed in the top section.

Clipping

Using the clip menu you can change the dynamic range of your data. The clip menu displays two lines. The green line is a histogram of your dataset. Specifically the data is mapped into 256 different regions. In byte mode this corresponds to the 256 different values that a byte can take on. In float mode the regions range from the minimum to maximum clip value.

The second black line shows the mapping from these 256 different regions to the 256 different colors in the colormap. By default the first region maps to the first color map value, the second to the second, etc. The horizontal axis is the regions and the vertical axis is the color map index. The black line is constructed by linearly interpolating between control points, (note the squares at the far ends of the display). The right mouse button creates additional control points can be added (the center button deletes the points). If you wish to increase the dynamic range, you can introduce additional control points. Figure 5 shows an example. The left panel is the original data; the right panel is after adding control points that map outlying data values to the same color map index. Note how the right panel of Figure 5 shows much more dynamic range.

Auto picking

Currently there is a single auto-picking mechanism in **Hyperview**. The user selects a series of control points. These points are honored and then the algorithm attempts to fill in between the selected points. The user then has the ability to add in additional control points and re-pick, or delete the auto-pick points entirely.

The method is a modified version of the Viterbi algorithm described in Clapp (2008). This is a purely 2-D picking scheme. **Hyperview** loops through all of the currently displayed views. For each view it finds all of the points in the current active pick color. These picks are used as control points, and the Viterbi algorithm attempts to find a smooth path between

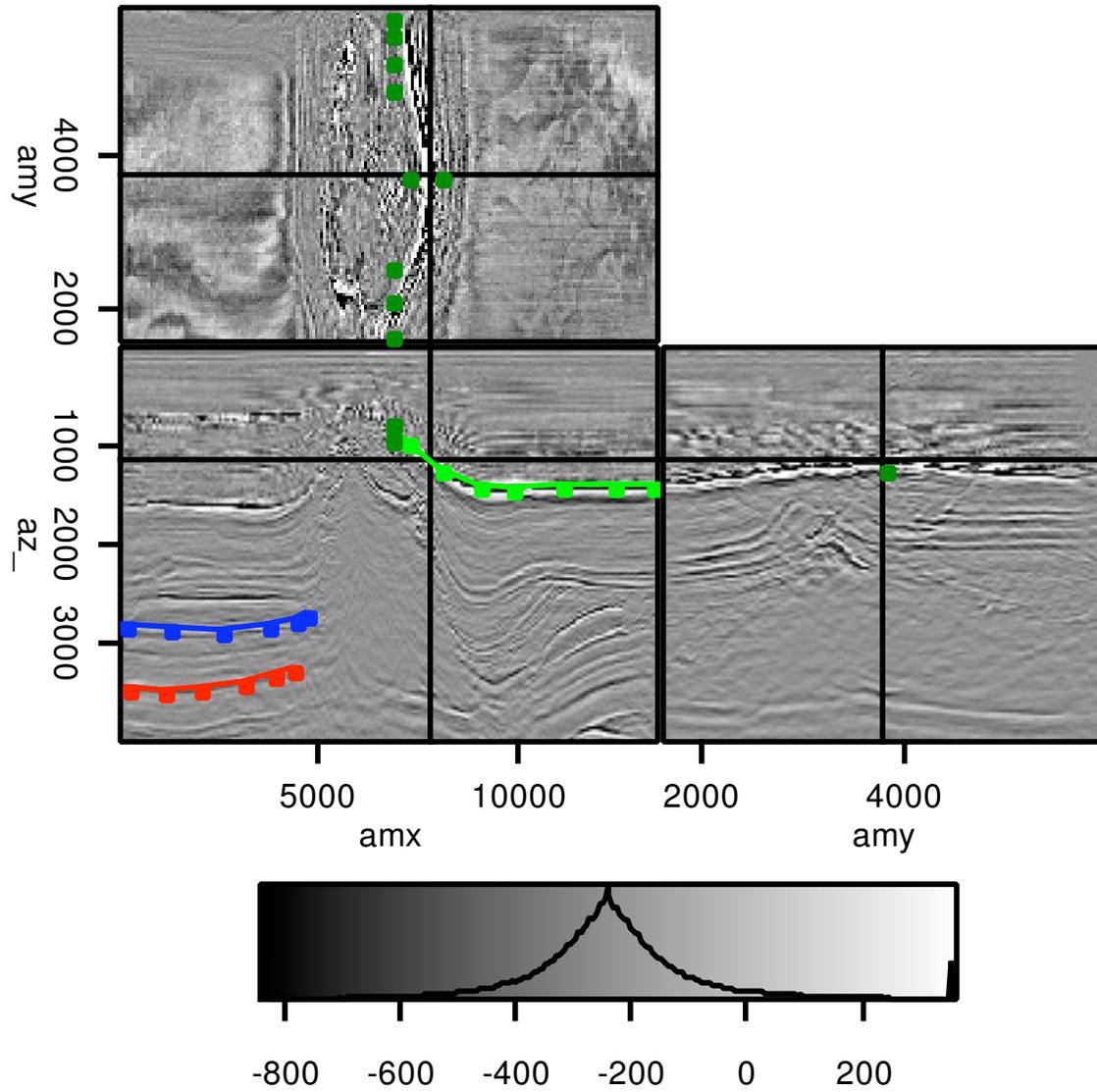


Figure 4: The result of picking three reflectors. Note the use of a colorbar and how picks show up as line in the front and side panel but as points in the top panel. [NR] bob2/. picks

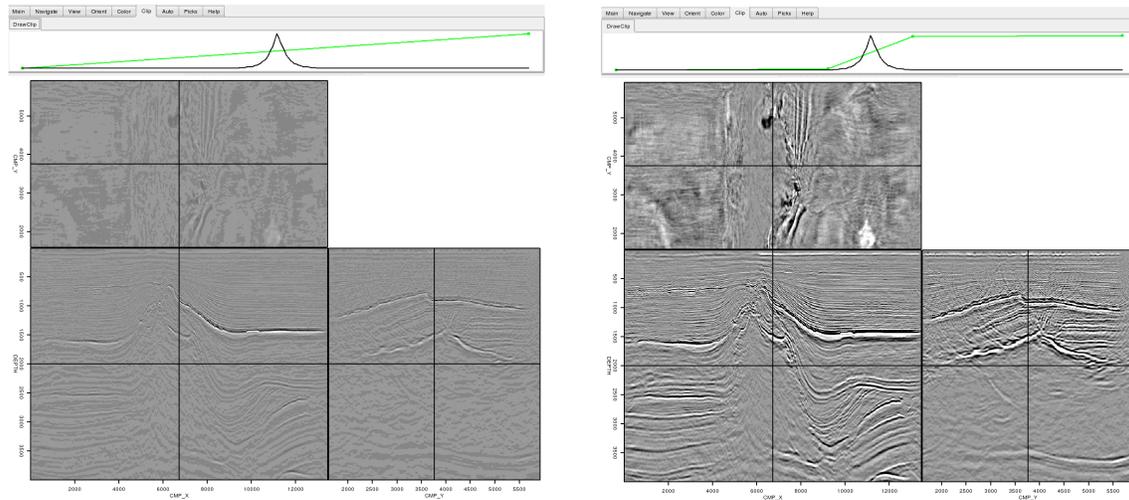


Figure 5: The effect of adjusting the clip. The left panel shows the original data and clip. In the right panel the clip has been adjusted by adding control points closer the range of most of the data's value. Note how the image on the right shows significantly more dynamic range. [NR] `bob2/. clip`

these points. By looking for large amplitudes between the picks (semblance) or finding the path with peaks in local correlation (reflectors).

An additional auto-picking sub-menu is `Lloyd`. The Lloyd's algorithm approach, described in Clapp (2008), is an attempt to make QCing easier. It takes a set of picked points and tries to find a reduced set of points which contains virtually the same information. The Lloyd's approach is currently a 2-D operation. It attempts to throw away points which are nearly linear combinations of neighboring points.

KEYBOARD SHORTCUTS

In addition to switching between modes described earlier, there are additional keyboard shortcuts.

Key	Action
<code>c</code>	Cycle forward one colormap.
<code>></code> , <code><</code>	Cycle forward or back one dataset.
<code>v</code>	Cycle through different perspectives.
<code>x</code> , <code>X</code>	Cycle forward or back along the x (second axis).
<code>z</code> , <code>Z</code>	Cycle forward or back along the z (first axis).
<code>y</code> , <code>Y</code>	Cycle forward or back along the y (third axis).
<code>r</code>	Start a movie running.
<code>s</code>	Stop a movie.
<code>q</code>	Quit current Hyperview session.

HISTORY

Hyperview borrows from **Gocad** (among other software packages) the ability to save and script actions. Every action through the menus, keyboard, or by the mouse is recorded by the program. Actually, what is recorded is both the action and how to counter the action. You can see the recorded actions in the history window. By clicking on an action you can undo that action all subsequent actions. In addition you can save all the actions you do during a run. This functionality is useful in two ways. First, it make it possible to consistently find a view that best illustrates a given point. Second, you can edit and rerun your file for a presentation. You can also add `sleep X` commands that will pause a presentation for a given number of seconds. The history functionality enables figure reproducibility. By saving the history rerunning it in batch mode the figure can be recreated from the command line.

DESIGN

Hyperview was designed to be easily modified for a variety of purposes. In this section we will cover the core objects, point out important public functions, and discuss some possible extensions that would add additional capabilities.

Window

Each view is a different `DrawWindow` object. It is a container class for all of the different objects associated with building, displaying, and interacting with a given dataset. It stores the current dataset being viewed, the colormap used to view the data, and how to display the data. Many of the menu and keyboard options modify the objects stored in this class.

Colormap

The `image_factor` object and its children describe how to take a 2-D field of unsigned chars and create an image of a given size. Currently this is limited to making a raster plot of the data based on a given colormap. One potential extension would be to add the ability to draw wiggle plots, graphs, contour plots.

Slice

The `slice` object's job is to display a 2-D plane. It requests from a data object a 2-D plane, potentially draws its axes, and the current location, and then calls the `image_factory` object to display the slice's contents.

View

The `view` object's children contain information about how a given `DrawWindow` object is going to display its dataset. Currently the view object has two children: `single`, which

displays a single 2-D plane from a dataset, and **multiple** which display multiple views of a given dataset. There are several potential additions to the view object. For example **Ricksep** allows an array of planes from a given dataset and a fences view of a given dataset.

Picks

The **pick_groups** object contains a series of **pickset** objects. The **picksets** are associated with a given color. Each **pickset** contains a series of **pick**'s. The **pick**'s contain its location and a integer type flag. Currently this flag is used to differentiate between picks made by the user and autopicked locations. The **pick_groups** object also contains how to draw the picks on a given slice.

Updater

One of the most important objects is the **update** object. This object is called by all menu, mouse, and keyboard actions. It stores a given action in its history along with how to undo the action. It then calls the appropriate object functions to accomplish its tasks. Any additional functionality must include an addition to the **update** object.

Data

The most important object is the **h_data** object. The data object influences the functionality of many other portions of **Hyperview**. The **h_data** object how to read the data its linked to in memory. Normally, this takes the form of reading one of the five data types described above. The **h_data** object knows to return a 2-D slice at a given location. Conventionally this involves extracting a portion of the buffer that was read in, but it could be expanded on to read from disk instead.

The data object also allows you turn off some default behavior if it is inappropriate for a given dataset. For example, each **h_data** says whether or not you can navigate within a view using the object, whether you can pick on the object, and what action to perform given a pick on the dataset. The contents of the dataset can also change. The **update_h_data** function is called before viewing in **DrawWindow** object. The **h_data** object has the option (based on the current position, picks that have been made, etc.) to change the data associated with a given view. Finally, the data object has the concept of fake axes. Generally **Hyperview** is limited to five dimensions. In fact an eight dimensional position is stored. These other three axes can be used to display datasets that conform to a subset of the first five dimensions. In the next section I will demonstrate how to use a number of these features.

SEMBLANCE PICKING

One of the most basic interactive processing operations is picking move-out. Writing a move-out application involved creating two new classes inherited from the **h_data** object, **nmo_semblance** and **nmoed**.

The `nmo_semblance` object is initialized with the data object associated with the pre-stack gathers. The `nmo_semblance` object takes advantage of the fake axis concept mentioned above. Instead of having an offset axis, a sixth axis, velocity, was added to the dataset. The data is generated from the CMP gather so the `read_h_data` does not perform any action. Finally, the `get_h_data` is modified. When initialized, every time the position changes the `nmo_semblance` object grabs the current CMP gather. It performs semblance on the gather and returns the resulting field. The center panel of Figure 6 shows the semblance calculated at the given CMP location.

The `nmoed` object is initialized with the both the pre-stack data and the `nmo_semblance` objects. Like the `nmo_semblance` object, it does not use the `read_h_data` function. The `get_h_data` function is also modified. The current CMP gather is grabbed from the data object, and the RMS picks closest to the current CMP are taken from the `nmo_semblance` object. The picks are then used to form a RMS velocity function and the CMP gather is NMOed with this function. The result is shown in the right panel of Figure 6. Picking is only allowed on the `nmo_semblance` object, and navigation is only allowed on the data object.

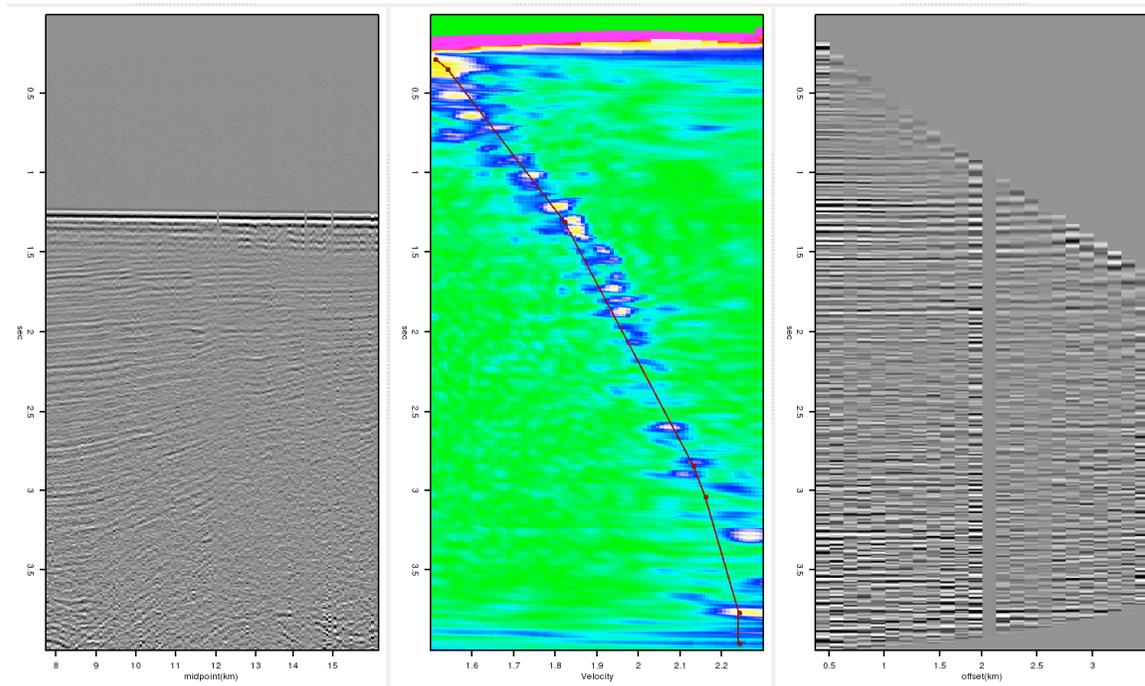


Figure 6: A typical semblance analysis display. The left panel show a depth-midpoint section. The center panel show the semblance at the given CMP overlain by the user selected v_{rms} function. The right panel is the result of NMO using the selected RMS function. [NR] `bob2/. semb`

FUTURE DIRECTIONS

Throughout the paper we discussed potential extensions to the current viewer. The most important feature to improve is the ease with which a programmer can retrieve information

from the GUI to send back an improved image. The current mechanism is effective but requires too much knowledge of the underlying libraries.

CONCLUSION

Interacting with multi-dimensional hypercubes is an essential tool in seismic exploration. The program, *Hyperview*, is a powerful viewing and interacting platform.

REFERENCES

- Berlioux, A., 1994, Building models with GOCAD: SEP-Report, **80**, 617–634.
- Chen, D. M. and R. G. Clapp, 2006, Data-fusion of volumes, visualization of paths, and revision of viewing sequences in ricksep: SEP-Report, **125**.
- Claerbout, J. F., 1981, On-line movies: SEP-Report, **28**, 1–4.
- , 1987, Interactive filter design in the z -plane: SEP-Report, **56**, 263–272.
- , 1991, Interactive one dimensional seismology program ed1D: SEP-Report, **71**, 293–294.
- Clapp, R., 2001, Ricksep: Interactive display of multi-dimensional data: SEP-Report, **110**, 163–172.
- Clapp, R. G., 2008, Lloyd and viterbi for qc and auto-picking: SEP-Report, **134**, 2000–2001.
- Clapp, R. G., B. Biondi, and M. Karrenbach, 1994, AVS as a 3-D seismic data visualizing platform: SEP-Report, **82**, 97–106.
- Cole, S. and D. Nichols, 1992, Xtpanel: An interactive panel builder: SEP-Report, **75**, 497–520.
- , 1993, Xtpanel update: Interactivity from within existing batch programs: SEP-Report, **77**, 409–416.
- Mora, C. B., R. G. Clapp, and B. Biondi, 1995, Velocity model building in AVS: SEP-Report, **89**, 133–144.
- Ottolini, R., 1982, Interactive movie machine user’s documentation: SEP-Report, **32**, 183–196.
- , 1983, Movie cubes: SEP-Report, **35**, 235–240.
- , 1988, Movies on the Macintosh II: SEP-Report, **59**, 255–268.
- , 1990, Seismic movies on the XView graphics system: SEP-Report, **65**, 315.
- Sword, C. H., 1981, SEP goes to the movies: SEP-Report, **28**, 11–20.
- van Trier, J., 1988, An interactive interface for velocity optimization using geological constraints: SEP-Report, **59**, 241–254.

APPENDIX A

The script `Hyperview` has numerous command line arguments. The general form is

```
pname file1 [file file3 pars]
```

where `file1...3` are data files and `pars` are series command line arguments of the form `param=value`. Generally, the options break down into three categories, commands that are data specific, view specific, and everything else.

Data specific options - By default `Hyperview` checks the suffix of all input files. It then tries to guess the data format. It makes the following suffix assumptions:

Suffixes	Data type
H,h,T,t,HH	SEPLib
rsf, RSF	RSF
su, SU	SU
segy,SEGYSgy, SGY	SEGYS
sp,SP	SeisPak

If the suffix of the file does not match one of these defaults or the default type is incorrect the user can add `typeX=FORMAT` where X is the order of the dataset (starting with 0) on the command line and `FORMAT` is either `SEGYS`, `SEP` (`RSF` is equivalent to `SEP`), `SU`, or `SEISPAK`.

View specific options - View specific options take the form `paramX=Y` where X is the view number.

Option	Default	Description
viewX	THREE	Style of view, must be <code>FRONT</code> , <code>SIDE</code> , <code>TOP</code> , <code>CUT</code> , <code>CUBE</code> , or <code>THREE</code> .
orderX	1,2,3,4,5,6,7,8	Axis order for the given view.
backgroundX	red	Background color, must be <code>red</code> , <code>green</code> , <code>grey</code> , <code>black</code> , or <code>white</code> .
overlayX	green	Overlay(text) color, must be <code>red</code> , <code>green</code> , <code>black</code> , or <code>white</code> .
colortableX	gray	Colortable to use view the given view, must be <code>flag</code> , <code>cbl</code> , <code>cgsi</code> , or <code>rainbow</code>

Other options - Two additional command line arguments are available: `nviews` and `position`. The `nviews` option tells how many different views to bring up in the display window; by default a single view is created. The `position` argument is the initial location to view in the cube. By default the center cube is displayed.

Interpolation of near offsets using multiples and prediction-error filters

William Curry and Guojian Shan

ABSTRACT

Most conventional marine reflection seismic data lack sampling of near offsets. We address this problem by interpolating data with a nonstationary prediction-error filter (PEF) that is first estimated from fully-sampled training data and then is used to interpolate missing data to produce an interpolated output. These training data need not be perfect, and may differ in amplitude and phase but should contain the local multi-dimensional amplitude spectra of the data we wish to recreate. We generate pseudo-primary data by crosscorrelating multiples and primaries in the recorded data. These pseudo-primary data can be generated at missing near offsets, but contain many artifacts, so it is undesirable simply to replace the missing data with the pseudo-primaries. Fortunately, many of the problems with the pseudo-primaries do not influence PEF estimation, so a desirable PEF can be obtained from these data, and then used to interpolate the missing near inline offsets to produce a result that is superior to direct substitution of the pseudo-primaries into the missing offsets.

INTRODUCTION

In most data sets, after the over-sampled time axis, the inline receiver axis is the most densely-sampled axis; the receivers are attached to a single cable so the sampling is dense and regular along this cable. Since both the air-gun source and receiver cable are usually towed by the same boat, the maximum inline offset, the largest inline distance between the source and a receiver, is limited by the length of the cable. The near end of the receiver cable is not at the source for obvious reasons; instead the receiver cable is towed a fixed distance behind the source. This distance, the inline near-offset, is consistent throughout the survey and is typically several times the inline receiver sampling interval. An example of a single shot from a 2D marine survey showing this near-offset gap is in Figure 1.

Near-offset traces are particularly valuable. Many methods attempt to recreate zero-offset data from larger offsets; standard multiple-removal techniques, such as surface-related multiple elimination (SRME) (Verschuur et al., 1992), require zero-offset data. Moveout differences between primaries and free-surface multiples are slight at these near-offsets thus compromising the performance of radon-based multiple removal algorithms that discriminate based on differential moveout.

There are many methods that could be used to reconstruct this near-offset gap. A simple way of recreating the missing near offsets is to replace the missing offsets with an NMO-corrected trace from the nearest offset. More sophisticated radon-based methods (Sacchi and Ulrych, 1995; Trad et al., 2002) are commonly used, as are Fourier-based methods (Liu and Sacchi, 2001; Xu et al., 2005). These methods all use existing data recorded adjacent

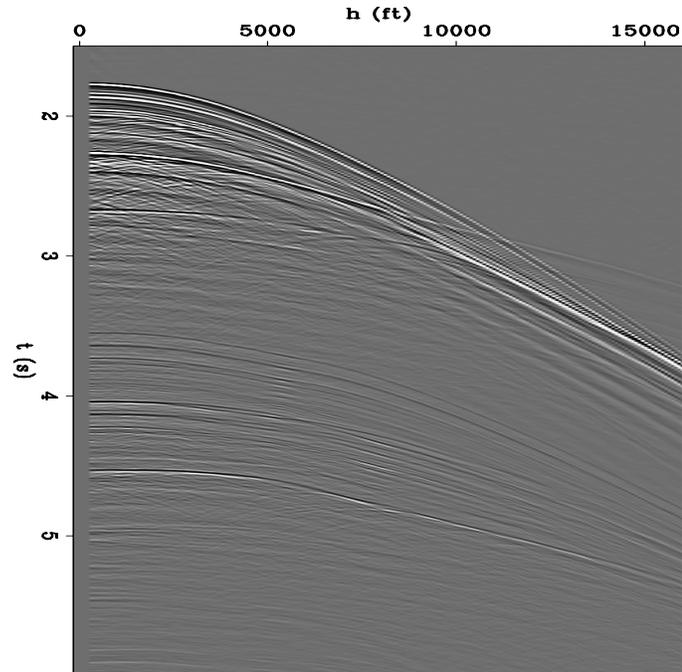


Figure 1: A single shot profile from a 2D Gulf of Mexico seismic survey. The nearest offset is at 330 ft for a near-offset gap of four traces. [ER]

bill2/. realshot

to the missing near offsets to create the missing data, so the performance degrades as the gap increases in size, common in undershooting situations.

Another approach to generating data at the missing near offsets starts by first creating pseudo-primary data. Pseudo-primary data are created by crosscorrelating every trace with every other trace within a shot (Berkhout and Verschuur, 2003, 2005). The free-surface multiples correlate with the primaries at lags comparable to times when a primary reflection would arrive if one of the receiver locations was the source. Since the receivers now act as virtual sources, near-offset traces can be generated by crosscorrelating traces from nearby receivers, and zero-offset traces by autocorrelating a single trace. These crosscorrelated traces contain many spurious events from correlations other than those between primaries and free-surface multiples or between free-surface multiples and higher-order free-surface multiples, so that the pseudo-primary signal-to-noise ratio for a single crosscorrelation is poor. This can be improved by summing crosscorrelations of the same receiver location pair for many source positions, so the desired pseudo-primary correlations sum while the other correlations interfere destructively. The resulting pseudo-primaries are data that honor the kinematics of the recorded data, but also contain noise, have a different amplitude scale, and have a squared wavelet compared to the recorded data.

Simply substituting these pseudo-primaries for the missing data does not produce an adequate result, as the data contain a squared wavelet, a high level of noise, and spurious events that do not correspond to primary reflections. Instead, here we use the pseudo-primary data as training data for a nonstationary prediction-error filter; whereas the data are inadequate as an interpolation result, they are quite acceptable as training data. The PEF estimation process is relatively insensitive to the phase, amplitude, and squared-wavelet issues that make direct substitution undesirable. This PEF is then used for the interpolation step to fill in the inline near-offset gap.

This approach of using the pseudo primaries as training data for a PEF can be performed in the time and offset ($t-h$) domain, the time, offset and source ($t-h-s$) domain or the frequency, offset and source ($f-h-s$) domain. In the $t-h$ domain, we interpolate each shot record independently, using a separate non-stationary 2D $t-h$ PEF generated from the corresponding pseudo-primary shot, while in $t-h-s$ a single PEF is estimated for the entire data set. In the $f-h-s$ domain, the process is performed on overlapping time windows of shot records NMO-corrected using water velocity. A non-stationary 2D PEF is estimated on each frequency slice of each time window of the pseudo primaries, which is then used to interpolate the near-offset gap of the data for the same frequency slice and time window of the original data. The interpolated result is then inverse Fourier transformed back to the time domain, the time windows reassembled, and the NMO correction removed.

GENERATING PSEUDO-PRIMARIES

Multiple reflections are typically viewed as undesired noise to be removed from reflection seismic data. One way to do this is first to predict the multiples and then subtract them from the data. Free-surface multiples can be predicted by autoconvolving data so that the convolution of primaries with themselves creates events at arrival times that are the same as those of multiple reflections with a single bounce point at the free surface. This approach creates free-surface multiple reflections with correct kinematics without the need for any additional subsurface information except the recorded data (Riley and Claerbout, 1976; Reiter et al., 1991; Verschuur et al., 1992).

Autocorrelation can be used to extract synthetic active source data from incoming waves that reflect at the free surface and return within the recording array (Claerbout, 1968; Cole, 1995; Schuster, 2001; Wapenaar, 2003; Artman, 2007). This has traditionally been thought of in a passive context, where random noise is assumed to be arriving from all locations. The reflecting waves from active-source experiments can be treated in the same manner, where the primary reflections correlate with the free-surface multiples (Reiter et al., 1991; Berkhout and Verschuur, 1994, 2003; Shan, 2003). Autocorrelating data, d , (implemented as multiplication of complex-conjugates in the f domain) for two receiver points, r_1 and r_2 , both for a single shot, s , gives the pseudo primaries, p ,

$$p(s, r_1, r_2, f) = d(s, r_1, f)\bar{d}(s, r_2, f). \quad (1)$$

One of the receiver coordinates, r_1 , becomes the virtual source location, while the other receiver coordinate, r_2 , remains the receiver location, or vice-versa, and the \bar{d} denotes the complex-conjugate of d . This is similar to 2D surface-related multiple prediction, where instead of cross-correlation the data are convolved with itself, so that the convolution of primaries with primaries produces surface-related multiples (Verschuur et al., 1992).

Figure 2a shows an example of a fully-sampled split-spread shot from the Sigsbee2B data set, while Figure 2b shows a slice of the pseudo-primary output of autocorrelating the zero-offset trace with all of the other traces within that single shot. Clearly, the original shot and the pseudo primaries generated from the autocorrelation are different. The water-bottom reflection at zero offset is present as are the first group of diffractors, but they are not present at the more distant offsets. The many other correlations between events produce undesired noise in the output. Primaries can correlate with other primaries, multiples with multiples, and noise with other noise or signal. Much of these undesired correlations, however, vary as

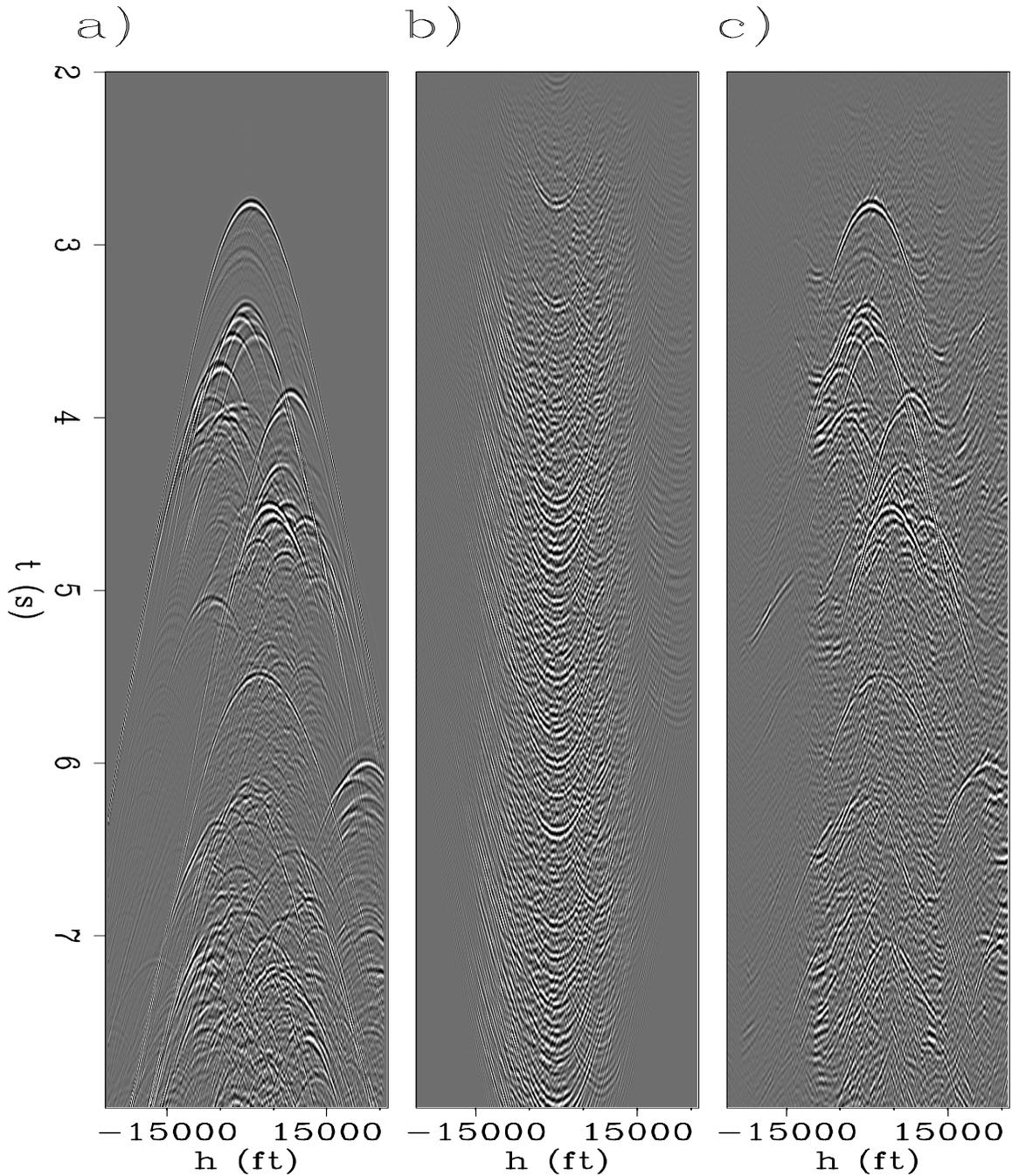


Figure 2: Crosscorrelation of a single shot. (a): Original fully-sampled split-spread shot; (b): The same shot recreated from crosscorrelating the traces in (a) using equation 1. (c): correlations are made for 496 shots and are then summed using equation 2. The quality of the pseudo-primaries is poor for a single shot, but improves after the summation of many shots. All data are scaled by t . [CR] `bill2/. sigsbee1shot`

a function of source position, s , so when the correlations for the same receiver pair r_1 and r_2 are performed for many different source locations, s , and are then summed, the unwanted events destructively interfere, while the correct pseudo-primaries constructively interfere.

$$p(r_1, r_2, f) = \sum_s d(s, r_1, f) \bar{d}(s, r_2, f). \quad (2)$$

As shown in Figure 2c, where 496 sources were used, this summing over multiple source positions thus can greatly improve the pseudo-primary signal. The range of usable offsets is greatly improved, as is the signal-to-noise ratio. The pseudo-primaries now look more similar in character to the original data.

Pseudo-primaries generated from recorded multiples are interesting in part because they have different illumination than do the recorded primaries. Figure 3a illustrates a desired near-offset primary ray-path that is not recorded because of the gap between the source and the nearest receiver. In Figure 3b, the source is positioned such that the raypath first travels from the source to the water-bottom and back up to the water surface within the recording array. This ray then reflects back into the subsurface and eventually returns into the recording array at another receiver. This four-segment raypath is a free-surface multiple, which can be reconsidered as two distinct events: the first, a recorded primary event from the source s to the first receiver r_1 and the second, another primary from the virtual source r_1 to another receiver r_2 . Crosscorrelating the multiple recorded at r_2 with the primary recorded at r_1 produces a pseudo-primary trace with a virtual source at r_1 and a receiver at r_2 . Comparing Figures 3a and 3b, we see that we can transform multiples that we record into pseudo-primaries with virtual source locations where we did not originally record data. This is most useful at unrecorded near offsets.

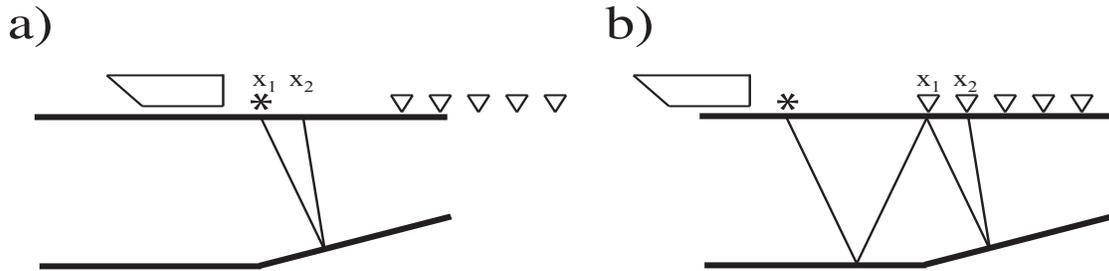


Figure 3: Raypaths of a primary and multiple reflection. (a): an unrecorded primary reflection that hits the surface at near offset. (b): a recorded multiple that first reflects at the surface within the recording array, then returns within the recording array. [NR]

bill2/. nearoffsetcartoon

The generated pseudo-primaries differ from the recorded data in several ways. First, the crosscorrelation of the data squares the amplitude spectrum of the pseudo-primaries relative to the original data, meaning that the wavelet of the pseudo-primaries will become zero-phase, different than that of the input data. Second, the amplitudes of the pseudo-primaries will differ from the amplitudes of the actual primaries, as the pseudo-primaries are a correlation and summation of different events. This amplitude difference will be both on a global scale difference between the pseudo-primaries and the primaries, as well as different

relative amplitudes within each data set. Third, spurious events from other correlations may still exist in the data as the number of sources is limited.

Because of these differences between the pseudo-primaries and primaries, a direct substitution of the created near-offset pseudo-primaries is not adequate. The pseudo-primaries, however, can be used as training data for a nonstationary prediction-error filter, as the PEF is insensitive to the phase of the training data. This PEF is then used to interpolate the missing near offsets, so that the negative aspects of the pseudo-primaries, such as the incorrect wavelet and extra noise, do not affect the PEF, while the positive aspects of the pseudo-primaries, contained in the autocorrelation of these data, are used.

INTERPOLATING WITH NONSTATIONARY FILTERS IN T - X OR IN F - X

Interpolation can be phrased in two steps, where a nonstationary prediction-error filter is first estimated from fully-sampled *training* data and is then used to interpolate *missing* data. The training data required by this method should have the same local autocorrelation as that of the desired output interpolated data. Pseudo-primaries, having roughly the same dip as the missing data, serve as training data for a PEF. Some of the problems with the pseudo-primary data, such as the different phase and amplitude, do not influence the PEF.

To estimate a nonstationary PEF, we solve,

$$\begin{aligned} & \min_{\mathbf{f}_{\text{ns}}} \|\mathbf{r}_d\|^2 + \epsilon^2 \|\mathbf{r}_f\|^2, \text{ where} \\ \mathbf{r}_d &= \mathbf{D}_{\text{ns}} \mathbf{K}_{\text{ns}} \mathbf{f}_{\text{ns}} + \mathbf{d} \\ \mathbf{r}_f &= \mathbf{R} \mathbf{f}_{\text{ns}} \end{aligned} \quad (3)$$

Here, the unknown nonstationary PEF coefficients, $\mathbf{K}_{\text{ns}} \mathbf{f}_{\text{ns}}$, are estimated from the pseudo-primary data, \mathbf{d} , a convolutional matrix, \mathbf{D}_{ns} , that is a function of \mathbf{d} , and a regularization operator, \mathbf{R} , that applies a Laplacian filter across the spatial axes of the PEF coefficients. This system of equations is solved to estimate a multi-dimensional, nonstationary prediction-error filter from a set of fully-sampled pseudo-primaries, such as was generated in the previous section.

Once this filter has been estimated from the pseudo-primary data, the filter is used in a second least-squares problem. In this problem, we estimate an interpolated output, \mathbf{m} , composed of missing data, $\mathbf{m}_{\text{unknown}}$, and known data, $\mathbf{m}_{\text{known}}$, that, when convolved with the nonstationary PEF, produces a minimized output,

$$\mathbf{r}_m = \mathbf{F}_{\text{ns}} \mathbf{m}_{\text{unknown}} + \mathbf{F}_{\text{ns}} \mathbf{m}_{\text{known}}. \quad (4)$$

Here, the known nonstationary PEF convolution matrix, \mathbf{F}_{ns} , derived from \mathbf{f}_{ns} obtained in the previous step, is multiplied with both the known and unknown data values, $\mathbf{m}_{\text{known}}$ and $\mathbf{m}_{\text{unknown}}$, to create a known quantity and a term with the unknown values to be interpolated. These terms are summed to form the residual \mathbf{r}_m . We minimize the L_2 norm of this residual to estimate the interpolated data values.

Interpolation in time and space

The output pseudo primaries from equation 2 are in frequency, source position, and receiver position. We can reorganize this to time, offset, and source position, and then estimate a nonstationary PEF, either shot-by-shot by estimating 2D PEFs in time and offset, solving equation 3 for each shot, or on the entire data set with a single 3D PEF in time, offset, and source, solving a much larger version of equation 3. This PEF, or series of PEFs, is then used to fill in the missing near-offset gap using equation 4 either once or for each shot.

Interpolation in frequency and space

Interpolation in the frequency domain requires a different approach than in the time domain. Spitz (1991) shows that 2D plane waves can be predicted for a single frequency by using a 1D spatial prediction filter, because a single plane wave at each frequency appears as a complex sinusoid in space. The wavenumber of this sinusoid increases as a function of frequency, so each frequency requires a unique 1D PEF. Data containing a combination of plane waves appear as a summation of complex sinusoids at each frequency, which can still be predicted by a reasonably short 1D PEF. In three dimensions, this filter would be a 2D filter, in four dimensions a 3D filter, and so on, making the frequency-based approach a series of smaller problems whereas the same region in time would be solved as a single larger problem.

Nonstationarity is addressed in a slightly different manner in f - x . Along the spatial axes, a nonstationary PEF can be used to capture slopes that change as a function of position. Since a Fourier transform is performed on the time axis to convert both the training and interpolated data to the frequency domain, we implicitly assume that the slopes of the plane waves do not vary as a function of time.

We address this problem of time nonstationarity by breaking up the problem into overlapping time windows that we assume to be stationary. First, perform a water-velocity normal move-out correction on both the pseudo primaries and the original recorded data to (roughly) flatten them, and then break up the data into overlapping windows along the time axis. We perform the NMO to reduce the amount of energy crossing the boundaries between patches. From there, each time window of both the pseudo primaries and the recorded data is Fourier transformed along the time axis, so that the data are sorted into source, offset, frequency, and time window. A unique 2D nonstationary complex-valued PEF is estimated in source and offset on each frequency of each time window of the pseudo primary data by solving equation 3. This nonstationary source-offset (h - s) PEF is then applied to fill the missing data in offset on the corresponding time window and frequency of the recorded data by solving equation 4, and this series of problems is repeated for each frequency of each time window.

After the data are interpolated, they are first inverse-Fourier transformed back to time, then the time-windows are reassembled with appropriate weighting in overlapping regions, and finally are inverse NMO-corrected to return the data to their original form of time, source, and offset.

SIGSBEE DATA EXAMPLE

In order to test the effectiveness of this pseudo-primary-based method, we use a synthetic example where we already know the answer. To create a challenging test case, we take a split-spread version of the Sigsbee2B synthetic data set, with a zero-offset section and a shot gather shown in Figures 4a and 4b, respectively, and remove the nearest 2100 ft of offset on either side, for a total of 4200 ft of missing offset, or a gap of 29 traces at the 150 ft sampling in offset, shown in Figure 4c. While this gap is overly large for a single boat, a two-boat undershooting of an offshore platform could have gaps this large.

Let us first examine the cross-correlated pseudo-primaries prior to the summation in equation 2 for a single output trace from one receiver pair for all shots, both using the fully sampled input data, shown in Figure 5b, and using the data missing the near offsets, shown in Figure 5c. This would correspond to $p(s, r_1 = 41000, r_2 = 41000, t)$ in equation 1. We refer to these images as pseudo-primary-contribution gathers, as they fulfill a role similar to multiple-contribution gathers in the SRME algorithm (Dragoet and Jericevic, 1998). The summations of these gathers, the predicted pseudo-primary traces, are shown in Figure 5a, plotted alongside the original trace. We see that the traces in Figure 5a look quite similar, especially the two pseudo-primary traces, so the missing sources at near-offsets did not significantly detract from the result.

Moving up from the single output trace in Figure 5, we now look at an entire zero-offset section in Figure 6. Figure 6a shows the original zero-offset section while 6b shows pseudo-primaries generated using fully-sampled data, and 6c shows pseudo-primaries generated from data with missing near offsets.

Note four important differences between Figures 6a and 6c. First, the pseudo-primaries have different illumination than do the true primaries, so the relative amplitudes within each image differ. The amplitude of the water-bottom reflection in the pseudo-primary image with limited offsets in Figure 6c is more variable than is that in the original image in Figure 6a, or even that in the pseudo-primaries generated with all of the offsets in Figure 6b. Also, the subtle reflections below the water bottom on the left side of the image at 4-5s are much less pronounced when the input offsets to the pseudo-primary generation are limited. Second, the pseudo-primaries compared to the original data exhibit cross-talk. The cross-talk is composed of both coherent events such as those above the water bottom and water-bottom multiple in Figure 6c, and more random noise. This cross-talk increases only slightly when the offsets are limited in the input to the pseudo-primary generation. Third, the wavelet of the pseudo-primary data differs from that in the original data as a result of the cross-correlation of the primary wavelet and the multiple wavelet. Finally, the amplitude scale of the pseudo-primary images (6b-c) is roughly a factor of 20 higher than in the original data.

Figure 7 shows the nearest 10000 ft of offset of the pseudo-primaries generated from all of the split-spread shots from Figure 4c that are missing the near offsets. The common-offset section on the left panel is at zero offset. Looking at the shot gather on the right panel, we see that the water-bottom reflection is the strongest event at early times, but a considerable amount of cross-talk is present before the water-bottom reflection. The time slice shown on the top panel of Figure 7 shows that in addition to the desired reflections and diffractions extracted from the multiple reflections, there are also strong spurious events that are not easily identifiable as cross-talk; in particular, note the event at roughly 60000 ft and near

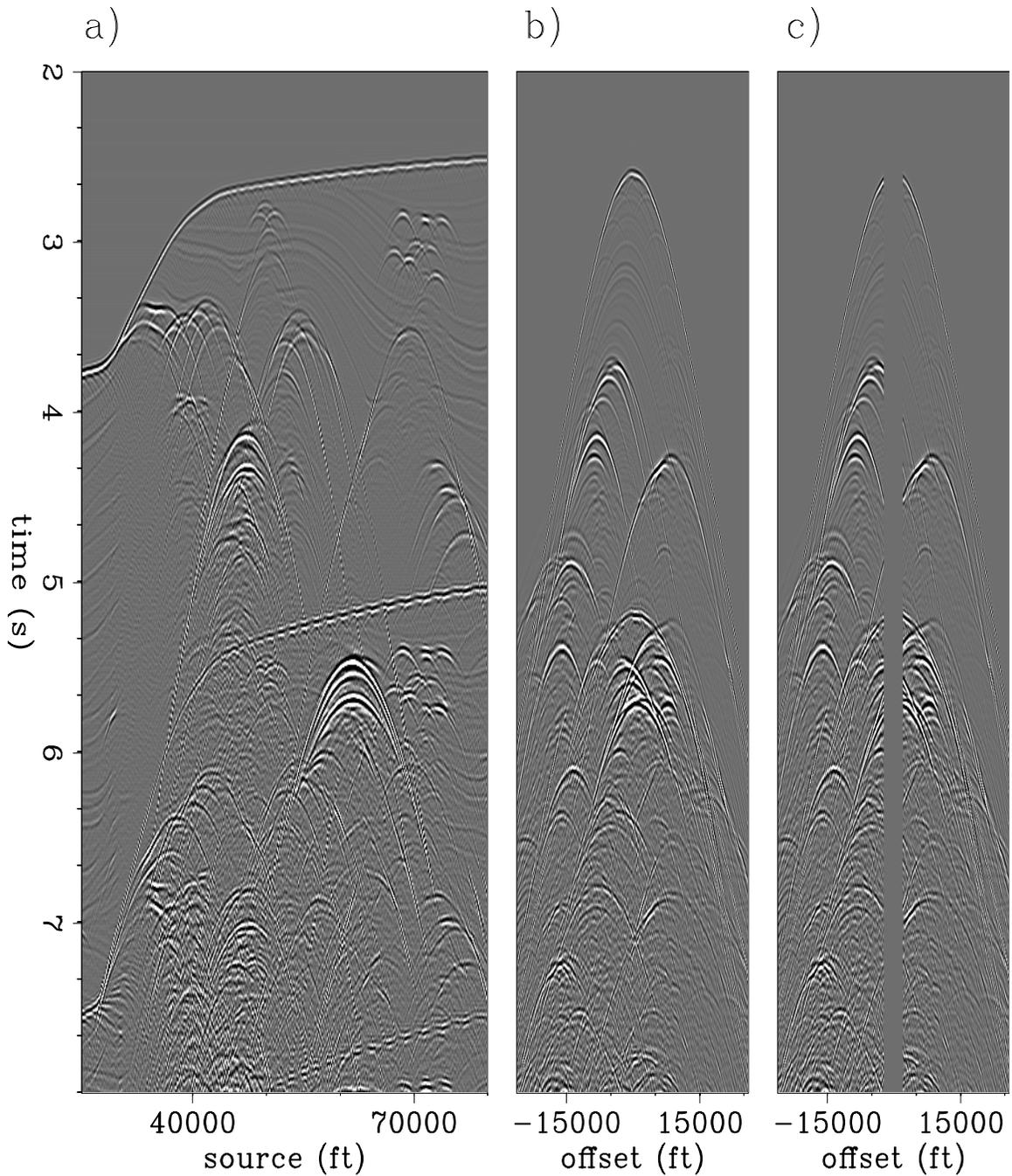


Figure 4: Original Sigsbee data. (a): zero-offset section. (b): one shot. (c): resampled shot. The near 4200ft or 29 traces of offset were excluded. All figures have amplitude scaled by $t^{0.8}$. [ER] `bill2/. splitspread`

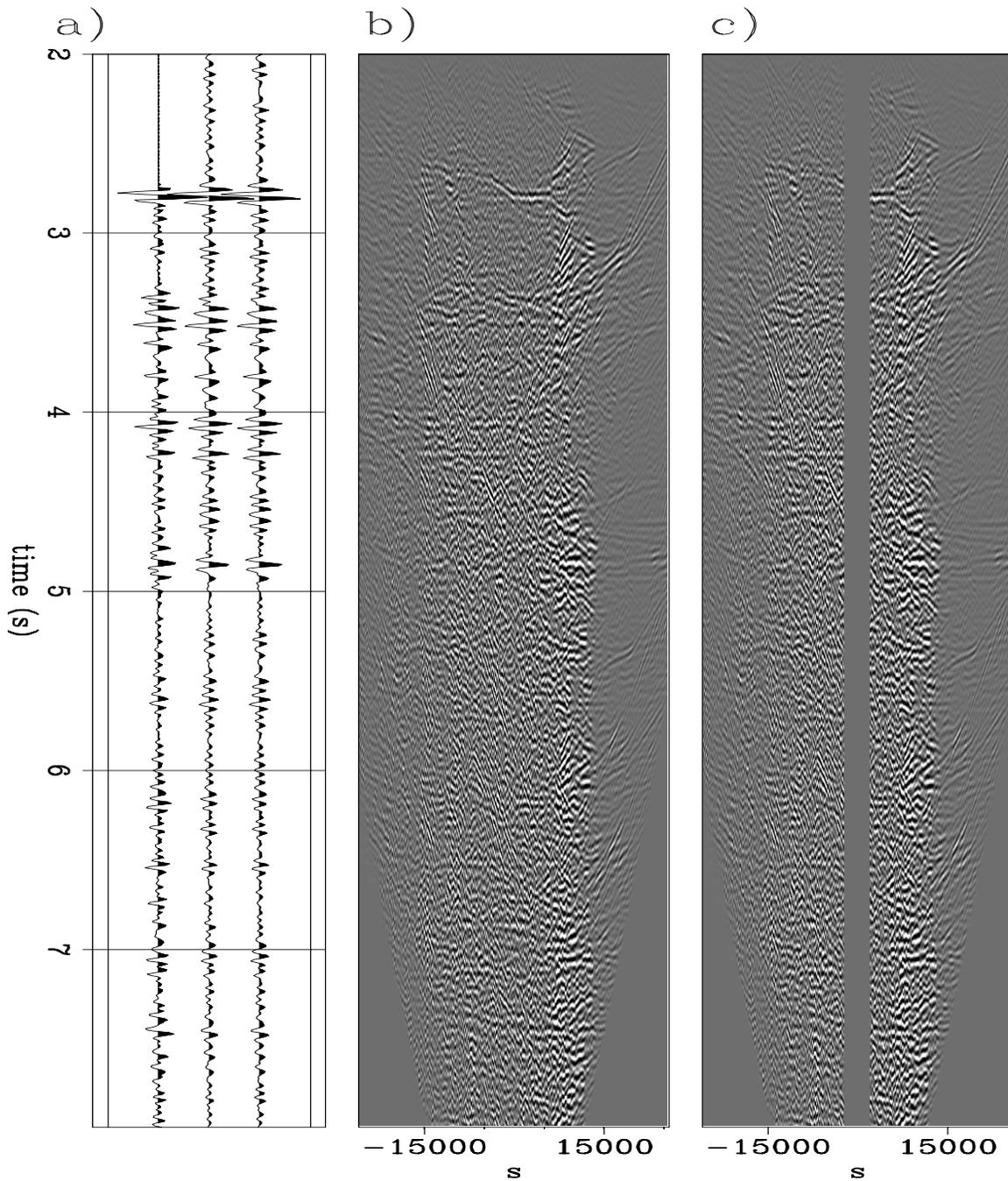


Figure 5: Generation of a pseudo-primary trace for $p(s, r_1 = 41000, r_2 = 41000, t)$. (a): comparison of traces of (left to right) original data, pseudo-primary data generated from fully-sampled input data, and pseudo-primary data generated from data with missing near offsets. (b): A pseudo-primary contribution gather, where the horizontal axis is shot location s . (c): The same pseudo-primary contribution gather as (b), but with the missing near offsets. The images are scaled by $t^{0.8}$ for display purposes. [CR] bill2/. pseudocontrib

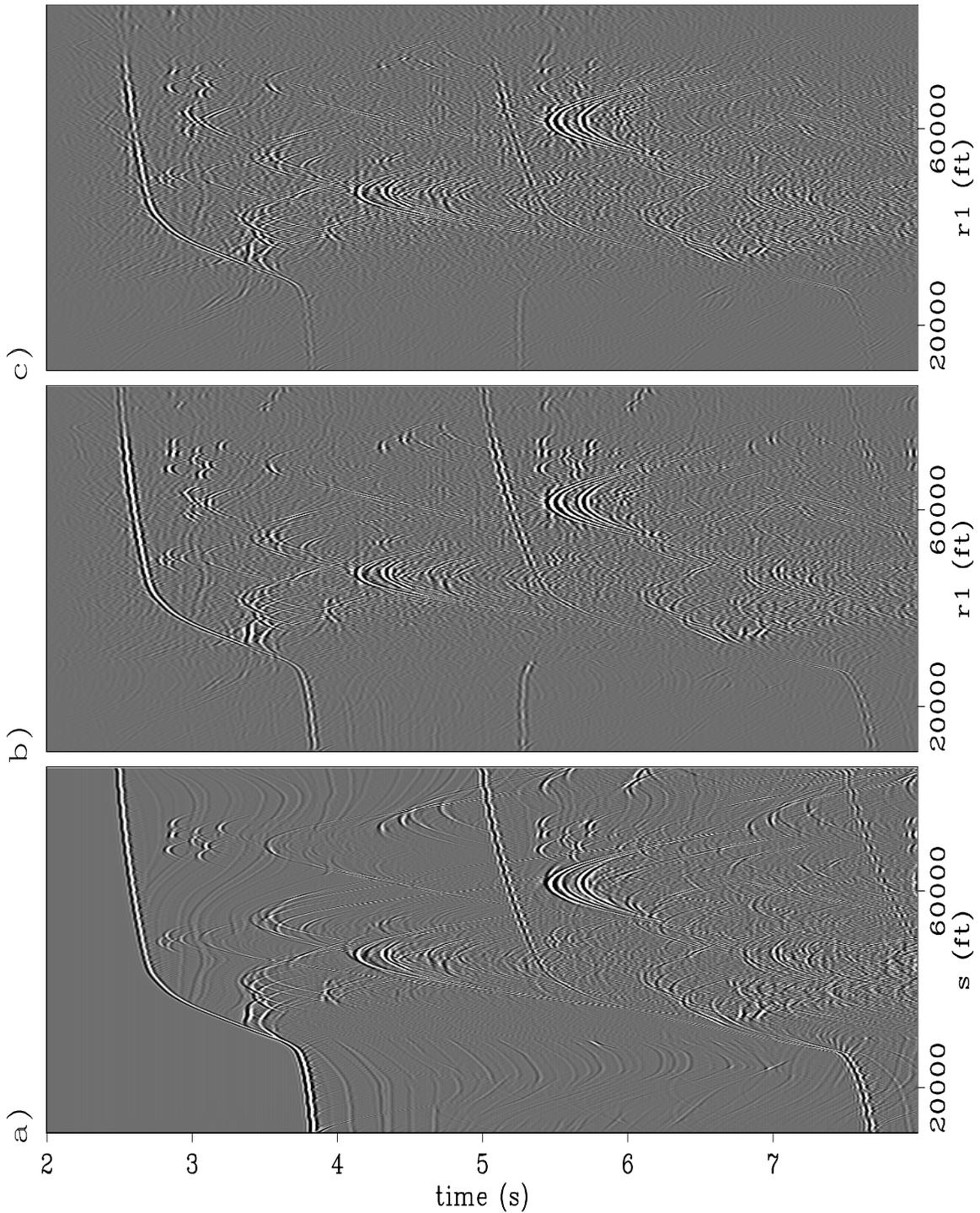


Figure 6: A comparison of zero-offset sections: (a) original data, (b) pseudo-primaries generated using all offsets, (c) pseudo-primaries generated using all offsets other than the missing near offsets. The quality of the pseudo-primaries degrades slightly with missing near offsets. The images are scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. pseudocontribslice`

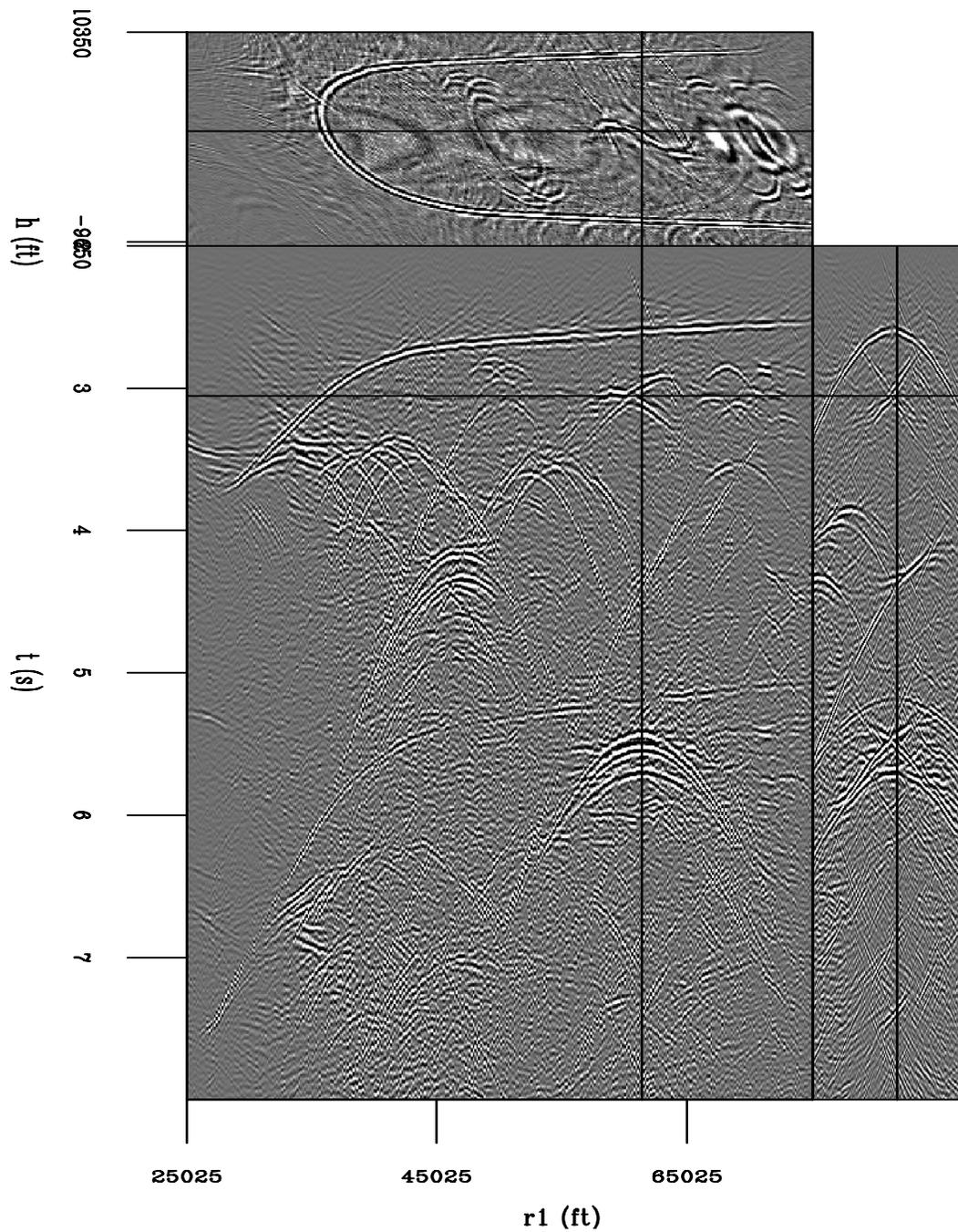


Figure 7: The near 10000 ft of offsets of pseudo-primaries generated from input data missing the 2000 ft of near offsets. The front panel is a constant-offset section, the right panel is a single shot, and the top panel a time slice. The image is scaled by $t^{0.8}$ for display purposes.

[CR] bill2/. pseudo

offsets.

The most straightforward way to use these pseudo-primaries to interpolate the missing near offsets would be to replace the missing traces with the pseudo-primaries, shown in Figure 8. The pseudo-primaries were scaled by a factor of 0.05 to match the mean amplitude of the sampled data. The cross-talk and squaring of the wavelet are both obvious in this Figure 8, making the region of interpolated data easily distinguishable, as do the spurious events seen at (1) and (2), and the extra slope in the more complex areas like (3). As we see next, we can improve greatly on this result by using the pseudo-primaries as training data for a prediction-error filter.

Interpolation of Sigsbee in time and offset

A single output shot record of the pseudo-primaries ($p(r_1, r_2, t)$) generated in Figure 7 is used as the training data (\mathbf{d}) in equation 3 to generate a nonstationary PEF, \mathbf{f}_{ns} . The regularization operator in equation 3, \mathbf{R}_{ns} , is a two-dimensional Laplacian that operates over time and offset for each local filter lag. Once estimated, the PEF is then used to interpolate the missing near offsets of the sampled data, $\mathbf{m}_{\text{unknown}}$, using equation 4. This process is repeated for all of the shots of both the pseudo-primary data and the original sampled data.

Figure 9 was generated by interpolating the near-offset gap using a t - h nonstationary PEF estimated on each shot record of the pseudo-primary data. Each 2D PEF has ten points in time and six points in offset, and varies every four points on each axis for a total of 76000 free coefficients for each shot. We applied 400 iterations of a conjugate-direction solver to solve both the filter-estimation and interpolation problems.

There are several interesting things to see in the t - h interpolated result in Figure 9. First, little cross-talk is present in the interpolation, such as the spurious event at (1) and (2). This is one of the benefits of using the t - h PEF-based approach, since low amplitude at the edges of the known data results in no large residuals in equation 4, regardless of the crosstalk present in the training data for the PEF. Second, the events present in the original data in a single shot shown in the right panel are interpolated, in most cases with good results, although the quality of the interpolation degrades deeper in the section. This can largely be traced to the quality of the input pseudo-primaries. It also appears that problems with the relative amplitudes in the pseudo-primaries are amplified in this result. Third, the common-offset section shows significant differences from shot to shot. This is because a 2D PEF is used and each shot is interpolated independently. This is different when a 3D PEF in time, offset *and* shot is used, where correlations between shots are used and the inconsistencies between shots penalized, discussed next. Finally, the wavelet issues in the pseudo-primaries appear to be mostly removed in the t - h interpolated result.

While the t - h approach gives a reasonable result when viewing a single shot, the extremely choppy constant-offset section shows one limitation of the t - h approach. We next use a 3D nonstationary t - h - s PEF that is $10 \times 5 \times 5$ elements and varies every 10 elements on the time axis, 3 elements on the offset axis, and 4 elements on the source axis. We solve for this PEF using 100 iterations of a conjugate-direction solver. Once estimated, we use this PEF, with 121 million coefficients, to interpolate the missing data, using 200 iterations of a conjugate-direction solver on equation 4, using a starting model for the unknown near

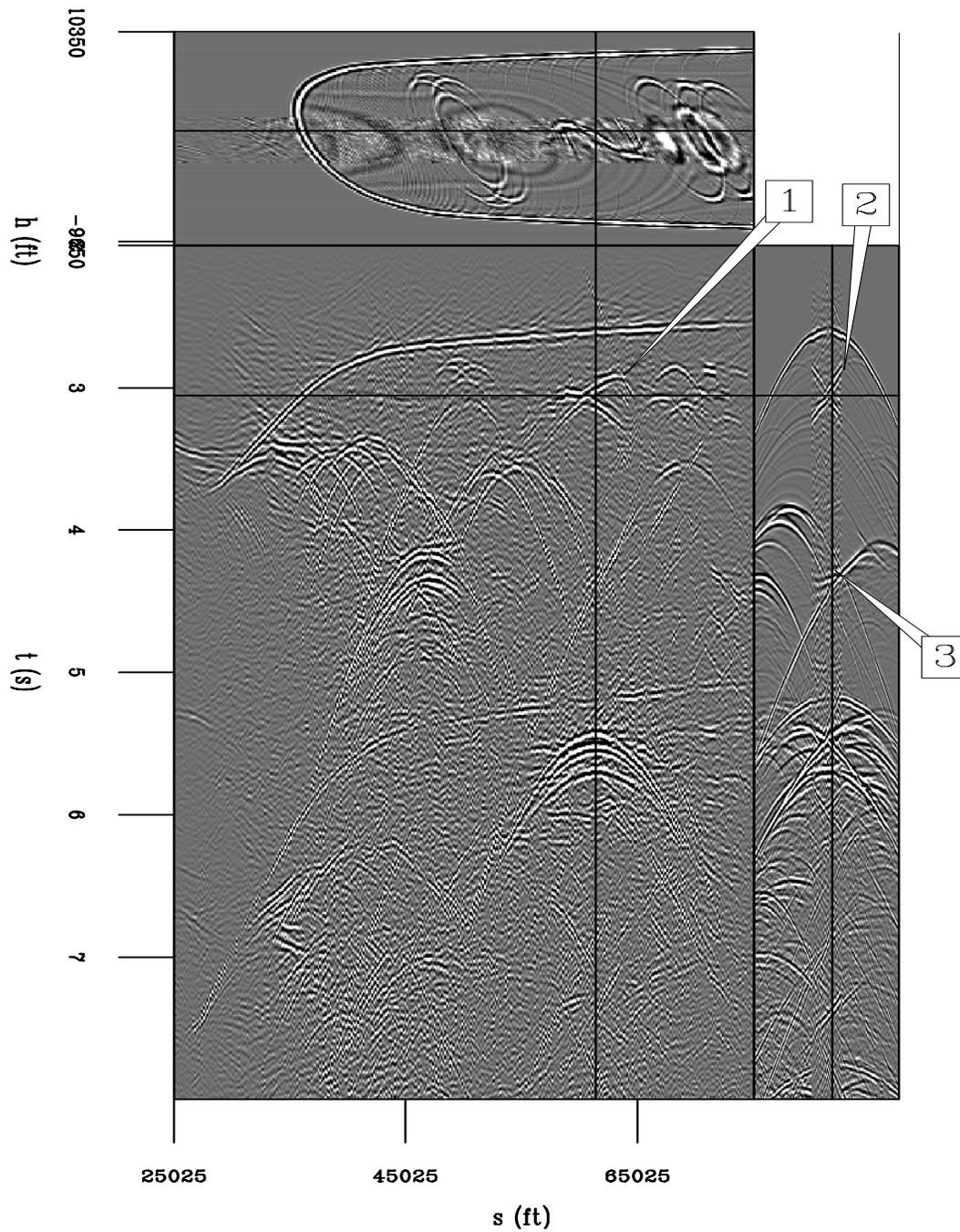


Figure 8: Original data missing the near offsets, with pseudo-primaries spliced into the locations where traces were missing. Note three points of interest, a spurious event caused by the correlation of primaries with other primaries, seen in both the constant-offset section (1) and a shot (2), and an area with crossing events and noise (3). The image is scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. pseudocut`

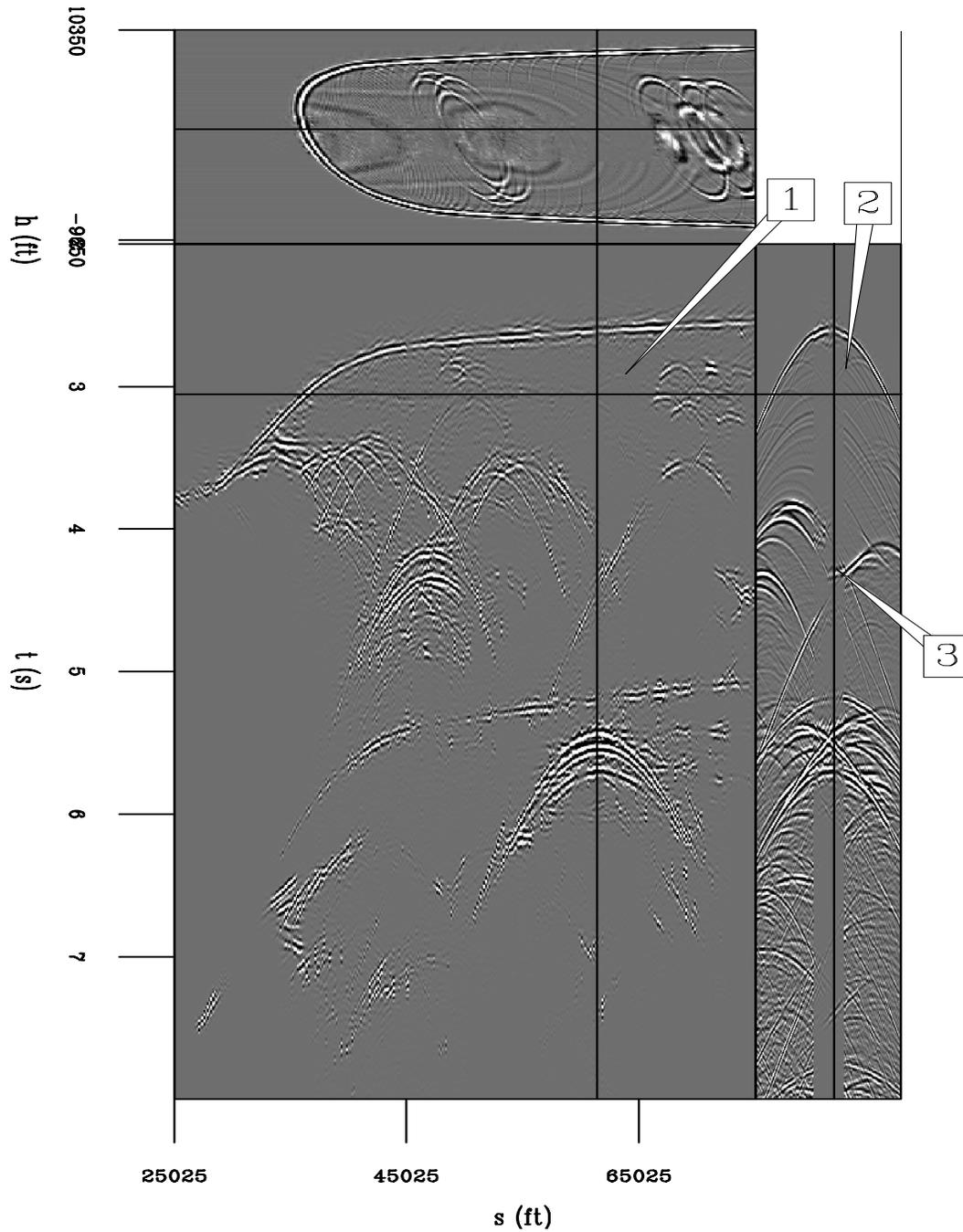


Figure 9: Interpolation with pseudo primaries and a t - h PEF. The front panel is an interpolated constant-offset section, the right panel is a single shot record, and the top panel a time slice. The spurious event in (1) and (2) is not present, but the region with conflicting slopes (3) is only partially interpolated, with an incorrect slope present. The image is scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. txinterp`

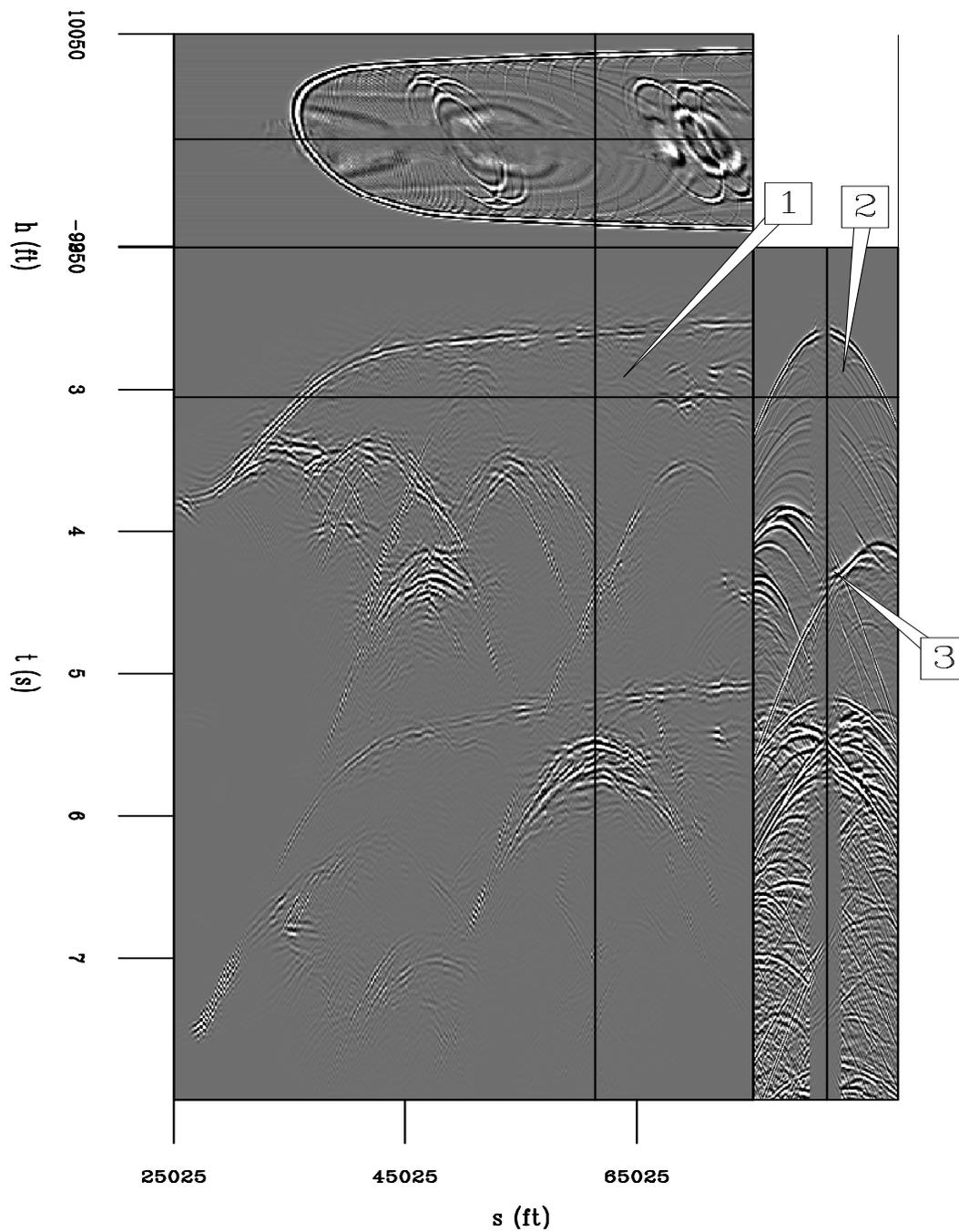


Figure 10: Interpolation with pseudo primaries and a 3D t - h - s PEF. The front panel is an interpolated constant-offset section, the right panel is a single shot record, and the top panel a time slice. The spurious event at (1) and (2) is not present and the region with crossing slopes (3) is better interpolated than in the 2D example. The water-bottom is more variable in amplitude, but the constant-offset section is more consistent from shot to shot. The image is scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. txyinterp`

offsets of NMO-corrected copies of the nearest recorded offset for that midpoint.

This 3D result, shown in Figure 10, is in part an improvement over the 2D approach. The roughness of the 2D approach along the source axis is gone, with a smoother result that contains slightly more noise. The spurious event present at (1) and (2) is still gone, and the the diffractions in the constant-offset section are more continuous, although the water-bottom amplitude is more variable. Next, we show that using a 2D PEF on frequency slices provides most of the benefits of using a 3D t - h - s PEF, but with a much lower computational cost and memory requirement.

Interpolation of Sigsbee in frequency, offset, and shot

Figure 11 shows the result of near-offset interpolation using 2D PEFs in the f - h - s domain. Here the training and sampled data were first NMO-corrected at water velocity, then were broken into overlapping windows along the time axis. Both sets of these time windows were then Fourier transformed to frequency. Each frequency slice of each window is treated independently, where we estimate a nonstationary PEF on a frequency slice of the pseudo-primaries, solving equation 3 with the frequency slice of the pseudo-primaries serving as training data (\mathbf{d}) to generate a nonstationary complex-valued PEF \mathbf{f}_{ns} . This PEF is then used in equation 4 as \mathbf{F}_{ns} , where $\mathbf{m}_{\text{known}}$ is the corresponding frequency slice of the recorded data, and $\mathbf{m}_{\text{unknown}}$ are the missing near offsets of this frequency slice.

In Figure 11 the time axis (1500 samples long) of both the input sampled data and the pseudo-primaries were broken into 40 overlapping windows of 64 points each. Both sets of windows were then Fourier transformed to produce 2560 2D frequency slices. Each frequency slice is treated as a separate problem, with a 2D nonstationary PEF estimated on the pseudo-primary frequency slice; the 2D PEF was four samples long on both the offset and shot axes, and the filter varied every four points on each axis, for a total of 67456 filter coefficients for each 496-shot by 136-offset frequency slice, estimated with 100 iterations of a conjugate-direction solver. The data were then interpolated using 200 iterations of a conjugate-direction solver on equation 4, again with an initial model of the missing data an NMO-corrected nearest-offset-trace from the same midpoint.

The frequency-domain interpolation in Figure 11 differs in several respects from the t - h domain interpolation in Figure 9. First, the f - h - s result still has some crosstalk present, while the t - h and t - h - s results have almost no crosstalk from spurious events. This is most visible prior to the water-bottom reflection. We believe this is in part due to the stationarity assumption within each time window. However, the large spurious event at a (1) and (2) is not present in the result in either case. Second, the f - h - s result shows less shot-to-shot variation than does the t - h result, and less amplitude variability than the t - h - s result. This is because in the t - h result each shot is interpolated separately as an independent problem, while the f - h - s interpolation estimates a PEF that spans both the offset and source axes, minimizing this jitter. Finally, the f - h - s result appears to contain more detail in the result than does the t - h result. In particular, events below the water-bottom in the shot that were not interpolated in the t - h case, in the right panel of Figure 9, were interpolated in the f - h - s case on the right panel of Figure 11.

This synthetic data example produced promising results for interpolation of a large gap using information only from the recorded data. The f - h - s approach improved results and

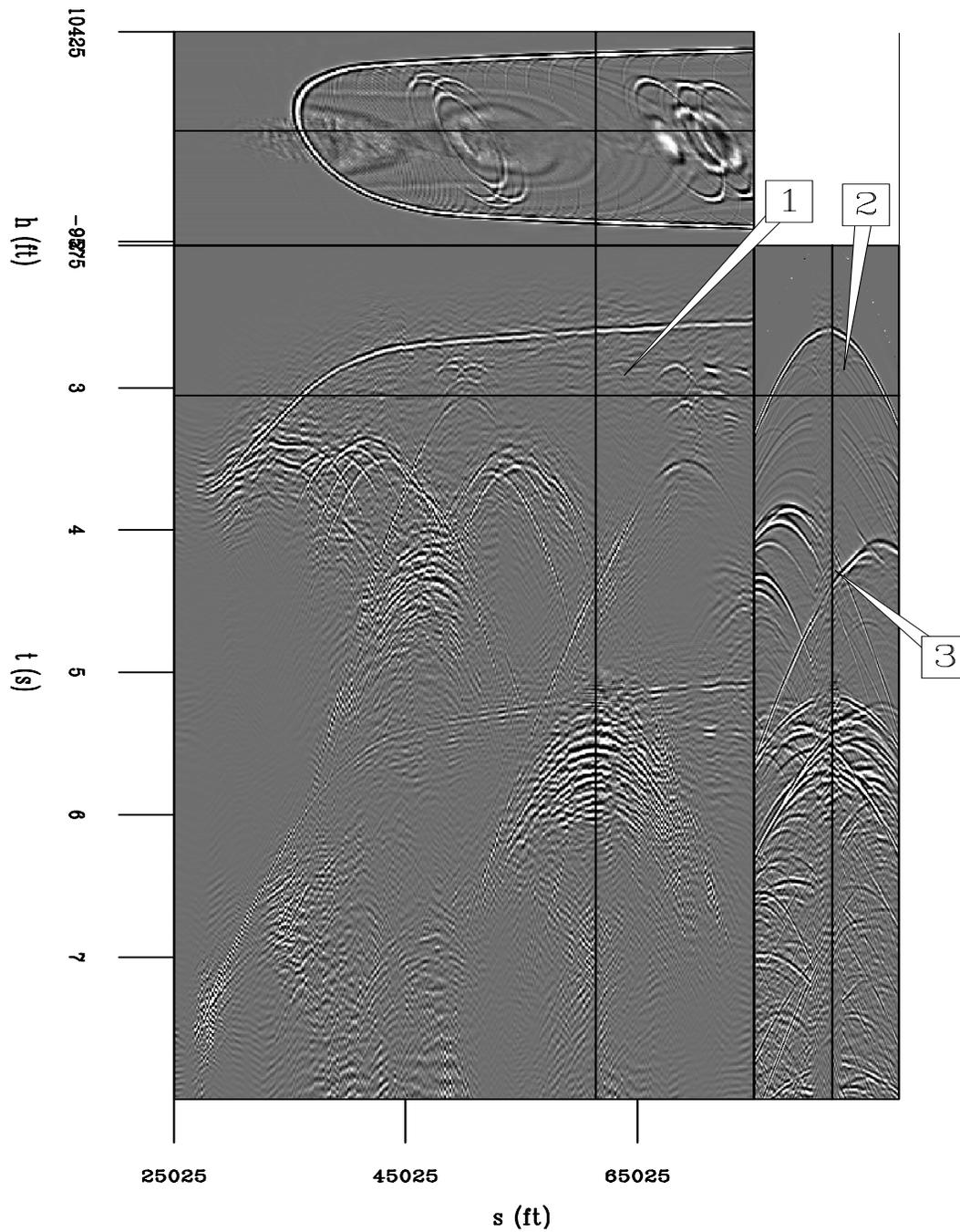


Figure 11: Interpolation with pseudo primaries with 2D f - h - s PEFs. The result is more consistent from shot-to-shot, but still contains some cross-talk (around the water bottom) from the pseudo-primary data. The spurious event (1 & 2) has been removed and the crossing slopes (3) are believably interpolated. The image is scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. fxinterp`

was still multiple times faster than the $t-h$ or $t-h-s$ interpolation methods. The synthetic data were noise-free, and only contained the desired primaries and multiples, with an idealized 2D geometry. Next, we examine how this method works on field data.

FIELD DATA EXAMPLE

Many have used the Mississippi Canyon multiples dataset from WesternGeco as a benchmark for testing multiple removal methods (Verschuur and Prein, 1999; Guitton and Cambois, 1999; Hadidi et al., 1999; Dragoset, 1999; Lamont et al., 1999; Lokshtanov, 1999; Matson et al., 1999; Berkhout, 1999; Weglein, 1999). We now use the strong free-surface multiples present in these data to generate pseudo-primaries. Figure 12 shows a source-offset cube of the recorded data; we created split-spread data from the off-end marine acquisition using source-receiver reciprocity. The near-offset gap in this case is six traces, which is much smaller than that in the Sigsbee2B example. However, these field data contain much more than just primary and multiple reflections.

Pseudo-primary generation from field data

The pseudo-primary-contribution gather, shown in Figure 13, does contain the water-bottom, top-of-salt, bottom-of-salt reflections, and water-bottom multiples, caused by the correlation between the water-bottom reflection and the first-order multiple of the water-bottom, top-of-salt, bottom-of-salt, and second-order water-bottom multiple, respectively. This cube is a series of cross-correlations of single traces, with no amplitude scaling or additional deconvolution performed before cross-correlation. The signal-to-noise ratio in the data is lower than that in the synthetic example. The side panel shows the same receiver-pair correlation for multiple shots that are later summed to produce a single output trace.

Figure 14 shows the volume of pseudo-primaries generated by summing the cross-correlations from all of the sources in the recorded data. The pseudo-primary data are somewhat poorer than the recorded field data in Figure 12, with the long wavelet present in the original data strongly featured in the pseudo-primaries. Additionally, when compared to the pseudo-primaries in the synthetic case, the quality of the pseudo-primaries for this example is significantly worse. Reviewing three of the points we made about the pseudo-primaries in Sigsbee, we see that these problems are even more pronounced in this field data example. For example, while there were variations in the relative amplitude of the Sigsbee pseudo-primaries, the variations in the pseudo-primaries from the field data are much more obvious; where the water-bottom reflection is dipping, the pseudo-primary reflection is nearly absent, and the pseudo-primaries of the salt body have higher amplitude than anywhere else. Second, the amplitude spectrum in the Sigsbee example was squared as a result of the cross-correlation involved in pseudo-primary generation. The much longer wavelet in the field data makes this more obvious, with the side-lobes of the water-bottom reflection appearing before the water-bottom reflection in the recorded data. This ringing is probably associated with a water bubble from the source. Third, the cross-talk in this field data is also present in this pseudo-primary result. These slight smile-shaped events have a slope the direction opposite that of the desired pseudo-primaries, and are more obvious at farther offsets, although the base of these smiles are at roughly zero-offset in Figure 14.

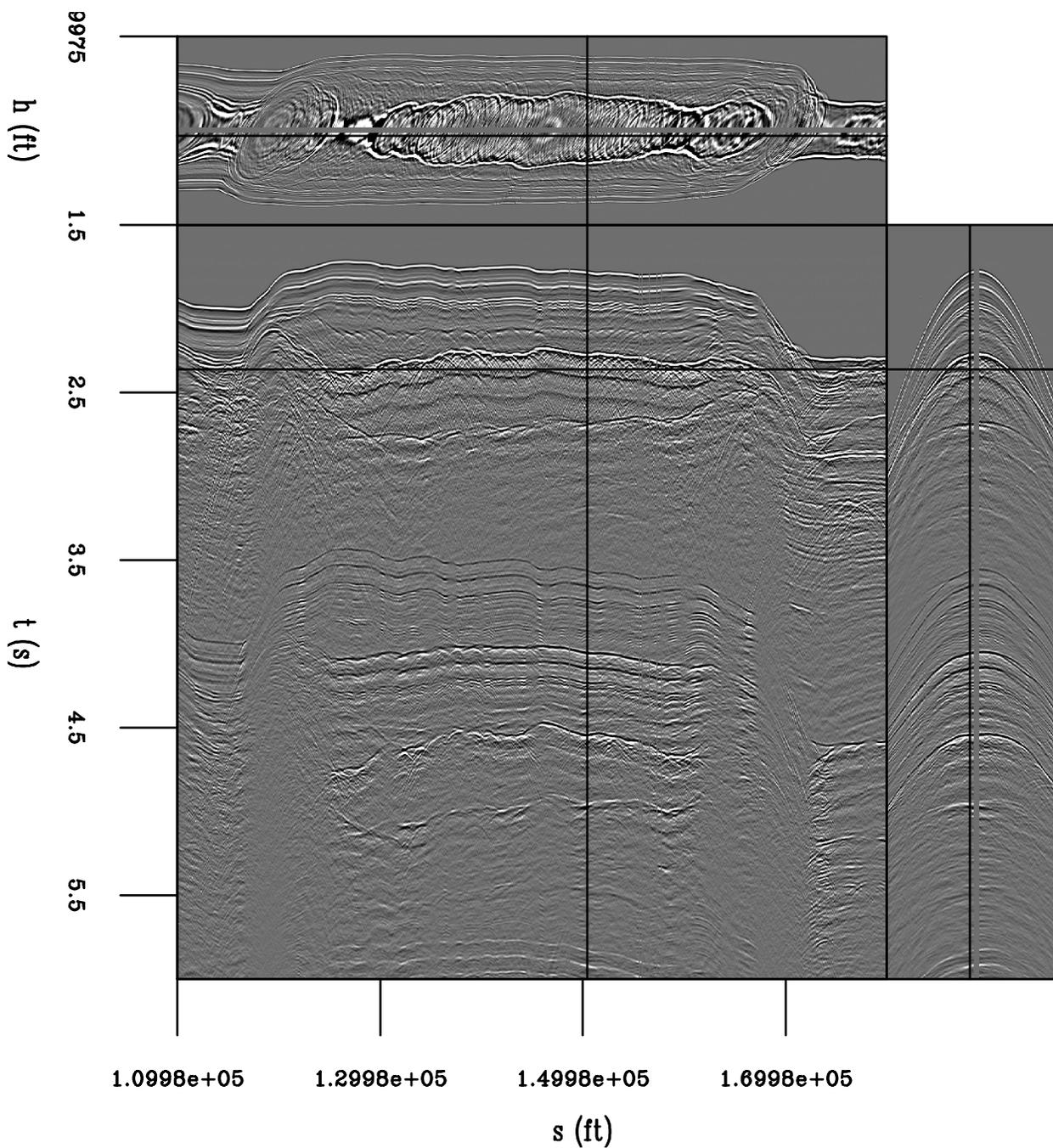


Figure 12: A Gulf of Mexico dataset. The front panel is a constant-offset section, and the side panel is a single shout, with the negative offsets predicted by reciprocity. The image is scaled by $t^{0.8}$ for display purposes. [ER] `bill2/. fieldin`

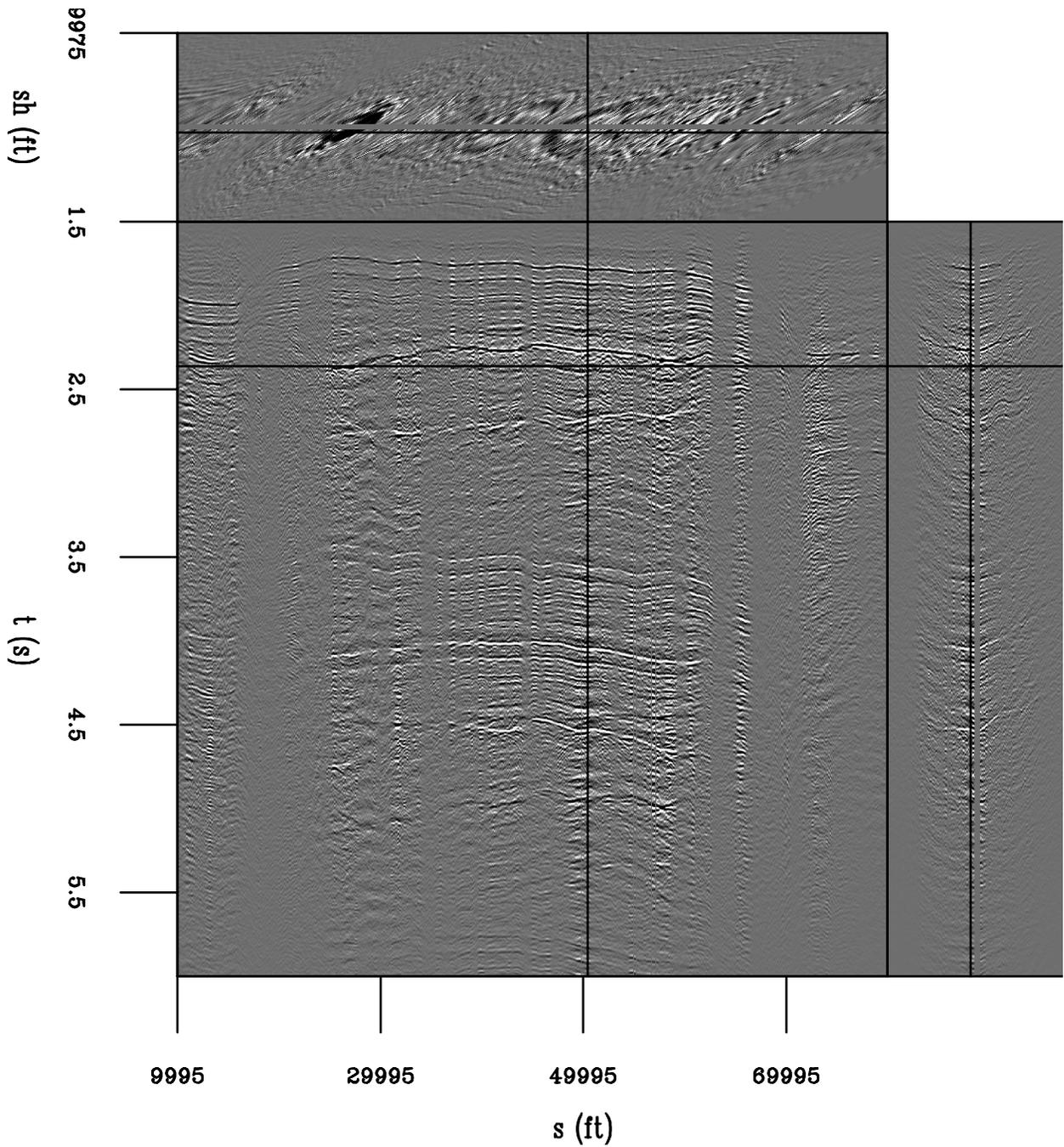


Figure 13: Pseudo-primary contribution gather. The front face is a constant-offset section, while the side face contains the shots that will be summed to produce the pseudo primaries. The image is scaled by $t^{0.8}$ for display purposes. [CR] [bill2/. fieldcontrib](#)

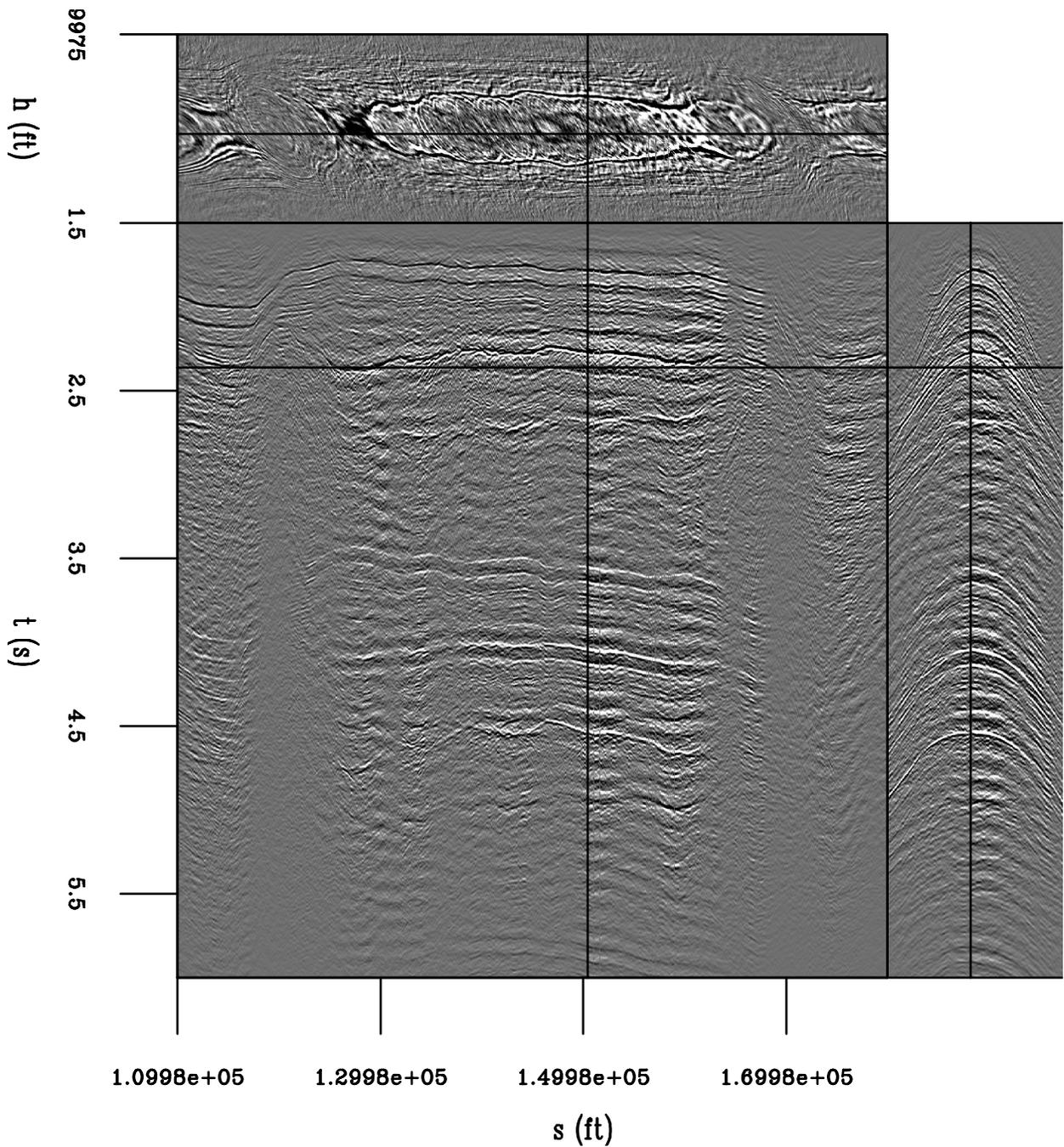


Figure 14: Pseudo-primaries of data. The front face is a constant-offset section, and the side face is a shot gather. The gross structure of the original data is present, but the squared wavelet strongly dominates the data. The image is scaled by $t^{0.8}$ for display purposes. [CR]

bill2/. fieldpseudo

Interpolation of field data in f - h - s

As in the Sigsbee example, we first take the pseudo primaries in Figure 14, perform a water-velocity normal moveout, and then break the result into 150 overlapping time windows of 32 samples each. These time windows are then Fourier transformed into 150×16 source-offset frequency slices. We then estimate a 3×4 nonstationary PEF that varies every sample along the source and offset axes for each of these slices, and use it to interpolate the corresponding slice of the recorded data with the near-offset gap. For each frequency of each time patch, we use 80 iterations of a conjugate-direction solver both to estimate the PEF and to interpolate the missing data. The data are then transformed back to time, the windows are reassembled, and the NMO correction is reversed. This result is shown in Figure 15.

The interpolation result in Figure 15 is greatly improved over that in the pseudo-primary data. The most obvious differences between the pseudo primaries and the interpolated result are that the ringing wavelet has been removed and the polarity of the images differ. The dipping portions of the water-bottom that were not present in the pseudo primaries were not interpolated. While this gap is much smaller than that in the Sigsbee example, the result is not as impressive. The water bottom, top of salt, bottom of salt, and the associated multiples are reasonably well interpolated, but the subtle stratigraphic reflectors outside of the salt body and the steeply dipping portions of the water bottom are not present in the interpolated result. For example, the diffractions below the bottom of salt reflection or at the edges of salt are not present in the interpolated result. The subtler stratigraphic reflections flanking the salt body are well interpolated, even at later times.

PSEUDOPRIMARIES IN 3D

The pseudoprimaries generated so far have been either for a two-dimensional synthetic model or for two-dimensional field data example with strong multiple reflections. We now attempt a 3D pseudoprimaries generation on 3D prestack synthetic data where the geology is fully known and the acquisition is along ideal overlapping straight lines.

Figure 16 is an illustration of the source distribution for a single receiver line from a 3D synthetic dataset, courtesy of ExxonMobil. The data are composed of a horizontal water-bottom reflection under which there is a prism filled with point diffractors. At slightly below the arrival of the diffracted multiples, three reflectors are present. The acquisition geometry, shown in Figure 16, is ideal, so that at each subsequent sail line the cables shift by a distance exactly equal to that between four receiver cables. As such, any one receiver cable location occurs for three sail lines given the recording aperture of 550 m, which when multiplied by the two crossline source positions per sail line, gives six crossline source contributions for any one crossline receiver location under these ideal circumstances.

Figure 17 is the result of creating pseudoprimaries for zero crossline offset for a single receiver cable, where each active receiver location is cross-correlated with every other receiver location along the same cable for all sources in the three sail lines. These 640 sources for each of the three sail lines are added together to produce the pseudoprimaries image in Figure 17, generated from roughly 4 TB of individual crosscorrelations. While the water-bottom reflection initially appears to be in the correct place, the diffracted multiples are a blurred mess at near offsets and are absent at the far offsets where the limbs of the

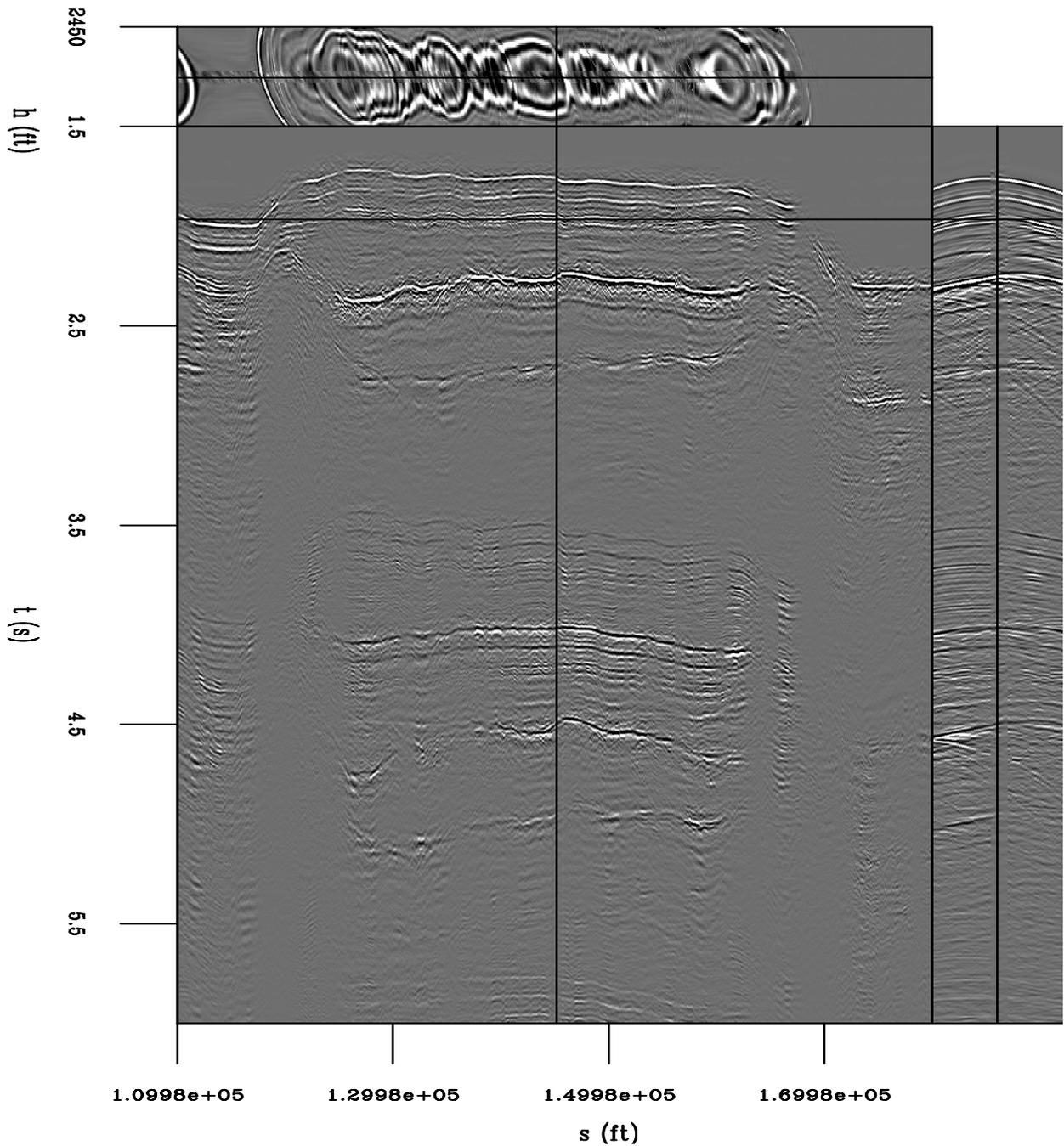


Figure 15: Interpolation of field data with a 2D f-h-s PEF using pseudo-primaries. The front face is a constant-offset section, and the side face is a shot. The near offsets appear much more reasonable than the pseudo-primaries in Figure 14. The top and side panels are zoomed in relative to earlier figures in order to emphasize the interpolated values. The image is scaled by $t^{0.8}$ for display purposes. [CR] `bill2/. fieldfx`

Figure 16: Crossline source distributions for a pseudoprimaries generated from a single receiver cable at zero offset. [NR] `bill2/. 3dpseudo`

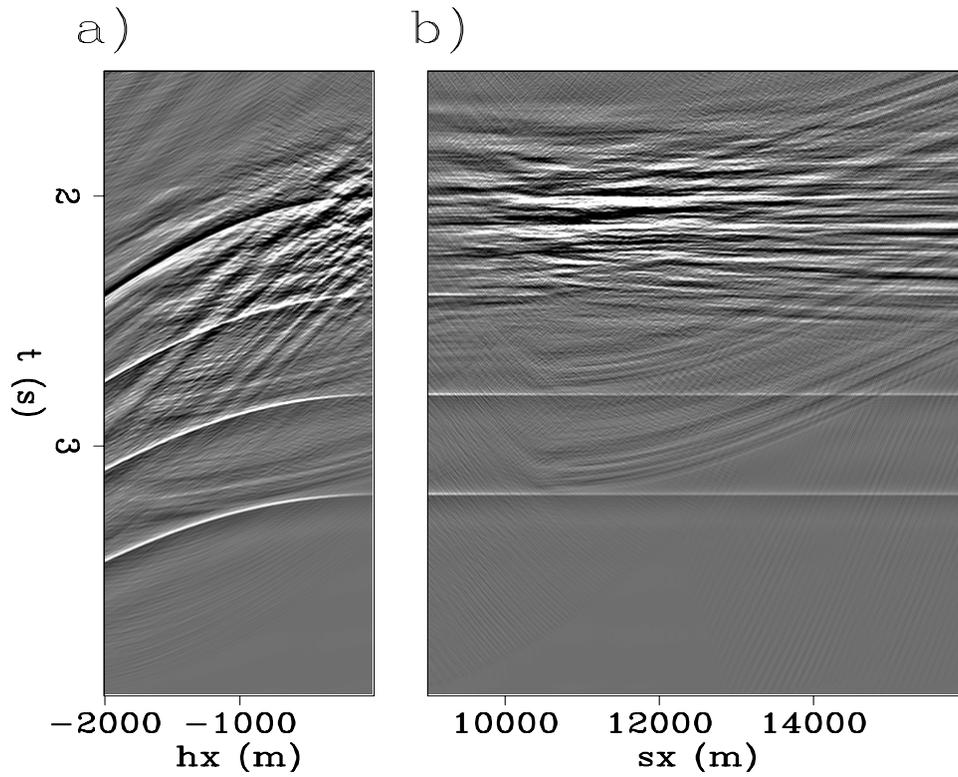
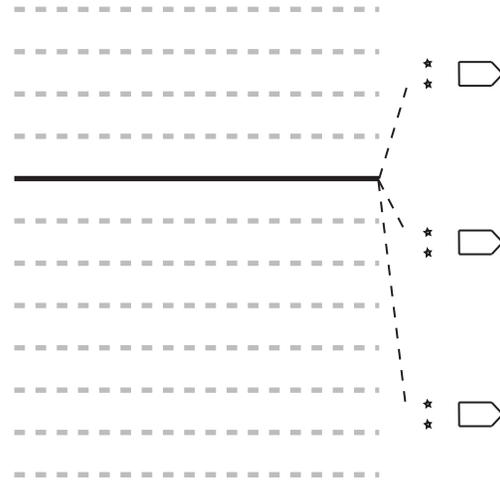


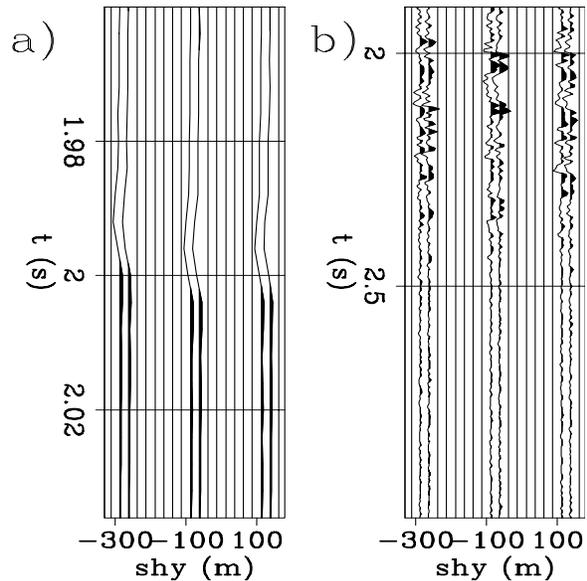
Figure 17: pseudoprimaries generated for a single receiver cable from six different crossline source positions. [CR] `bill2/. 3dsynpseudo`

recorded diffractions are in the recorded data. There are also apparent reflectors below the water bottom that are created from the correlation of the water-bottom primary reflection with the much deeper series of three primary reflections at below the water-bottom multiple arrival time, illustrating the potential pitfalls of unwanted correlations.

In order to determine the variability of the predicted pseudoprimaries with crossline source position, we have produced a crossline gather similar to the pseudoprimary contribution gathers we produced earlier in this chapter. While the full 3D uncollapsed cube of crosscorrelations would have 7 dimensions: time, inline and crossline source, inline and crossline first receiver, and inline and crossline second receiver, we only show the predicted multiples for a single first receiver (virtual source), second receiver, and inline source position, showing how the predicted multiples vary as the crossline location of the source varies. Figure 18 contains two of these gathers, where 18a is a zoom in on an area of the data with little diffracted multiples at the water bottom. The arrival time of the water bottom changes as a function of distance from the receiver cable to the source, with the stationary phase point at zero. Looking at a larger section of the time axis for a different location located in the cloud of diffractions below the water-bottom in Figure 18b, there is some similarity between the traces in adjacent flip-flop shots, but very little similarity between the three different sail lines, which in part explains the incoherence of the diffracted pseudoprimaries in Figure 17.

Figure 18: A crossline pseudoprimary contribution gather. The location of the water-bottom reflection changes as a function of the crossline distance from the inline location of the pseudoprimaries to the six sources. [CR]

bill2/. 3dsynpseudogather



CONCLUSIONS AND FUTURE WORK

The results in both the synthetic and field data examples show that this method works at capturing large inline features in the data. The field data example shows in particular that the PEF is insensitive to issues with the squared spectrum, amplitude scale and phase, which avoids the need for preprocessing involving amplitude scaling and deconvolution prior to cross-correlation.

The utility of this method in 3D is limited by the poor spatial distribution of sources, which has also been the case with passive interferometric methods. However, future ac-

quisition methods may increase source density and aperture, which could in turn produce more useful pseud primaries in the crossline direction. Alternatively, this approach could be combined with moveout operators to produce data to add to the correlation.

REFERENCES

- Artman, B., 2007, Passive seismic imaging: PhD thesis, Stanford University.
- Berkhout, A. and D. Verschuur, 2003, Transformation of multiples into primary reflections: 73rd Ann. Internat. Mtg., 1925–1928, Soc. of Expl. Geophys.
- Berkhout, A. J., 1999, Multiple removal based on the feedback model: *The Leading Edge*, **18**, 127–131.
- Berkhout, A. J. and D. Verschuur, 2005, Transformation of multiples into primary reflections, 1039–1042. Soc. of Expl. Geophys.
- Berkhout, A. J. and D. J. Verschuur, 1994, Multiple technology: Part 2, migration of multiple reflections: 64th Ann. Internat. Mtg, 1497–1500, Soc. of Expl. Geophys.
- Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: *Geophysics*, **33**, 264–269.
- Cole, S. P., 1995, Passive seismic and drill-bit experiments using 2-D arrays: PhD thesis, Stanford University.
- Dragoset, B., 1999, A practical approach to surface multiple attenuation: *The Leading Edge*, **18**, 104–108.
- Dragoset, W. H. and Z. Jericevic, 1998, Some remarks on surface multiple attenuation: *Geophysics*, **63**, 772–789.
- Guitton, A. and G. Cambois, 1999, Multiple elimination using a pattern-recognition technique: *The Leading Edge*, **18**, 92–98.
- Hadidi, M. T., M. Sabih, D. E. Johnston, and C. Calderon-Macias, 1999, Mobil's results for the 1997 workshop on multiple attenuation: *The Leading Edge*, **18**, 100–103.
- Lamont, M. G., B. M. Hartley, and N. F. Uren, 1999, Multiple attenuation using the mmo and isr preconditioning transforms: *The Leading Edge*, **18**, 110–114.
- Liu, B. and M. Sacchi, 2001, Minimum weighted norm interpolation of seismic data with adaptive weights: 71st Ann. Internat. Mtg, 1921–1924, Soc. of Expl. Geophys.
- Lokshantov, D., 1999, Multiple suppression by data-consistent deconvolution: *The Leading Edge*, **18**, 115–119.
- Matson, K. H., D. Paschal, and A. B. Weglein, 1999, A comparison of three multiple-attenuation methods applied to a hard water-bottom data set: *The Leading Edge*, **18**, 120–126.
- Reiter, E. C., M. N. Toksoz, T. H. Keho, and G. M. Purdy, 1991, Imaging with deep-water multiples: *Geophysics*, **56**, 1081–1086.
- Riley, D. C. and J. F. Claerbout, 1976, 2-D multiple reflections: *Geophysics*, **41**, 592–620.
- Sacchi, M. D. and T. J. Ulrych, 1995, Model re-weighted least-squares Radon operators: 65th Ann. Internat. Mtg, 616–618, Soc. of Expl. Geophys.
- Schuster, G., 2001, Theory of Daylight/Interferometric Imaging - Tutorial: 63rd Mtg., Session: A-32, Eur. Assn. Geosci. Eng.
- Shan, G., 2003, Source-receiver migration of multiple reflections: 73rd Ann. Internat. Mtg., 1008–1011, Soc. of Expl. Geophys.
- Spitz, S., 1991, Seismic trace interpolation in the F-X domain: *Geophysics*, **56**, 785–794.
- Trad, D., T. J. Ulrych, and M. D. Sacchi, 2002, Accurate interpolation with high-resolution

- time-variant Radon transforms: *Geophysics*, **67**, 644–656.
- Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: *Geophysics*, **57**, 1166–1177.
- Verschuur, D. J. and R. J. Prein, 1999, Multiple removal results from delft university: *The Leading Edge*, **18**, 86–91.
- Wapenaar, K., 2003, Synthesis of an inhomogeneous medium from its acoustic transmission response: *Geophysics*, **68**, 1756–1759.
- Weglein, A. B., 1999, How can the inverse-scattering method really predict and subtract all multiples from a multidimensional earth with absolutely no subsurface information?: *The Leading Edge*, **18**, 132–136.
- Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, V87–V95.

An algorithm for interpolation using Ronen's pyramid

Jon Claerbout and Antoine Guitton

ABSTRACT

Shuki Ronen has shown that a dip spectrum in 2-D may be characterized by a 1-D Prediction Error Filter (PEF) after his "pyramid transform" where (t, x) -space is transformed to $(\omega, u = \omega x)$ -space. The transform to (ω, u) -space creates empty locations (missing data). Thus we have the question of unknown PEF along with missing data. Here we propose to find both simultaneously by iterative linear least squares.

INTRODUCTION

A Prediction Error Filter (PEF) is good at characterizing a 1-D spectrum. A short simple PEF will easily model a narrow bandwidth, a broad one, or a combination — whatever fits. Unfortunately, the spectrum characterized by a PEF is a temporal spectrum or a spatial spectrum, when in practice it is more generally the dip spectrum that we seek a good representation for.

Although the helix transform brings us the power of PEFs to 2-D and 3-D data it does not directly offer us exactly what often seek, a dip spectrum. Shuki Ronen's pyramid transform does. It brings PEF power from time and space axes to the dip axis.

First we review Ronen's principles noticing that raw data traces are moved to radial traces on a new grid. The 1-D PEF must be found on this new grid. But the new grid has empty spaces (missing data) between the known radial traces. Thus we approach the problem as one of nonlinear least squares which we propose to solve by iterative linear least squares.

FIRST PRINCIPLES

Let us review Ronen's pyramid transform (Hung et al., 2005). Fourier transform data from (t, x) to (ω, x) . Ronen then shrinks and stretches the space axis to transform from x to $u = \omega x$. All the information on a trace (constant x_0), in the (t, x) plane, lands on the line $u = \omega x_0$ in the $(\omega, u = \omega x)$ -plane. This is a line through the origin. Ordinary traces have become radial traces.

Look ahead to Figure 1 and you will see an example of the basic property of Ronen's space, the property we will prove next, namely: Whereas dips in (t, x) space are curves in (ω, x) space, they return to straight lines in (ω, u) space. Dips in physical (t, x) space become frequencies on Ronen's u -axis, the same u -space frequency for all ω . To track the three events in Figure 1 into Ronen's space, it's helpful to notice that each has a different ω spectrum.

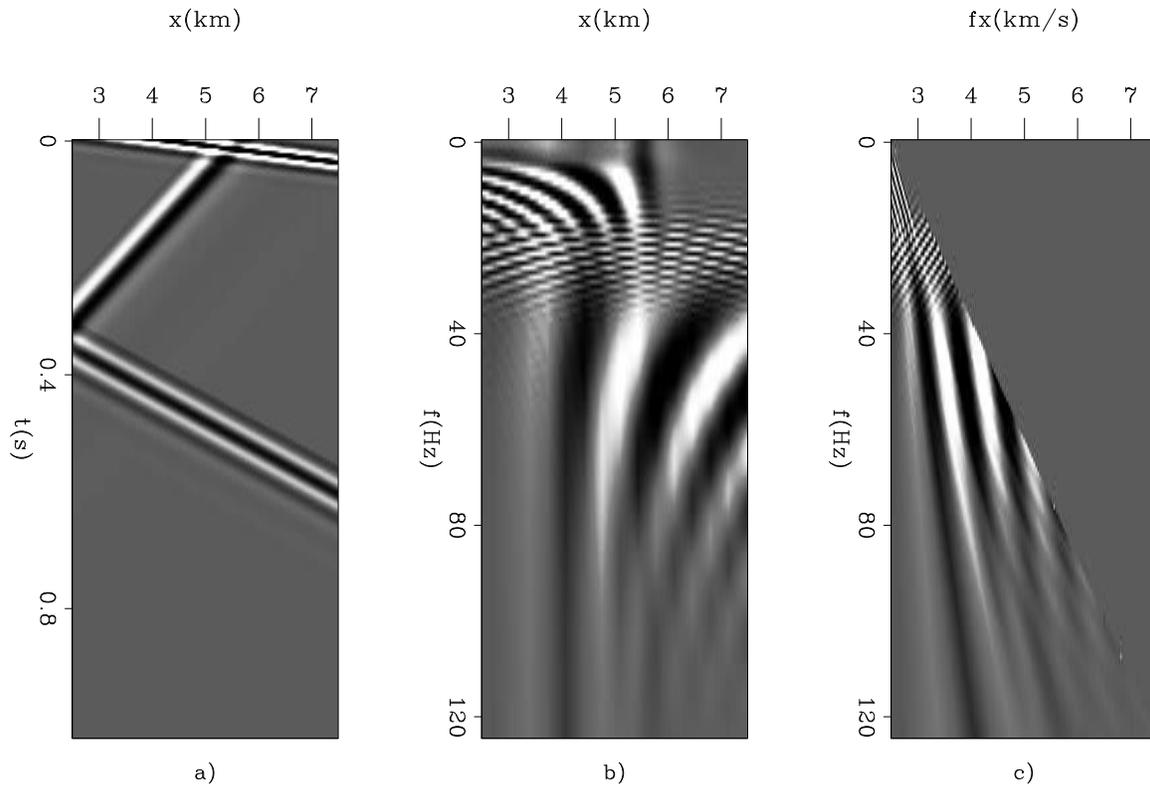


Figure 1: LEFT: Input data (t, x) with three events, each with different frequency content (and spatial extent). MIDDLE: Real part of (ω, x) -space of the panel on the left. RIGHT: Same in $(\omega, u = \omega x)$ -space, Ronen's pyramid space in which we again see straight lines as we did in (t, x) -space. Observe three frequencies on the u -axis, each corresponding to a different slope in the (t, x) -plane. To help identify which is which, notice each of the three events in the (t, x) -plane has a different ω spectrum. [ER] `jon1/. synth3`

To establish the basic property of pyramid space we consider a dipping signal in (t, x) -space, say $g(t, x) = g(t - px)$. Transform time t to frequency ω . Now we have a plane of complex valued information say $G(\omega, x)$. What is it in Ronen's (ω, u) space? Say $g(t, x = 0)$ fourier transforms to $G(\omega)$. Then $g(t - t_0)$ transforms to $G(\omega)e^{i\omega t_0}$. Take the time shift t_0 to be distance x times slope p . Then $g(t - xp)$ transforms to $G(\omega)e^{i\omega xp} = G(\omega)e^{ipx}$. Add some extra plane waves.

$$G(\omega, u) = G_1(\omega)e^{ip_1u} + G_2(\omega)e^{ip_2u} + \dots \quad (1)$$

Examine any ω . You see sinusoids on the u -axis. The frequencies of these sinusoids are (p_1, p_2, \dots) . We see those same frequencies for each ω . Thus the PEF on the u -axis informs us of a spectrum on the p -axis. To a seismologist, p is a dip axis though it would be more correct to call it an axis of stepout.

What happens in 3-D? Instead of (x, y) -axes we have $(u, v) = (\omega x, \omega y)$ axes. It is not immediately clear what applications will bring us to do in the (u, v) -plane. Some applications may call for a 2-D PEF there, while others call for two 1-D PEFs. What are we expecting to describe in (p_x, p_y) -space? Globally, there may be a bounding range in a slowness circle corresponding to a slowest material velocity. Locally things will look more like plane waves.

APPLICATION: AN ALGORITHM FOR INTERPOLATION

In seismology we often deal with instruments spaced more widely than they should be, more widely than they should be for typical data processing such as Fourier transform, more widely than is suitable for data display. Fundamentally something is lost, but that does not detract from our goal of regular spaced data on a dense enough mesh. Given regular data, finding a PEF is a linear problem. Given a PEF, interpolating data is a linear problem (demonstrated by many examples in Claerbout's free on-line book, "Image Estimation by Example"). In both cases we are minimizing the energy of filtered data. A more generalized approach minimizes energy in the filtered data where some unknowns are in the PEF while others are missing data values among the knowns. This minimization is nonlinear (because the PEF multiplies missing data).

The main difficulty of trying to utilize Ronen's pyramid in practice is the issue of bringing the x -space to the u -space. On first glance it seems to require interpolation. On trial, this interpolation seems to need to be done extremely carefully. An alternate approach, which we take here, is to sample the u -space very densely. This, of course, introduces many locations not touched directly by the data. We have traded the interpolation problem for a missing data problem, nonlinear because we must estimate this new missing data at the same time we estimate the PEF. We'd like to come up with a reliable pyramid method of interpolating aliased data that is devoid of low frequency information. An attractive feature is that the pyramid concept does not require the original data on a regular mesh in x -space. Our method will build a dense regular mesh in model u -space.

The problem is non-linear because of the product of unknowns, the PEF multiplying the missing data. Here we approach the problem by multistage linear least squares which can be iterated to solve the nonlinear problem. In any nonlinear problem the initial guess must be "near enough". Hopefully the proposed method will not demand unaliased low frequency information.

Imagine five unevenly spaced traces on the x -axis. The data space is defined as $d(\omega, x)$ at five known values of the coordinate x . Define a model space $m(\omega, u = \omega x)$ that is dense (many uniformly spaced points) on u -space. We are interested in fitting

$$\mathbf{0} \approx \mathbf{L}m - d$$

The operator \mathbf{L} linearly interpolates from the dense u -axis to the sparse x -space (which need not be regularly sampled on x). In the limit of an extremely dense u -space we might choose \mathbf{L} to be “extraction” basically “nearest-neighbor inverse binning”. A zeroth order model space is $m_0 = \mathbf{L}'d$. This model is simply dropping the several data traces into radial traces in pyramid space. Because u -space has very many points, $m_0(\omega, u)$ has many empty regions (triangularly shaped). Thus we use preconditioning. Let us precondition with the convolutional roughener $A_0 = (1, -1)$ on the u -axis. This particular A_0 leads to a solution $m(u)$ that linearly interpolates the given data. Such preconditioning should be good at large ω where the radial traces are far apart. Thus $p = \mathbf{A}_0 m$ or $m = \mathbf{A}_0^{-1} p$. To find p and m (independently for each ω) we iterate on the regression

$$\mathbf{0} \approx \mathbf{L}\mathbf{A}_0^{-1}p - d$$

The resulting m we will call m_1 (because we will eventually improve on it getting an m_2).

Next let us upgrade \mathbf{A}_0 . At each ω from the model space m_1 , make an operator \mathbf{M}_1 for convolution over the u axis. Simultaneously for all ω we find the regression for an upgraded PEF a (which is constant over ω).

$$\mathbf{0} \approx \mathbf{W}\mathbf{M}_1 a$$

Here \mathbf{W} is a diagonal weighting matrix defined next. As mentioned earlier, there are large empty spaces in the zeroth order model space $m_0 = \mathbf{L}'d$. Although our improved model space $m_1(\omega, u)$ has filled the holes with artificial data, we don't want to use regression equations except where the model space points directly to real data, namely where $m_0(\omega, u)$ is non-zero. Thus we define \mathbf{W} to be a diagonal matrix of ones and zeros, zeros where $m_0(\omega, u)$ is zero.

Although we are planning to iterate, we will never change \mathbf{W} . From solution of the regression above we have the vector a which we use to make the filter operator \mathbf{A}_1 . Use it in place of \mathbf{A}_0 in the regression for p above. Iterate to get an m that is improved over m_1 . Call it m_2 . Iterate.

CONCLUSIONS

We have an algorithm. We should code it and test it.

REFERENCES

Hung, B., C. Notfors, and S. Ronen, 2005, Seismic data reconstruction using the pyramid transform: 67th Meeting, EAGE, Expanded Abstracts, A039.

Data interpolation in pyramid domain

Xukai Shen

ABSTRACT

Pyramid domain is defined as a frequency-space domain with different spatial grids for different frequencies. Data interpolation in pyramid domain is preferable since for stationary events only one prediction error filter (PEF) is needed for estimating all offsets and frequencies. However, when it is necessary to estimate both missing data and the PEF, it becomes a nonlinear problem. By solving both iteratively, we can linearize the problem and frame it as a least-squares problem. Initial results show that the iterative method will recover the missing data and PEF quite well.

INTRODUCTION

Data interpolation is an important step in seismic data processing that can greatly affect the results of later processing such as multiple removal, migration and inversion. There are many ways to interpolate data, including PEF-based approaches (e.g., Spitz, 1991; Crawley, 2000) and Fourier transform-based approaches (e.g., Xu et al., 2005). The first paper interpolate data using prediction filters in the frequency-space domain while the second one does it by using prediction-error filters (PEFs) in the time-space domain. On the other hand, the third one uses Fourier transform on irregularly sampled data to perform interpolation. The first two belong to the class of methods which is PEF-based data interpolation. A PEF is a filter that predicts one data sample from n previous samples, where n is the length of the PEF. One important feature of a PEF is that when convolved with known data, it minimizes the convolution result in the least-square sense, thus making it a good representation of correlations between samples of known data. Previous work includes Claerbout (1999), Curry (2007) and Crawley (2000). However, one disadvantage is that if PEF estimation is done in frequency-space (f - x) domain, one PEF is needed for every frequency.

Pyramid domain is a resampled representation of ordinary f - x domain. Although it has the axes of frequency and space, the spatial sampling is different for different frequencies. This is attractive because we can use sparser sampling to adequately sample the data at lower frequencies, which makes uniform sampling for all frequencies unnecessary. Therefore in pyramid domain, coarser grid spacing is used for lower frequencies, while finer spacing is used for higher frequencies. This makes it possible to capture the character of all frequency components of stationary events with only one PEF, making data interpolation more efficient.

In this paper, I present a way to interpolate data in pyramid domain based on PEF estimation. The paper is organized as follows: I first show how to transform data into pyramid domain and describe the method for missing-data interpolation and PEF estimation. I then show synthetic data examples, and discuss possible future work.

METHODOLOGY

There are two important parts of the pyramid-based interpolation algorithm, the first one is forward and inverse pyramid transform. The more accurate these transforms are performed, the better the result we can get for missing data interpolation. The second step is data interpolation and PEF estimation in pyramid domain, which are done alternately in an iterative way.

Forward and inverse pyramid Transform

A pyramid domain is derived from f - x domain, where instead of a uniform spatial grid spacing for all frequencies, larger grid spacing is used for lower frequencies. As a result, low-frequency data components are represented by fewer data points in pyramid domain, while higher-frequency data components are represented by more data points. However the grid spacing is not allowed to be smaller than the given grid spacing in f - x , since our data does not contain information at finer scales. This means that even for high frequencies, the maximum spatial samples that can be used is equal to the number in original f - x data.

Significant work has been done on this topic, a good example of which can be found in Sun and Ronen (1996), who transform the data from t - x space to f - x space, where f denotes frequency; the spatial grid spacing is then calculated for each frequency f using the equation

$$\Delta x(f) = \max(\Delta x_0, v/f), \quad (1)$$

where Δx_0 is the uniform spatial grid spacing in the original f - x data and v is the velocity that controls the slope of the pyramid. The equation gives the maximum grid spacing for a certain frequency without aliasing stationary events in the t - x domain. Next, an inversion scheme is employed to transform data in f - x space to pyramid domain as follows.

$$\mathbf{Lm} - \mathbf{d} \approx \mathbf{0}, \quad (2)$$

where \mathbf{m} is the model to be solved, which is data in pyramid domain, \mathbf{d} is the known data in f - x space, and \mathbf{L} is the interpolation operator. In this paper, I use simple linear interpolation; however, a more accurate interpolation would improve the PEF estimation. Sun and Ronen (1996) suggests using 8-point sinc interpolation to produce a relatively accurate result. The inverse transform uses a similar equation to the one mentioned above.

$$\mathbf{d} = \mathbf{Lm}. \quad (3)$$

Instead of solving an inverse problem, it is now a forward modeling problem where \mathbf{m} is known and \mathbf{d} is unknown data in f - x space.

Missing data estimation and PEF estimation

When transforming data from the f - x domain to the pyramid domain, there will be missing data in the pyramid domain if there are missing data in the f - x domain. We must interpolate the missing data and estimate the PEF in pyramid domain. If they are done simultaneously, the problem becomes non-linear. Instead of directly solving this non-linear problem, we

alternately solve two linear problems at each iteration, which are PEF estimation and missing data interpolation.

For missing data estimation, we start with a known PEF \mathbf{A} , and try to solve the following least-squares problem (Claerbout, 1999):

$$\begin{aligned} \mathbf{W}(\mathbf{d}_{interp} - \mathbf{d}_{known}) &\approx \mathbf{0} \\ \epsilon \mathbf{K} \mathbf{A} \mathbf{d}_{interp} &\approx \mathbf{0}, \end{aligned} \quad (4)$$

where \mathbf{W} is a diagonal masking matrix that is 1 where data is known and 0 elsewhere, ϵ is a weight coefficient that reflects our confidence in the PEF, \mathbf{K} is a diagonal masking matrix that is 1 where the convolution can be used and 0 elsewhere.

For PEF estimation, I try to solve the following problem assuming training data are \mathbf{D}_{known} , and W is the same as explained above while \mathbf{a} is the unknown PEF (Claerbout, 1999):

$$\mathbf{K} \mathbf{D}_{known} \mathbf{a} \approx \mathbf{0}, \quad (5)$$

EXAMPLE

Here I show the result of the missing data interpolation using PEF estimated on a synthetic data set consisting of 3 plane waves of different frequency bands, locations and dips (Figure 1a). After transforming data in f - x domain (real part shown in Figure 1b) into pyramid domain (real part shown in Figure 1c), I selected a subset of the data, keeping only one trace out of every twenty in the radial direction (Figure 2a). This sub-sampling is done by first creating a missing data mask in f - x domain, then transform this mask into pyramid domain, then apply the mask onto known data in pyramid domain, thus creating missing data in pyramid domain. This is different from how it should be done normally, which is decimate data in f - x domain before go into pyramid domain. But it is a good start to understand how things work in pyramid domain. Notice that the interpolated data (Figure 2b) looks like the original data (Figure 1c). The different between interpolated data and original data at known locations (real part shown in Figure 2c) may be further minimized by employing more accurate interpolation scheme to go from f - x domain to pyramid domain.

FUTURE WORK

Currently a linear interpolation operator is used for the forward and inverse pyramid transform, and the nonlinear iteration have not yet been performed. A better understanding of how data behaves in pyramid domain will allow me to perform the nonlinear iterative method which estimates PEF and interpolates missing data at the same time. I find that linear interpolation operator introduces artifacts that affect PEF estimation, in future tests, I will try an 8-point sinc interpolation operator, which may yield a better result for PEF estimation and data interpolation. Adding a weighting function may also improve PEF estimation by enabling increased reliance on known data, and less on interpolated data. Also, I plan to test the algorithm on more realistic data set. Since in the Pyramid domain only one PEF is needed for all frequencies, interpolation should be faster than the f - x domain PEF estimation scheme described in Curry (2007).

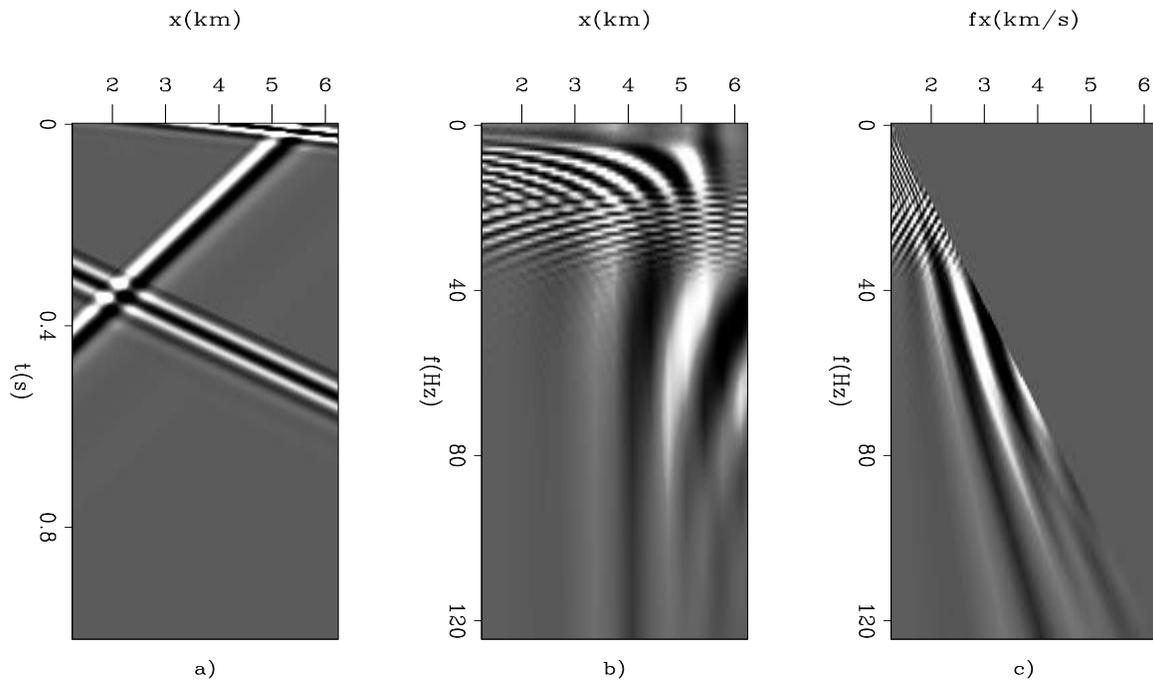


Figure 1: a) Input data, consisting of three plane waves with different spatial extents and frequency contents. b) real part of the input data in the f - x domain. c) real part of input data in pyramid domain. [ER] `xukai1/. synth3`

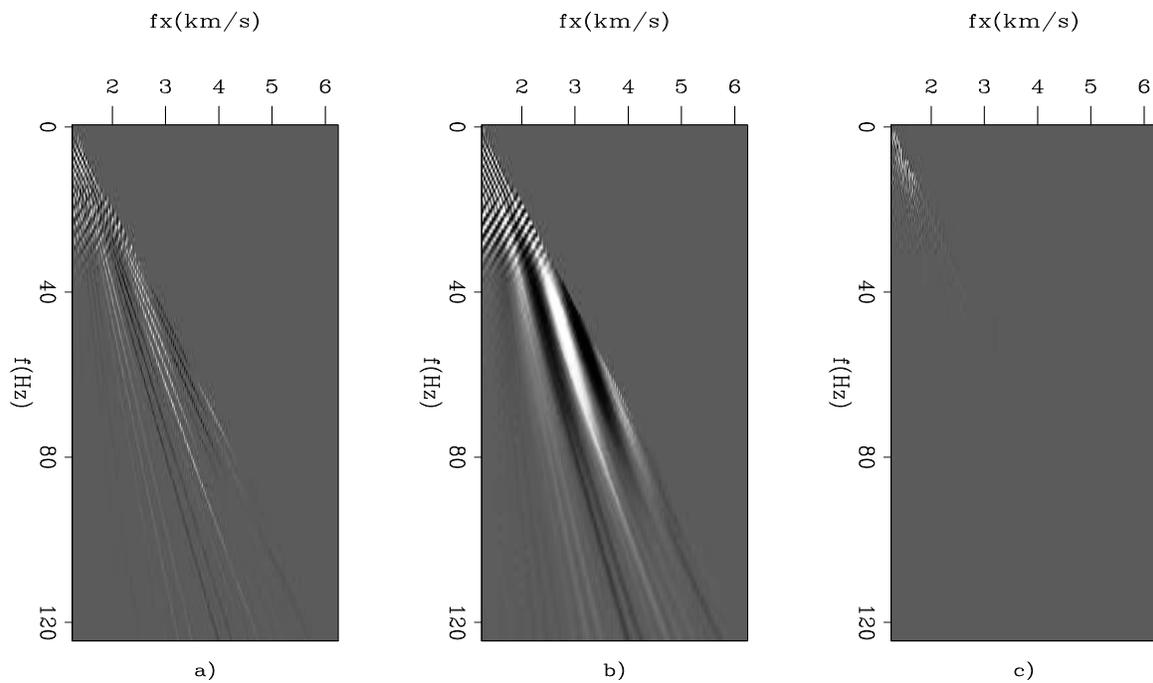


Figure 2: a) real part of known data in pyramid domain. b) real part of the interpolated data in pyramid domain. c) real part of weighted difference between original data and interpolated data in pyramid domain, clipped at 0.01 for display purpose, with a) and b) clipped at 0.2. [ER] `xukai1/. Interp3`

CONCLUSIONS

Pyramid domain is a very promising domain for missing data interpolation. The synthetic example demonstrates that with good initial PEF estimate, we can interpolate the missing data relatively accurately. With only one PEF needed for the whole data set, the efficiency of the algorithm is competitive.

ACKNOWLEDGMENTS

I would like to thank Jon for ideas about data interpolation in pyramid domain, and also thank Antoine and Bill for numerous debugging suggestions.

REFERENCES

- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project.
- Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: PhD thesis, Stanford University.
- Curry, W., 2007, f-x domain non-stationary interpolation in patches: SEP-Report, **131**.
- Spitz, S., 1991, Seismic trace interpolation in the F-X domain: *Geophysics*, **56**, 785–794.
- Sun, Y. and S. Ronen, 1996, The pyramid transform and its application to signal/noise separation: SEP-Report, **93**, 161–176.
- Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: *Geophysics*, **70**, V87–V95.

Biot-Gassmann analysis of partial and patchy saturated reservoirs for both reflected and transmitted seismic waves

James G. Berryman

ABSTRACT

This paper is a tutorial on the relationships between Biot-Gassmann theory for AVO-AVA analysis of reflection seismic amplitudes and also for the related analysis of propagation speeds for transmitted seismic waves in applications such as VSP or crosswell seismic. The main observation of special note is that the so-called “fluid-line” in AVO analysis appears to be related to patchy saturation of fluids when the Biot-Gassmann analysis of laboratory data is considered.

INTRODUCTION

In this paper, I present a short analysis of and comparison between the well-known Amplitude Versus Offset (AVO) approach (Ostrander, 1984; Shuey, 1985; Allen and Peddy, 1993; Castagna, 1993; Mavko et al., 1998) — as well as for Amplitude Versus Angle [for AVA, see Brandsberg-Dahl et al. (2003) and Contreras et al. (2006)] — for identifying fluid content in reservoirs from seismic reflection data and the predictions of Gassmann’s equations for AVO reflectivity analysis. In addition other distinct types of seismic wave analysis based on wave velocities alone will be considered in the same context (Krief et al., 1990; Goodway et al., 1997; Berryman et al., 2002). Some of these cases of most interest may not involve any reflections, only transmission of the waves through the reservoir of interest. In such cases, the wave speed/velocity analysis presented can in fact be used in the absence of reflectivity data, for example, in well-logging applications, VSP, or crosswell seismic tomography.

AVO analysis is a technique used to determine various physical properties of reservoirs, including density, porosity, bed thickness, velocity, and (most importantly) fluid content of rocks. This type of analysis tends to be more successful in young, poorly consolidated rocks, such as those in the Gulf of Mexico. Older, better consolidated, or well-cemented sediments from continental regions are not such good candidates for AVO studies.

One common failure of AVO comes from the inability to distinguish a gas-filled reservoir from one having only partial gas saturation (also known in this application as “fizz water”). As the Biot-Gassmann theory shows, shear wave energy of mode-converted waves at reservoir interfaces permits some levels of gas saturation to be distinguished. And one of the better ways of doing this is to use time-lapse analysis of seismic reflection data.

The analysis presented here will emphasize relations to the Biot-Gassmann theory (Biot, 1956, 1962; Gassmann, 1951) of fluid-saturated porous media. Although clearly important for explaining “bright spots” due to gas over oil, as well as many other issues in AVO, the Biot-Gassmann theory has seldom been treated carefully in this context due to the presence of so many other conflicting practical issues with the data analysis. I will nevertheless work

through the predictions of Biot-Gassmann here, so our knowledge of this part of the science is clearly laid out for those who do find situations where it may be appropriate to take advantage of it. The Biot-Gassmann theory, when extended to the cases of partial and patchy saturation, can help to eliminate some of these ambiguities.

AVO ANALYSIS AND ALTERNATIVES FOR SEISMIC REFLECTIONS

AVO analysis in isotropic layered systems uses the reflectivity moveout for compressional wave scattering (Koefoed, 1955, 1962; Aki and Richards, 1980; Ostrander, 1984; Shuey, 1985; Foster et al., 1997)

$$R(\theta) = A + B \sin^2 \theta, \quad (1)$$

where R is the reflectivity, θ is the angle of incidence measured from the vertical so that $\theta = 0^\circ$ corresponds to vertical incidence, and $\theta = \pi/2$ ($= 90^\circ$) is horizontal incidence. The coefficients A and B are called respectively the intercept and slope, because of their significance at $\theta = 0^\circ$ (for the intercept), and $\theta \neq 0^\circ$ (for the slope). Layering of physically distinct types of earth materials is implicitly required by this method, since the wave reflectivity vanishes if there is no contrast in the pertinent wave-related physical properties at such interfaces. Wave-related properties include: the elastic constants K (bulk modulus) and G (shear modulus), and also the inertial mass density ρ . A list of important physical properties could also include wave attenuation, but I will not discuss that issue in this paper.

Coefficients A and B depend on the elastic properties of the system through the wave speeds v_p (compressional) and v_s (shear) determined by the well-known relations

$$v_p^2 = (K + 4G/3)/\rho, \quad (2)$$

and

$$v_s^2 = G/\rho. \quad (3)$$

The bulk modulus and shear modulus are related to the Lamé elastic parameters λ and μ for isotropic media (Love, 1944; Landau and Lifshitz, 1986) by $K = \lambda + 2\mu/3$ and $G = \mu$. For long-wavelength analysis, the density $\rho = M/\Omega$ includes all the mass M (of solid, liquid, and/or gas) present in the volume Ω .

Following Aki and Richards (1980) [but also see Ewing et al. (1957), Achenbach (1973), Castagna (1993), Foster et al. (1997), and Mavko et al. (1998)], we have equations for the reflectivity coefficients which are:

$$A = \frac{\Delta v_p}{2v_p} + \frac{\Delta \rho}{2\rho} = \frac{\Delta(v_p^2)}{4v_p^2} + \frac{\Delta \rho}{2\rho} \quad (4)$$

and

$$B = \frac{\Delta v_p}{2v_p} - 4 \frac{v_s^2}{v_p^2} \left(\frac{\Delta v_s}{v_s} + \frac{\Delta \rho}{2\rho} \right) = \frac{\Delta(v_p^2)}{4v_p^2} - 2 \frac{v_s^2}{v_p^2} \left[\frac{\Delta(v_s^2)}{v_s^2} + \frac{\Delta \rho}{\rho} \right]. \quad (5)$$

The meaning of the Δ in these formulas is clear, since I am referring to the differences in these quantities as the wave passes from one layer to the next. Thus, for example, Δv_p is the change in the compressional wave speed v_p across the interface. Since — for small

changes (consistent with the analysis of Aki and Richards) — the reflectivities should be the same (except for the difference in sign due to the sign of the Δ) whether the waves are up-going or down-going, it is clear that the quantities not differenced in these formulas must be averaged, *i.e.*, take half the sum of these quantities on each side of an interface. If these differences across the interfaces are very small, then the layer averaging step can be ignored altogether. Linearized parameter estimation techniques (Ursin et al., 1996) are also consistent with these assumptions together with the additional caveat that reasonably good estimates of the background parameters must also be available from other means. The second version of each of these two formulas in (4) and (5) is presented to emphasize the fact that these quantities are really equivalent to logarithmic derivatives, and as such can be written in many different ways. The advantage of the second choice in each case is that, by squaring the velocities, we eliminate the square roots of the bulk and shear moduli, as well as those of the densities. These square roots would naturally fall away as we compute the relationships between coefficients A , B , and physical constants K , ρ , G . But it seems a bit easier to scan the math and, thereby, see where terms originate by introducing this intermediate step.

If the only change across the boundary of interest occurs in the nature — *i.e.*, the physical properties — of the pore fluid, then for isotropic materials I can take advantage of Gassmann's equations to note that the shear modulus G does not depend on the fluid bulk modulus K_f , and, therefore, G does not change under these circumstances within mixed-fluid-saturated reservoirs. Then, using the definitions of the wave speeds for isotropic elastic/poroelastic media, I have:

$$A = \frac{\Delta K}{4(K + 4G/3)} + \frac{\Delta \rho}{4\rho} \quad (6)$$

and

$$B = \frac{\Delta K}{4(K + 4G/3)} - \frac{\Delta \rho}{4\rho}. \quad (7)$$

By taking linear combinations of these two expressions, we can separate out the dependencies on the changes in bulk modulus K and the density ρ , giving respectively

$$\frac{\Delta K}{2(K + 4G/3)} = A + B \quad (8)$$

and

$$\frac{\Delta \rho}{2\rho} = A - B. \quad (9)$$

Thus, these relations represent a very clean decoupling of the changes in the reflectivity coefficients *due to the fluid effects alone* (since, by present assumption, the fluid effects are the only variables present) in terms of the elastic and inertial physical parameters K and ρ . And, at least for isotropic media, these are the two parameters that determine relationships between fluid properties and wave speeds.

DENSITY AND BIOT-GASSMANN RELATIONS FOR THE BULK MODULUS

Of the two parameters K and ρ , the easier physical parameter to analyze in the foregoing text is certainly the inertial density, which is given explicitly by

$$\rho = (1 - \phi)\rho_s + \phi[S\rho_l + (1 - S)\rho_g], \quad (10)$$

where ϕ is the porosity, ρ_s is the solid density of the rock material (with no porosity), ρ_l is the density of any liquid present (assuming only one such is present), and ρ_g is the density of any gas present (again assuming only one such is present). The saturation $S = S_l$ refers to the liquid saturation, meaning that the parameter $0 \leq S_l \leq 1$ is a measure of the fraction of the pore space that is occupied by liquid. Then, the gas saturation is $S_g = 1 - S$, which is just the fraction of the pore space left to be occupied by the gas. This way of accounting for the density contribution assumes there are no dynamic effects of density in the seismic/acoustic/ultrasonic frequency domain of interest. For Biot theory (Biot, 1956, 1962), this assumption is an oversimplification, since there can be motion of the fluids, creating damping effects that make the inertial density effectively a complex number due to out-of-phase motion between the pore fluid and the surrounding solid matrix materials. I am typically assuming here that the frequencies of interest are those of field/seismic experiments, so they are usually less than about 20 kHz for well-logging, and more likely to be the range 10 to 100 Hz for reflection seismic work. Cross-well tomography can cover the frequencies between these two extremes. But all of these frequencies, up to about 1 kHz for sure (Berryman et al., 2002), are low enough that the scalar constant density assumption made here is usually appropriate. The biggest concerns — if any — will come into play only at the very highest frequencies, and then for laboratory experiments, especially in the ultrasound range. (Laboratory data analyzed later is at higher frequencies, but — as is demonstrated — not so high as to invalidate the present discussion.)

Again, for low enough frequencies and isotropic solid matrix, it is normally adequate to treat the shear modulus as independent of both frequency and fluid-saturation. So we may assume that $G = \mu = \text{constant}$ [following the isotropic results of Gassmann (1951)]. But this assumption can also go wrong at high enough frequencies, or in the presence of anisotropic elastic behavior (Pride *et al.*, 2004; Berryman, 2005). We will ignore both of these effects in the following *simplified* analysis.

One alternative simplification of Equation (5) follows from this constancy of G in isotropic poroelasticity. In particular, I find that

$$\left[\frac{\Delta(v_s^2)}{v_s^2} + \frac{\Delta\rho}{\rho} \right] \equiv 0, \quad (11)$$

since G is independent of the physical mechanisms that (I have assumed) are varying, so the G dependence of this expression disappears, so that the remaining contribution is just:

$$\left[\frac{\Delta(\rho^{-1})}{\rho^{-1}} + \frac{\Delta\rho}{\rho} \right] = -\Delta \ln \rho + \Delta \ln \rho \equiv 0. \quad (12)$$

So it vanishes identically. Thus, for the Biot-Gassmann examples considered here, the slope B is given simply by

$$B = \frac{\Delta(v_p^2)}{4v_p^2}. \quad (13)$$

Thus, only spatial variations in v_p determine B , while variations in both v_p and ρ play roles in determining the value of $A = B + \Delta\rho/2\rho$ for this Biot-Gassmann analysis.

Another important observation based on the result (B-2) is the fact that the logarithmic derivative of the density can be replaced everywhere by the logarithmic derivative of v_s^2 , using the relation:

$$\frac{\Delta\rho}{\rho} \equiv -\frac{\Delta(v_s^2)}{v_s^2}, \quad (14)$$

which is always true for the situations under consideration here. Thus, in these circumstances, I find that

$$A = \frac{\Delta(v_p^2)}{4v_p^2} - \frac{\Delta(v_s^2)}{2v_s^2}, \quad (15)$$

and

$$A - B = -\frac{\Delta(v_s^2)}{2v_s^2}. \quad (16)$$

Seismic impedance analysis

Seismic impedance analysis (Goodway et al., 1997) is based on ideas similar to those already presented, but the effort is concentrated now on the quantities $I_p \equiv \rho v_p$ and $I_s \equiv \rho v_s$, which are the seismic impedances that determine the values of the reflection and transmission coefficients for normal incidence at interfaces between layers. Clearly, we can reanalyze everything we have presented so far in terms of these impedances, since exactly the same physical quantities are involved: ρ , v_p , and v_s . To the order of approximation we are considering now, the coefficient A is seen from (4) to be given by

$$A = \frac{\Delta I_p}{I_p}. \quad (17)$$

Then the intercept A is just the logarithmic derivative of I_p (which is basically how and why I_p was defined in the first place). Obtaining a similar expression for the logarithmic derivative of I_s , requires consideration of similar measurements for the reflections of shear waves at boundaries. On the other hand, for the Biot-Gassmann analysis, I have $I_s = \sqrt{\rho G}$, with G being constant across these boundaries, and therefore

$$\frac{\Delta I_s}{I_s} = \frac{\Delta \rho}{2\rho}, \quad (18)$$

so, from (9), I have

$$B = A - \frac{\Delta \rho}{2\rho} = \frac{\Delta I_p}{I_p} - \frac{\Delta I_s}{I_s}. \quad (19)$$

Thus, B can also be interpreted as the difference in the two logarithmic derivatives of I_p and I_s .

If direct measurements of these two impedances are available (perhaps from well-logs of density and wave speeds, for example), then other combinations of these quantities, may be considered and, in particular, I have $\rho\lambda = I_p^2 - 2I_s^2$ — where $\lambda = K - 2G/3$ is one Lamé constant. The advantage of studying the quantity $\rho\lambda$ is that this calculation has eliminated the part of the problem that is insensitive to fluids, namely the shear modulus G , which has disappeared altogether from this relationship. Thus, when various authors (Goodway et al., 1997; Contreras et al., 2006) compare plots of v_p , I_p , and $\rho\lambda$, it is typically found that the $\rho\lambda$ plots are the ones that are the most sensitive to fluid/liquid presence.

Well-mixed gas-liquid mixtures: partial (liquid) saturation

Finally, I need to consider the bulk modulus K for the fluid-filled system, and this modulus requires more discussion. First, for the isotropic case with a single fluid F , we have

Gassmann's equation (Gassmann, 1951; Biot and Willis, 1957; Berryman, 1999):

$$K_u = K_d + \frac{\alpha^2}{\frac{\alpha-\phi}{K_s} + \frac{\phi}{K_F}}, \quad (20)$$

where K_s is the (assumed constant/uniform) bulk modulus of the solid material, K_d is the drained modulus (for a porous system filled with air or other very compliant gas, this modulus is essentially the same as if all fluid is drained from the pores), K_u is the undrained modulus (either the fluid is trapped by the boundary conditions, or the wave frequencies are so high that the system cannot relax significantly during one cycle), $\alpha = 1 - K_d/K_s$ is the Biot and Willis (1957) coefficient (also sometimes called the effective-stress coefficient), and finally K_F is the bulk modulus of the assumed uniform fluid. The modulus K_F can also be treated as the modulus of a uniformly mixed fluid, such as a liquid-liquid mixture, or a liquid-gas mixture. There is however an assumption of immiscibility implicit in the resulting formula for K_F , which — for liquid and gas mixtures — is given by

$$\frac{1}{K_F} = \frac{S}{K_l} + \frac{1-S}{K_g}, \quad (21)$$

where K_l is the bulk modulus of the liquid constituent, and K_g is the bulk modulus of the gaseous constituent. As before, $S \equiv S_l$ is the liquid saturation level. Equation (21) is the volume weighted harmonic mean of the moduli of the two constituents [also known as the Reuss (1929) average], and it is the exact result for quasi-static deformation of a well-mixed liquid-gas mixture.

Next, the most important fact to notice about (21) is that $K_g \ll K_l$ typically holds for liquids and gases. So, (21) implies that $K_F \simeq K_g/(1-S)$, unless $S \simeq 1$, in which case the formula reverts to the near equality $K_f \simeq K_l$. Thus, there is usually a broad range of values of saturation $S \ll 1$ for which K_F behaves like K_g , and a narrow range of S 's, on the order of $0.95 \leq S \leq 1$, in which the liquid behavior becomes dominant. (See the discussion later on of Figure 1.)

Patchy saturation

Now suppose that the fluids are not well-mixed, but actually very poorly mixed, so that each of the two fluids essentially occupies its own private compartments of the rock. Then, the correct formula for the undrained modulus is *not* given by (20). Instead we need to consider these two quite differently occupied compartments separately. Each compartment may still obey Gassmann's equation, but with $K_F = K_l$ in one compartment, and $K_F = K_g$ in the other compartment. So we have two distinct results, both valid simultaneously but for different spatial regions determined by the fluid (g, l) saturant:

$$K_u^{(l)} \equiv K_d + \frac{\alpha^2}{\frac{\alpha-\phi}{K_s} + \frac{\phi}{K_l}}, \quad (22)$$

and

$$K_u^{(g)} \equiv K_d + \frac{\alpha^2}{\frac{\alpha-\phi}{K_s} + \frac{\phi}{K_g}}, \quad (23)$$

assuming that all the other properties except for the saturation levels are the same in the two compartments.

Next consider the observation that, since the shear modulus $G = \mu$ does not depend in any way on fluids (so $G_g = G_l = G$) if the porous material is locally isotropic, then we can take advantage of the rigorous and exact results for anisotropy due to layering. The known results on anisotropy due to layering by Postma (1955) and Backus (1962), and later for the bulk modulus of general heterogeneous materials (when composed of isotropic materials) by Hill (1963) all produce the same general formula:

$$K_u^* = \left[\sum_{f=g,l} \frac{S_f}{K_u^{(f)} + 4G/3} \right]^{-1} - 4G/3, \quad (24)$$

where the $*$ indicates the (overall) effective bulk modulus of the system as a whole, and again $S_g = 1 - S_l$. We should emphasize that the condition of constant shear modulus is not at all a trivial one, and it may not in fact be closely satisfied in many cases in practice. Furthermore, a more general pair of relations for bounds on overall bulk modulus is known, namely the Hashin-Shtrikman bounds Hashin and Shtrikman (1963). These bounding relations of Hashin and Shtrikman have exactly the same functional form as (24), although with different choices for the shear modulus parameter G (since G is not unique in the heterogeneous system) — and also possibly the undrained bulk modulus parameters $K_u^{(f)}$ may differ as well — since these bounds are usually written for two-component media having $G = G_1$ or G_2 . The logic of bringing in the Hashin-Shtrikman bounds into the discussion at this point is somewhat doubtful, however, since my main line of argument concerns a constant porous background medium with only variable fluid content. So I will not pursue this alternative point of view here. But it should be emphasized that use of these Hashin-Shtrikman bounds in the present context has considerable potential to produce misleading results, especially when the system has some local anisotropy present. [See Avellaneda and Milton (1989) and Milton (2002).] Nevertheless, if the local (when drained) porous medium is sufficiently isotropic and uniform, then these conditions, that are required for (24) to be valid, are then satisfied automatically. This somewhat limited result is then sufficient for the purposes of the present discussion.

Patchy saturation is clearly very pertinent to fluid distributions in the earth. It is also true that it must have frequently happened that laboratory data on partially saturated systems have been analyzed (in the absence of other, more detailed, information about fluid arrangement in the pore space) assuming that the spatial distribution of fluids did not matter, and then applying the Biot-Gassmann results for homogeneous saturation. One recent example of a recalculation for a system that was surely patchily saturated, but analyzed as if it were homogeneously saturated was the data of Murphy (1982), where he inappropriately analyzed his drainage data on porous cylinders using the Biot-Gassmann theory for homogeneous partial saturation. The observed results did in fact seemingly agree quite well with the theory in this case. But more recently Berryman and Pride (2005) have reanalyzed Murphy's data assuming instead that the drainage process created a dry annulus surrounding a fully wet core in the porous cylinder. This fluid distribution would be consistent with a drying process that starts at the outside cylinder surface, and gradually works its way into the center of the cylinder. Berryman and Pride (2005) show that the new analysis of these data actually agrees significantly better with the laboratory data than

did the original homogeneous analysis of these same data by Murphy (1982). The patchy analysis is more difficult technically than was the earlier analysis, and that is one reason that it took more than 20 years for this reanalysis to appear.

Saturation proxy

If the inertial density of the partially or patchily saturated medium is given by $\rho = (1 - \phi)\rho_s + \phi(S_l\rho_l + S_g\rho_g)$, where again $S_g = 1 - S_l = 1 - S$, then the inverse of the shear velocity squared can be used as a liquid saturation proxy via:

$$S_{proxy} \equiv \frac{(1/V_s^2)|_S - (1/V_s^2)|_{S=0}}{(1/V_s^2)|_{S=1} - (1/V_s^2)|_{S=0}} = \frac{\rho|_S/G - \rho|_{S=0}/G}{\rho|_{S=1}/G - \rho|_{S=0}/G} = \frac{\rho|_S - \rho|_{S=0}}{\rho|_{S=1} - \rho|_{S=0}} = S_l. \quad (25)$$

The point of showing the various steps in (25) is that, since the shear modulus G is independent of the fluids within isotropic poroelastic media [but see Mavko and Jizba (1991) and Berryman (2005)], G cancels out in the next to last step — thus producing the desired result. This approach requires the auxiliary data $V_s|_{S=0}$ and $V_s|_{S=1}$, which can presumably be estimated either from data in other differently saturated spatial locations in the same reservoir, or from time-lapse data at the same location in the reservoir, as the liquid saturation level declines or increases. This fact can be useful (Berryman et al., 2002) when saturation data are not available, but will not be considered further here.

EXAMPLES AND DISCUSSION

AVO analysis typically concentrates on observations of the so-called “fluid line,” which on a plot of B versus A is the line defined by $B = -A$, from the upper left through the origin at $(A, B) = (0, 0)$ down to the lower right. The nominal “hydrocarbon line” is the vertical line for all B values at fixed $A = 0$, and other points in a vertical band in the close vicinity of this line.

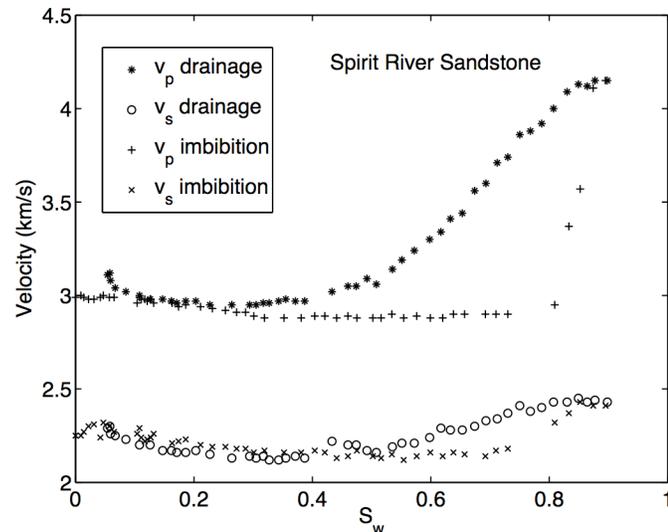
These “rules” are typically gleaned from field experience, rather than from tight connections to the physical principles embodied in the Biot-Gassmann theory of porous media containing fluids. Part of the problem regarding the physics versus the field experience is lack of sufficient knowledge — both of the porous earth medium in situ, and also of the various pertinent details of the properties and spatial arrangement of the fluids themselves. So it is not unusual to see discussions of these AVO methods without any reference at all to the Biot-Gassmann literature for seismic waves in fluid-saturated or partially (possibly patchily) saturated porous media.

Figures 1-9 all make use of laboratory data by Knight and Nolen-Hoeksema (1990) on Spirit River sandstone. This sandstone has porosity $\phi = 0.052$, and fluid permeability of $1\mu\text{D}$. Measurements were made with piezoelectric transducers having resonant frequencies of 1 MHz for the v_p measurements, and 600 kHz for the v_s measurements. (These frequencies are high compared to the earlier discussion, but the data displayed indicate the prior analysis is valid nevertheless.) Maximum saturation attained was $S_w = S_l = 0.9$. We can think of these results as time-lapse data with imbibition carrying the liquid saturation from zero to 0.9, and with drainage carrying the saturation from 0.9 back to zero. Clearly a type of hysteresis is observed in these data, as the curves do not lie on top of each

other. Rather, what is seen is surely (based on interpretations of the Biot-Gassmann theory for homogeneous and patchy saturation) homogeneous saturation along the imbibition curve, and patchy saturation along the the drainage curve (Berryman et al., 2002). For $S_w = S_l < \simeq 0.3$, the two curves essentially coincide. The precise location of this coincidence point is hard to predict as it must depend in detail on the connectedness of the pore space, and so is typically not known unless other methods such as x-ray imaging have been implemented to determine this normally missing information.

Figure 1: Spirit River sandstone data of Knight and Nolen-Hoeksema (1990) for both drainage and imbibition methods of changing the liquid saturation levels. Note that imbibition is a time-lapse experiment with S_w increasing from zero to its maximum value. Similarly, drainage is a time-lapse experiment with S_w decreasing from its maximum value to zero. Thus, the compressional wave data show explicitly effects due to the fact that the fluid saturation is patchy for drainage, but nearly homogeneous for imbibition. [NR]

jim1/. SpiritRiverAll



Color is used in Figures 2 through 9 to highlight those parts of the two data sets where the most rapid changes are occurring in v_p . For Figures 2-5, these values are from approximately $S_w = 0.8$ up to $S_w = 0.9$, and the points are shown in blue. For Figures 6-9, these values are from approximately $S_w = 0.4$ up to $S_w = 0.9$, and these points are shown in red.

Previously mentioned observations from AVO analysis tend to emphasize that the so-called “fluid line” ($B = -A$) and the hydrocarbon line ($A = 0$) are actually *not* prominent in the observed behavior in Figures 2-5 for homogeneous saturation. Thus, these imbibition data do not seem to show the usual behavior associated with AVO plots of field data. This fact may indicate instead that field data typically display heterogeneous — or patchy saturation — characteristics.

In Figures 6-9, important differences from Figures 2-5 are observed, including the fact that, whereas the homogeneous saturation results for Figure 2-5 do not display the anticipated results for data lying along the so-called “fluid line” at $B = -A$, it is nevertheless easy to see that the patchy saturation data (*i.e.*, the red data points in Figures 6-9 for $S_w = 0.4$ up to $S_w = 0.9$) clearly do fall along this $B = -A$ line in Figure 6, for these Spirit River Sandstone ultrasonic data. This observation suggests that the empirically observed “fluid line” may actually be a mixed or patchy saturation line — since viscous hydrocarbons and different densities of both oil and gas are commonly found in situ. Since hydrocarbons were not used as the fluids in these laboratory experiments, I cannot say more about the lack of any hydrocarbon line (*i.e.*, $A = 0$) behavior in either of these data plots.

Figure 2: The next four figures show a comparison of four ways to view wave speed data and/or wave speed change data (via measured reflection coefficients). Laboratory data of Knight and Nolen-Hoeksema (1990) on Spirit Rive sandstone is used as the example, because both imbibition (quasi-homogeneous saturation) and drainage (patchy saturation) examples are available. Figure 2 show the usual display for AVO data of intercept A versus slope B . [NR]

jim1/. KBAvBSRSIM

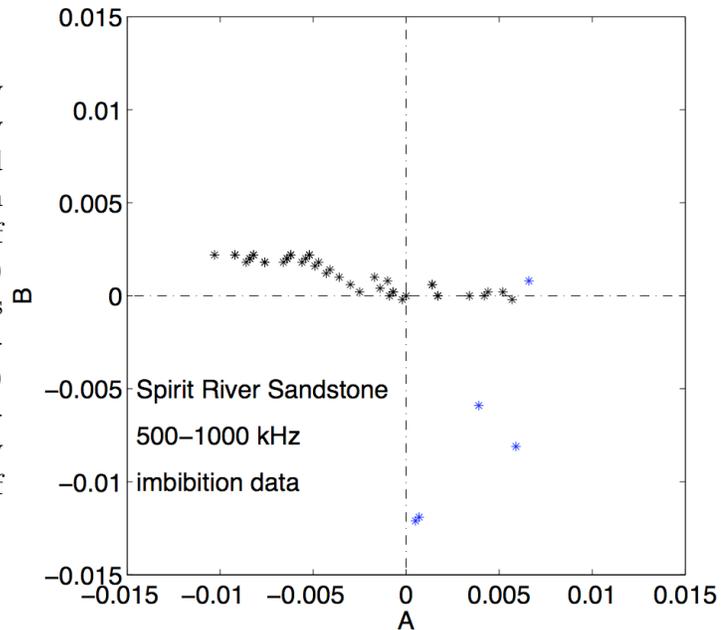


Figure 3: This figure shows how these same data from Figure 1 may be displayed by taking different combinations, namely $A - B$ and $A + B$. This approach is basically a rotation of the usual display in Figure 2, but the boundaries also expand by a factor of two. [NR]

jim1/. KBAmBvApBSRSIM

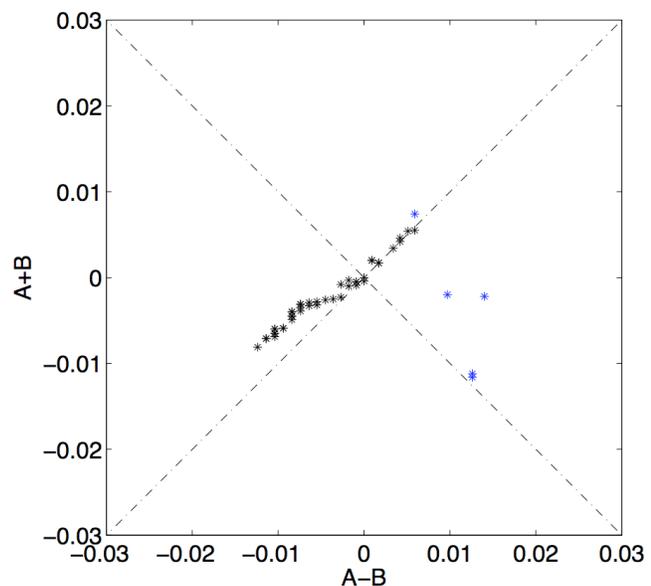


Figure 4: This third display of the imbibition data emphasizes $\rho\lambda = I_p^2 - 2I_s^2$ versus saturation, although it is actually $\rho\lambda/\mu^2$ that can be computed most easily from available velocity data. This is a display that could also be obtained using well-log data (ρ, v_p, v_s) . [NR]

jim1/. Brholamvmu2SRSIM

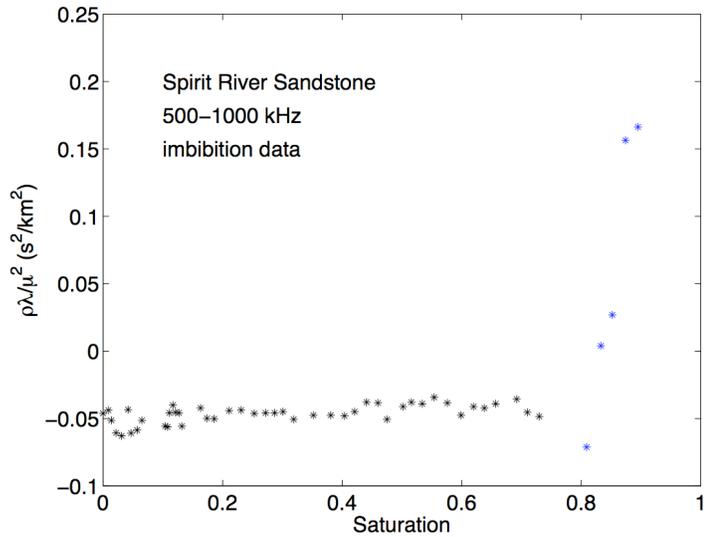


Figure 5: This fourth and last display of the imbibition data is very similar to Figure 4, but does not require independent information about the density ρ , so it is also pertinent to both VSP and crosswell seismic data. Note that in Figures 2-5, these imbibition data do not seem to show the usual behavior associated with AVO plots of field data, which may indicate that the patchy saturation model is the more appropriate one to use when analyzing field data. [NR]

jim1/. BlamvmuSRSIM

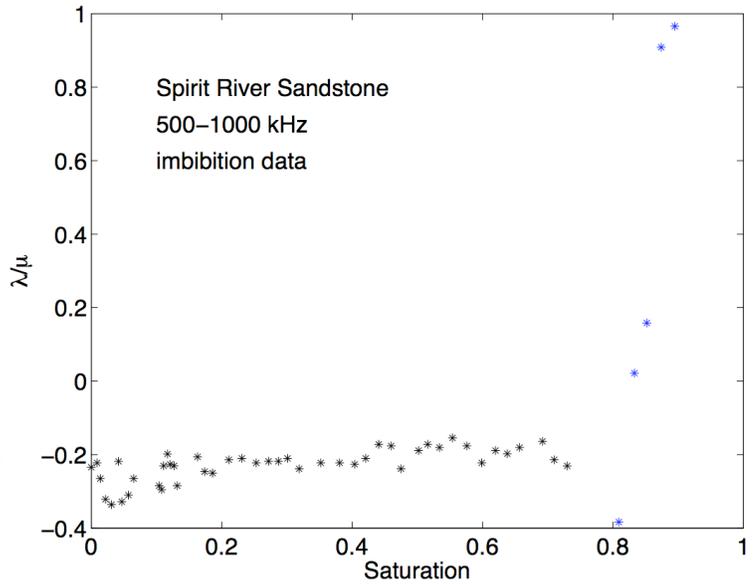


Figure 6: Like Figure 2, but for drainage data and therefore patchy saturation. Note that the most interesting parts of the data set (having more liquid present) do in fact fall along the $B = -A$ “fluid-line” in this Figure, as is also often seen empirically in field data. [NR]

jim1/. KRAvBSRSDD

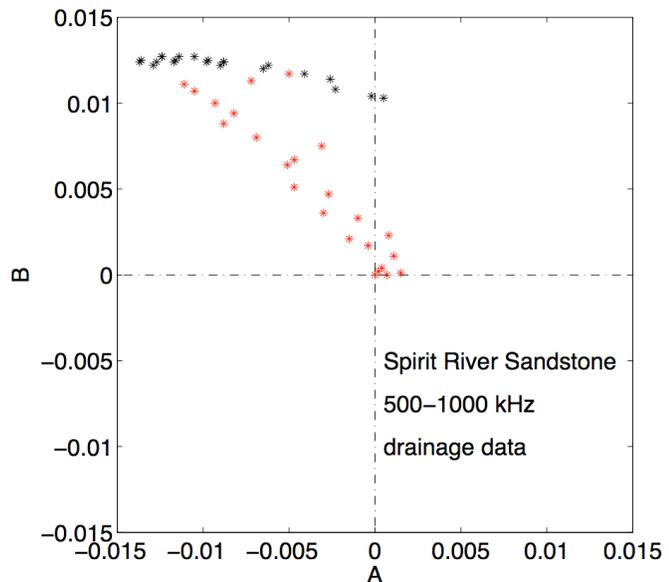


Figure 7: Like Figure 3, but for drainage data and patchy saturation.
 [NR] jim1/. KRAmBvApBSRSDD

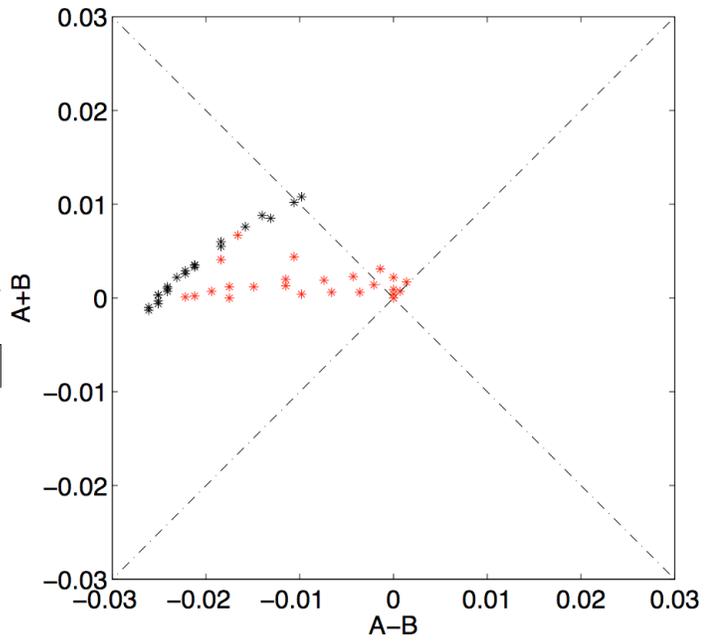


Figure 8: Like Figure 4, but for drainage data and patchy saturation.
 [NR] jim1/. KRrholamvmu2SRSDD

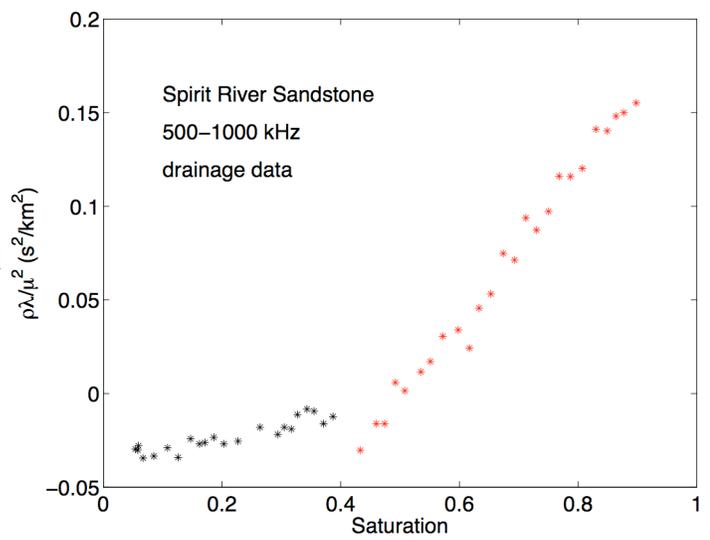
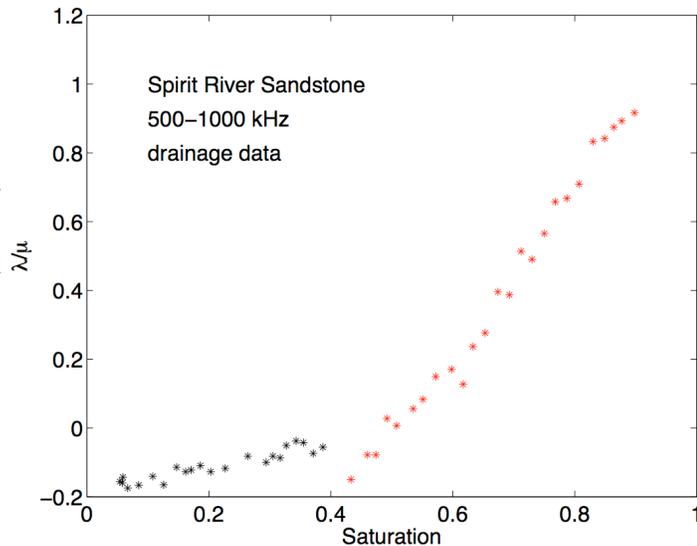


Figure 9: Like Figure 5, but for drainage data and patchy saturation. Again note how similar Figures 8 and 9 are, showing that most of the useful information is in λ , not ρ . [NR]

jim1/. KRlamvmuSRSDD



ACKNOWLEDGMENTS

I would like to thank Steven R. Pride for helpful comments and references.

REFERENCES

- Achenbach, J. D., 1973, Wave propagation in elastic solids: North-Holland/American Elsevier, New York.
- Aki, K. and P. Richards, 1980, Quantitative Seismology: Theory and methods, Volume 1: W. H. Freeman and Company, San Francisco.
- Allen, J. L. and C. Peddy, 1993, Chapter 4: Rock Properties: in Amplitude Variation with Offset: Gulf Coast Case Studies.
- Avellaneda, M. and G. Milton, 1989, Optimal bounds on the effective bulk modulus of polycrystals: SIAM Journal of Applied Mathematics, **49**, 824–837.
- Backus, G. E., 1962, Long-wave elastic anisotropy produced by horizontal layering: Journal of Geophysical Research, **67**, 4427–4440.
- Berryman, J. G., 1999, Origin of Gassmann's equations: Geophysics, **64**, 1627–1629.
- , 2005, Pore fluid effects on shear modulus in a model of heterogeneous rocks, reservoirs, and granular media: Journal of Geophysical Research, **110**, B07202.
- Berryman, J. G., P. Berge, and B. Bonner, 2002, Estimating rock porosity and fluid saturation using only seismic velocities: Geophysics, **67**, 391–404.
- Berryman, J. G. and S. Pride, 2005, Dispersion of waves in porous cylinders with patchy saturation: Formulation and torsional waves: Journal of the Acoustical Society of America, **117**, 1785–1795.
- Biot, M. A., 1956, Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range: Journal of the Acoustical Society of America, **28**, 168–178.
- , 1962, Mechanics of deformation and acoustic propagation in porous media: Journal of Applied Physics, **33**, 1482–1498.
- Biot, M. A. and D. Willis, 1957, The elastic coefficients of the theory of consolidation: Journal of Applied Mechanics, **24**, 594–601.

- Brandsberg-Dahl, S., M. de Hoop, and B. Ursin, 2003, Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers: *Geophysics*, **68**, 232–254.
- Castagna, J. P., 1993, AVO analysis - Tutorial and review: in *Offset Dependent Reflectivity - Theory and Practice*, edited by J. P. Castagna and M. Backus, Investigations in Geophysics, No. 8: Society of Exploration Geophysicists, Tulsa, Oklahoma.
- Contreras, A., C. Torres-Verdin, and T. Fasnacht, 2006, AVA simultaneous inversion of partially stacked seismic amplitude data for the spatial delineation of lithology and fluid units of deepwater hydrocarbon reservoirs in the central Gulf of Mexico: *Geophysics*, **71**, E41–E48.
- Ewing, W. M., W. Jardetsky, and F. Press, 1957, *Elastic waves in layered media*: McGraw-Hill, New York.
- Foster, D. J., R. Keys, and D. Schmitt, 1997, Detecting subsurface hydrocarbons with elastic wavefields: in *Inverse Problems in Wave Propagation*, edited by G. Chavent, G. Papanicolaou, P. Sacks, and W. Symes: Springer, New York.
- Gassmann, F., 1951, über die elastizität poröser medien: *eirteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **96**, 1–23.
- Goodway, W., T. Chen, and J. Downton, 1997, Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters: 67th Annual International Meeting, SEG, Expanded Abstracts, 183–186.
- Hashin, Z. and S. Shtrikman, 1963, A variational approach to the theory of the elastic behaviour of multiphase materials: *Journal of the Mechanics and Physics of Solids*, **11**, 127–140.
- Hill, R., 1963, Elastic properties of reinforced solids: Some theoretical principles: *Journal of the Mechanics and Physics of Solids*, **11**, 357–372.
- Knight, R. and R. Nolen-Hoeksema, 1990, A laboratory study of the dependence of elastic wave velocities on pore scale fluid distribution: *Geophysical Research Letters*, **17**, 1529–1532.
- Koefoed, O., 1955, On the effect of poisson's ratios of rock strata on the reflection coefficients of plane waves: *Geophysical Prospecting*, **3**, 381–387.
- , 1962, Reflection and transmission coefficients for plane longitudinal incident waves: *Geophysical Prospecting*, **10**, 304–351.
- Krief, M., J. Garat, J. Steffingwerff, and J. Ventre, 1990, A petrophysical interpretation using the velocities of p and s waves (full-waveform sonic): *Log Analyst*, **31**, 355–369.
- Landau, L. D. and E. Lifshitz, 1986, *The theory of elasticity*: Butterworth-Heinemann, Oxford.
- Love, A. E. H., 1944, *A treatise on the mathematical theory of elasticity*: Dover, New York.
- Mavko, G. and D. Jizba, 1991, Estimating grain-scale fluid effects on velocity dispersion in rocks: *Geophysics*, **56**, 1940–1949.
- Mavko, G., T. Mukerji, and J. Dvorkin, 1998, *The rock physics handbook: Tools for seismic analysis in porous media*: Cambridge University Press, Cambridge, UK.
- Milton, G. W., 2002, *The theory of composites*: Cambridge University Press, Cambridge, UK.
- Murphy, W. F., I., 1982, Effects of partial water saturation on attenuation in Massilon sandstone and Vycor porous-glass: *Journal of the Acoustical Society of America*, **71**, 1458–1468.
- Ostrander, W. J., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence: *Geophysics*, **49**, 1637–1648.
- Postma, G. W., 1955, Wave propagation in a stratified medium: *Geophysics*, **20**, 780–806.

- Reuss, A., 1929, Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle: *Z. Angew. Math. Mech.*, **9**, 49–58.
- Shuey, R. T., 1985, A simplification of the Zoeppritz equations: *Geophysics*, **50**, 609–614.
- Ursin, B., B. Ekren, and E. Tjåland, 1996, Linearized elastic parameter sections: *Geophysical Prospecting*, **44**, 427–455.

Maximum entropy spectral analysis

James G. Berryman

ABSTRACT

A review of the maximum entropy spectral analysis (MESA) method for time series is presented. Then, empirical evidence based on maximum entropy spectra of real seismic data is shown to suggest that $M = 2N/\ln 2N$ is a reasonable *a priori* choice of the operator length M for discrete time series of length N . Various data examples support this conclusion.

INTRODUCTION

When analyzing seismic traces, it is often useful to know what frequencies are present in the data. Filtering and smoothing of data should be done with knowledge of the frequency content. In the standard approach to spectral analysis, the Fourier transform of the trace (amplitude spectrum) is computed. This approach is quite reliable for long data sequences (1000 or more data points) and is satisfactory for somewhat shorter sequences. Unfortunately, this technique becomes unreliable for very short time samples due to the increased importance of end effects: (a) the resolution of true peaks in the spectrum becomes poor and (b) spurious peaks may be introduced because of the implicit (and incorrect) assumption often made that the known data sequence is repeated periodically in time.

A different approach to spectral analysis was introduced into the geophysical literature by Burg (1975). His idea was to obtain an estimate of the power spectrum (square of the amplitude spectrum) by maximizing the spectral entropy with the known autocorrelation values as constraints. In principle, this approach should give a power spectrum that is consistent with the available information, but maximally noncommittal with regard to the unavailable information. It turns out that the resulting mathematical problem can be solved exactly using linear matrix theory. In fact, the method requires computation of the minimum phase deconvolution operator [also known as the “prediction error filter” (Claerbout, 1976)], which has received much attention in the geophysical literature. The power spectrum is then given by the square inverse of the operator’s Fourier transform. Burg’s method is known as maximum entropy spectral analysis (MESA) and is closely related both to deconvolution and to autoregressive analysis of stationary random time series.

The method of computing the spectrum in MESA can be easily understood in terms of filter theory. If we apply a prediction error filter to an input time series, the output will be a white spectrum. It is well-known that the spectrum of the output is the spectrum of the input times the spectrum of the filter. Since a white spectrum is constant, an estimate of the input spectrum is given by the inverse of the spectrum of the prediction error filter.

MESA has one principal advantage over the standard Fourier transform method of spectral analysis: resolution of peaks in the power spectrum is enhanced for short data

sequences. MESA has two principal disadvantages: (a) computation time is increased (substantially for long data sequences) and (b) the best choice of for the operator length is not known (poor choices can give misleading results for short data samples). A possible solution to this second problem is discussed in the section on **Choosing the Operator Length**.

At least two other approaches to spectral analysis are possible. (a) The maximum likelihood method or MLM Lacoss (1971) has been shown by Burg (1972) to be the inverse of the arithmetic average of inverse maximum entropy spectra of increasing operator length. Thus, MLM weights the strongest peaks of MESA the least and cannot give very good resolution. (b) Using the terms of stochastic theory (Jenkins and Watts, 1968), the ordinary power spectrum assumes that the underlying process is a moving average (MA) process. Using MESA can be viewed as being equivalent to assuming the process is autoregressive (AR). In fact, a discretely sampled geophysical time series is most likely to be a combination of the two, namely an autoregressive-moving-average (ARMA) process. It is possible to estimate the spectrum under the ARMA assumption; however, a substantial increase in computation time is required (over MESA), while the resolution of peaks should remain nearly the same.

A brief discussion of the theory and practice of MESA has appeared previously in Berryman (1978). An expanded version of this account is given in the following pages. The work presented here leads to the conclusion that for short time series MESA may well be a useful tool, and that MESA is probably the best available alternative to standard methods for such short data processing problems.

THE VARIATIONAL PRINCIPLE

Given a discrete (possibly complex) time series $\{X_1, \dots, X_N\}$ of N values with sampling interval Δt (and Nyquist frequency $W = 1/2\Delta t$), we wish to compute an estimate of the power spectrum $P(f)$, where f is the frequency. It is well known that

$$P(f) = \lim_{N \rightarrow \infty} \frac{1}{N} \left| \sum_{n=1}^N X_n \exp(i2\pi fn\Delta t) \right|^2 = \sum_{n=-\infty}^{\infty} R_n \exp(i2\pi fn\Delta t), \quad (1)$$

where the autocorrelation function R is defined by (for $n \geq 0$)

$$R_n = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N-n} X_i^* X_{i+n} = R_{-n}^*. \quad (2)$$

Now suppose that we use the finite sequence $\{X_i\}$ to estimate the first M autocorrelation values R_0, \dots, R_{M-1} . (Methods of obtaining these estimates are discussed in the section on **Computing the Prediction Error Filter**.) Then, Burg (1975) has shown that maximizing the average entropy (see Appendix A for a derivation)

$$h = \frac{1}{4W} \int_{-W}^W \ln [2WP(f)] df, \quad (3)$$

subject to the constraint that (1) is equivalent to extrapolating the autocorrelation R_n for $|n| \geq M$ in the most random possible manner.

Doing the math, we find that

$$\frac{\delta h}{\delta R_n} = \frac{1}{4W} \int_{-W}^W P^{-1}(f) \exp(i2\pi f n \Delta t) df = \begin{cases} \lambda_n/2 & \text{for } |n| < M \\ 0 & \text{for } |n| \geq M. \end{cases} \quad (4)$$

The λ 's are Lagrange multipliers to be determined. That the variation of h with respect to R_n for $|n| \geq M$ should be zero is the essence of the variational principle. The value of h is then stationary with respect to changes in the R_n 's, which are unknown. We can infer from Equation (4) that

$$P^{-1}(f) = \sum_{n=-(M-1)}^{M-1} \lambda_n \exp(-i2\pi f n \Delta t). \quad (5)$$

Making the Z -transform to $Z = \exp(-i2\pi f \Delta t)$, Equation (5) becomes a polynomial of the complex parameter Z :

$$P^{-1}(f) = \sum \lambda_n Z^n. \quad (6)$$

Since P is necessarily real and nonnegative, Equation (6) can be uniquely factored as

$$P^{-1}(f) = 2W E_M^{-1} \left[\sum_m a_m Z^m \right] \left[\sum_n a_n^* Z^{-n} \right] = 2W E_M^{-1} \left| \sum_n a_n Z^n \right|^2, \quad (7)$$

with $a_0 = 1$. The first sum in (7) has all of its zeroes outside the unit circle (minimum phase) and the second sum has its zeroes inside the unit circle (maximum phase).

Fourier transforming Equation (1), we find that

$$R_n = \int_{-W}^W P(f) \exp(-i2\pi f n \Delta t) df. \quad (8)$$

Substituting (7) into (8), we find (after a few more transformations) that R_n is given by the contour (complex) integral

$$R_n = \frac{E_M}{2\pi i} \oint_{|Z|=1} \frac{Z^{n-1}}{\left| \sum a_m Z^m \right|^2} dZ. \quad (9)$$

The integrand of (9) can have simple poles inside the contour of integration at $Z = 0$ and at any zero of the maximum phase factor. The poles for $Z \neq 0$ can be eliminated by taking a linear combination of Equation (9) "for various values of n ." Using the Cauchy integral theorem, we find that

$$\sum_j a_j^* R_{n-j} = \frac{E_M}{2\pi i} \oint_{|Z|=1} \frac{Z^{n-1}}{\sum a_m Z^m} dZ = \begin{cases} E_m & \text{for } n = 0 \\ 0 & \text{for } n > 0, \end{cases} \quad (10)$$

since $a_0 = 1$. Equation (10) and its complex conjugate for the a_m are exactly the standard equations for the maximum and minimum phase spike deconvolution operators $\{a_m^*\}$ and $\{a_m\}$, respectively.

Notice that, if we define the $N \times N$ matrix T_{N-1} as the equidiagonal matrix of auto-correlation values whose elements are given by

$$[T_{N-1}]_{ij} \equiv R_{i-j}, \quad (11)$$

then Equation (10) may be seen as a problem of inverting the matrix T to find the vector $\{a_{N-1}^*, \dots, 1\}$. Equation (10) can be solved using the well-known Levinson algorithm for inverting a Toeplitz matrix (Wiener, 1949). Therefore, a power spectral estimate can be computed by using (10) to find the a_n 's, and (7) to compute the spectrum.

One gap in the analysis should be filled before we proceed. That the variational principle is a stationary principle (*i.e.*, $\delta h = 0$) is obvious. That it is truly a maximum principle however requires some proof. First note that the average entropy h computed from substituting (7) into (3) is exactly

$$h = \frac{1}{2} \ln E_M. \quad (12)$$

This fact can be proven by writing (3) as

$$2h = \ln E_M + \frac{M-1}{2\pi i} \oint \ln Z \times \frac{dZ}{Z} - \frac{1}{2\pi i} \oint \ln \left(\sum_n a_n Z^n \right) \frac{dZ}{Z} - \frac{1}{2\pi i} \oint \ln \left(\sum_n a_n^* Z^{M-n-1} \right) \frac{dZ}{Z}. \quad (13)$$

The first integral in (13) vanishes identically as is shown in Appendix B. The second integral vanishes because its argument is analytic for all $|Z| < 1$ except for $Z = 0$, and the residue there is $\ln a_0 = 0$. The third integral can be rewritten as

$$\frac{1}{2\pi i} \oint \ln \left(\sum_n a_n^* Z^{M-n-1} \right) \frac{dZ}{Z} = \ln a_0 + \sum_{n=1}^{M-1} \frac{1}{2\pi i} \oint \ln (Z - Z_n) \frac{dZ}{Z}, \quad (14)$$

where the Z 's are the $M - 1$ zeroes of the maximum phase factor ($|Z| < 1$). Each of the integrals on the right side of (14) vanishes because of the identities proven in Appendix B.

For small deviations from the constraining values of R_n , and from the values of R_n computed from (8) once $P_M(f)$ is known, we can expand h in a Taylor series:

$$h = \frac{1}{2} \ln E_M + \sum_{n=-(M-1)}^{M-1} \lambda_n r_n - \sum_{m,n=-\infty}^{\infty} H_{mn} r_m r_n^*. \quad (15)$$

The r_n 's are small deviations in the R_n 's. The λ_n 's are defined by (4). The matrix elements of H are given by

$$H_{mn} = -\frac{\delta^2 h}{\delta R_m \delta R_n^*} = \frac{1}{4W} \int_{-W}^W \frac{Z^{n-m}}{P^2(f)} df, \quad (16)$$

with $Z = \exp(-i2\pi f \Delta t)$. H is obviously Hermitian and is seen to be positive definite because

$$\sum_{mn} H_{mn} v_m v_n^* = \frac{1}{4W} \int_{-W}^W \frac{|\sum v_m Z^{-m}|^2}{P^2(f)} df \geq 0, \quad (17)$$

where $\{v_n\}$ is an arbitrary complex vector and the equality in (17) holds only when $\{v_n\}$ is identically zero.

The result (17) is sufficient to prove that h is not only stationary, but actually a maximum.

The analysis given in this section has at least two weak points: (a) For real data, we never measure the autocorrelation function directly. Rather, a finite time series is obtained and

an autocorrelation estimate is computed. Given the autocorrelation estimate, an estimate of the minimum phase operator must then be inferred. A discussion of various estimates of the autocorrelation is given in the next section on **Computing the Prediction Error Filter**, along with a method of estimating the prediction error filter without computing an autocorrelation estimate. (b) Even assuming we could compute the “best” estimate of the autocorrelation, that estimate is still subject to random error. The probability of error increases as we compute values of R_n with greater lag n . Since there is a one-to-one correspondence between the R_n 's and the a_n 's, the length of the operator can strongly affect the accuracy of the estimated MESA power spectrum. A method of estimating the optimum operator length for a given sample length N will be discussed in the subsequent section on **Choosing the Operator Length**.

COMPUTING THE PREDICTION ERROR FILTER

When the autocorrelation values R_0, \dots, R_{M-1} are known, Equation (10) is a linear set of M equations for the M unknown a_M^* 's. On the other hand, if a prediction error filter $\{a_m\}$ and prediction error E_m are known, Equation (10) together with (2) forms a linear set of equations that could be solved for the R_n 's. Thus, there exists a one-to-one correspondence between the prediction error filter and the autocorrelation function. This relationship is exploited by Burg (1975) in his algorithm for computing the minimum phase operator.

The autocorrelation function defined by (2) requires an infinite series, yet it can only be estimated from a series of finite length N . Given the data set $\{X_1, \dots, X_N\}$, a reasonable estimate of R_n for large N is given by

$$R_n = \frac{1}{N} \sum_{m=1}^{N-n} X_m^* X_{m+n} \quad \text{for } 0 \leq n \leq N-1. \quad (18)$$

This estimate has at least two shortcomings: (a) Conceptually, the autocorrelation should be an arithmetic average of the $N-n$ lag products in (18). The true arithmetic average is (for $n \geq 0$)

$$R'_n = \frac{1}{N-n} \sum_{m=1}^{N-n} X_m^* X_{m+n} \simeq \frac{N}{N-n} R_n. \quad (19)$$

Equation (19) might be used as the autocorrelation estimate instead of (18). Unfortunately, this is seldom possible because the Hermitian Toeplitz matrix T defined in (11) is not always nonnegative definite when the definition (19) is used (Parzen, 1964). A stable operator $\{a_m\}$ cannot be found if T is not nonnegative definite. We conclude that (19) is not a satisfactory estimate of R_n . (b) Suppose for the moment that the matrix T computed using (19) happens to be positive definite. Then each estimated R'_n is being computed from only $N-n$ measurements of the n -lag product, whereas R_0 is estimated from N measurements of the zero-lag product. From measurement theory, it is clear that the uncertainty increases approximately as $(N-n)^{-\frac{1}{2}}$. In fact, this increase in the uncertainty of R_n is unavoidable regardless what choice of estimate for R_n we use as long as N remains finite. One might try to alleviate this problem by using periodic boundary conditions, so that

$$R_n = \frac{1}{N} \sum_{m=1}^N X_m^* X_{m+n} \quad (20)$$

and

$$X_{m+N} \equiv X_m. \quad (21)$$

However, this approach merely trades one problem for another one. The periodic assumption introduces spurious peaks into the spectrum by making unfounded assumptions about time series behavior off the ends of the data. Although nevertheless a fairly common approach, this method really cannot improve the accuracy of the computed R_n 's for seismic traces having typical lengths.

We conclude that, if an autocorrelation must be computed, then Equation (18) should be used. However, Burg (1975) has observed that, in order to compute the maximum entropy spectrum, all that is required is an estimate of the minimum phase deconvolution operator. If this estimate can be computed without first estimating the autocorrelation values, then so much the better.

Suppose an estimate of the operator length M is known ($a_0 = 1$ for $M = 1$). How can the operator length be increased from M to $M + 1$? Note that by definition the forward prediction error is given by

$$f_{i+M}(M) = \sum_{j=0}^{M-1} a_j(M) X_{i+M-j} \quad \text{for } 1 \leq i \leq N - M, \quad (22)$$

and the backward prediction error is

$$b_i(M) = \sum_{j=0}^{M-1} a_j^*(M) X_{i+j} \quad \text{for } 1 \leq i \leq N - M. \quad (23)$$

Similarly, we have

$$f_{i+M+1}(M + 1) = \sum_{j=0}^M a_j(M + 1) X_{i+M+1-j}, \quad (24)$$

and

$$b_i(M + 1) = \sum_{j=0}^M a_j^*(M + 1) X_{i+j}, \quad (25)$$

which are the linear combinations of (22) and (23) given by

$$f_{i+M+1}(M + 1) = f_{i+M+1}(M) + C_{M+1} b_i(M) \quad (26)$$

and

$$b_i(M + 1) = b_i(M) + C_{M+1}^* f_{i+M+1}(M). \quad (27)$$

Assuming the value of C_{M+1} is known, (22)–(27) can be used to show that the recursion formulas for the a 's are:

$$\begin{aligned} a_0(M + 1) &= 1, \\ a_i(M + 1) &= a_i(M) + C_{M+1} a_{M-i}^*(M) \quad \text{for } 1 \leq i \leq M - 1, \quad \text{and} \\ a_M(M + 1) &= C_{M+1}. \end{aligned} \quad (28)$$

Equation (28) is exactly the recursion relation for a minimum phase operator when $|C_i| < 1$ for all $i \leq M + 1$. Thus, estimating the a 's reduces to estimating the C 's. A criterion for choosing C_{M+1} is still required.

Burg (1975) suggests that one reasonable procedure is to choose the C_{M+1} that minimizes the total power of the prediction errors. Setting

$$\frac{d}{dC_{M+1}^*} \sum_{i=1}^{N-M-1} [|f_{i+M+1}(M+1)|^2 + |b_i(M+1)|^2] = 0, \quad (29)$$

the estimate becomes

$$C_{M+1} = -\frac{2 \sum b_i^*(M) f_{i+M}(M)}{\sum [|f_{i+M}(M)|^2 + |b_i(M)|^2]}. \quad (30)$$

Substituting (30) for C_{M+1} into the total power, it is not difficult to show that

$$\begin{aligned} 0 &\leq \frac{1}{2} \sum [|f_{i+M+1}(M+1)|^2 + |b_i(M+1)|^2] \\ &= \frac{1}{2} \sum [|f_{i+M}(M)|^2 + |b_i(M)|^2] (1 - |C_{M+1}|^2) \\ &= E_M (1 - |C_{M+1}|^2) = E_{M+1}, \end{aligned} \quad (31)$$

where $E_0 = R_0$. Equation (31) guarantees that $|C_{M+1}| \leq 1$, as is required for a to be minimum phase.

Finally, the algorithm for computing the set $\{a_m\}$ is this: (a) Compute the C 's using Equation (30). (b) Store the C 's until the desired operator length M is attained. (c) Compute the a 's from the C 's using the recursion (28). This algorithm (simplified for real data) is the one used in the maximum entropy processor for MESA that I developed.

It is important to notice before proceeding further that the Burg algorithm has been constructed to remove the first difficulty discussed earlier in computing R_n . All the information $\{X_n\}$ has been used; the operator is minimum phase; but no explicit averaging of lag products was required. On the other hand, this algorithm does nothing to alleviate the second problem we discussed. It is still inherent in the finite time series problem that the numbers we compute become less reliable as the operator length increases.

A major difficulty in applying MESA is that there is no built-in mechanism for choosing the operator length. From the derivation of (5), it is clear the operator length should be N if the first N autocorrelation values are known precisely and unknown otherwise. However, the autocorrelation function has (normally) been estimated from the time series data and its estimated values are inaccurate for n close to N . How to choose a practical operator length M satisfying $1 < M < N$ is therefore the subject of the next section.

CHOOSING THE OPERATOR LENGTH

Numerous procedures for choosing the operator length have been discussed in the literature. In this section of the paper, first a discussion of the general principles behind the operator length optimization is given. Then several of the most prominent practical methods in use are critically reviewed. Finally, a new criterion is derived which is easier to apply and believed to be more appropriate for use in MESA.

A mean square error criterion

One theoretically sound procedure for choosing a truncation point is based on a mean square error criterion (Jenkins and Watts, 1968). Suppose that $P_M(f)$ is our M -th estimate of the

true spectrum $P(f)$. Then we might wish to minimize the square error

$$E[P_M(f) - P(f)]^2 = Var[P_M(f)] + B^2[P_M(f)], \quad (32)$$

where E is expectation, Var is variance

$$Var(P_M) = E(P_M^2) - E^2(P_M), \quad (33)$$

and B is bias

$$B[P_M(f)] = P(f) - E[P_M(f)]. \quad (34)$$

In general, as M increases, the bias decreases while the variance increases. Thus, (32) will have a minimum for some value of M .

This criterion is not of practical value unless it is possible to obtain reasonably good estimates of the variance and bias of P_M . This problem is not easily solved, but satisfactory approximate solutions can probably be found. However, this approach will NOT be pursued here.

This notation has been introduced to help the reader understand why one should *expect* such an optimum operator length to exist. Spectral estimates are nearly always designed to decrease the bias as M increases. (MESA is clearly designed this way.) However, when the bias *is* small, the variance is a measure of the ragged oscillations $P_M(f)$ makes around $P(f)$. Since most people prefer to study a smooth spectrum, a balance between variance and bias is our goal.

The common criteria

A number of fairly simple criteria are commonly discussed in the literature. Some of these will be reviewed here.

Treitel et al. (1977) suggest monitoring the magnitude of C_M to determine the operator length empirically. Their criterion is to choose that M value for which C_M first satisfies $|C_M| \ll 1$. The argument is that C_M computed from (30) is “a partial correlation coefficient,” measuring the correlation between the forward and backward prediction errors. When $|C_M| \simeq 1$, the correlation is high. When $|C_M| \ll 1$, the correlation is low — presumably because most of the predictable information in the data has been removed by the filter. However, they point out that this procedure fails to produce reliable results for series not purely autoregressive in character. Numerical studies of the author on real seismic traces have shown the fluctuations in $|C_M|$ to be too great for this approach to give a reliable criterion.

Berkhout and Zaanen (1976) review a number of possible approaches. The two which are probably easiest to apply are the F -test and the relative error coefficient test: (a) The F -test monitors E_M and checks whether the change in going from E_M to E_{M+1} is statistically significant according to some predetermined criterion. This method is limited by computer round off error for large data samples. It is also limited for small data samples because the predetermined criterion of statistically significant change may very well be met for all $M \leq N$. (b) The relative error coefficient test amounts to finding the minimum of the modified prediction error

$$E'_M = \frac{N}{N-M} E_M. \quad (35)$$

The prediction error is modified in this manner to account for the decrease in the degrees of statistical freedom for the time series as the operator length increases. Clearly, E_M decreases whereas the multiplicative factor increases as M increases. E'_M will therefore exhibit a minimum. A number of such minima can (and do) occur in practice. The parameter E'_M is easily monitored while computing the C_M 's using the Burg algorithm. Results obtained using this approach have been found satisfactory for moderate to large data samples. For small N , the variations in both factors in (35) can be dramatic and the results become less reliable.

Ulrych and Bishop (1975) review a number of alternatives and conclude that the final prediction error (FPE) criterion of Akaike (1970) is an objective basis for choosing the operator length. This criterion monitors

$$(FPE)_M = \frac{N + M}{N - M} E_M. \quad (36)$$

Like (35), this expression has a minimum since E_M decreases monotonically while the multiplicative factor increases monotonically with M . In fact, (35) and (36) have very similar behavior, the principal difference being that "when E_M is sufficiently smoothly varying," the minimum of (36) always occurs for smaller M values than that of (35). For short time series with sharp spectral lines, Ulrych and Bishop (1975) found that FPE did not give a clear minimum. Both (35) and (36) suffer from this same ambiguity. For data samples of length $20 \leq N \leq 40$ in their work, they found that $M = N/2$ was a satisfactory choice. This choice is also confirmed for short time series by the work of Chen and Stegen (1974).

Although each of these criteria has its merits, none of them is really satisfactory for a data sample of arbitrary length. Furthermore, none of them has been derived in the spirit of MESA, *i.e.*, with no assumptions about the data off the ends of the sample. Much has been said about the application of optimum criteria from autoregressive analysis to MESA (Ulrych and Bishop, 1975). But an important point should be made: The fact that an autoregressive process has the maximum entropy is interesting but irrelevant. The spectrum of an *arbitrary* time series (whether MA, AR, or ARMA of any order) can be *estimated* using MESA. But, making *any* assumption about the nature of the stochastic process that generated the series is contrary to the spirit of MESA.

Thus, it seems that the choice of operator length should be made without assumptions concerning the nature of the stochastic process involved. The argument in the next subsection is based only on information theory, and measurement theory. It is believed to free of these inconsistencies.

An information theory criterion

Suppose we have found an estimate of the prediction error filter of length M using the autocorrelation estimates R_0, \dots, R_{M-1} . In order to increase the operator length to $M + 1$, additional information is needed: namely, R_M . A quantitative measure of the information in the operator is easily obtained from the average entropy, which we know is given by $h'_m = \frac{1}{2} \ln E_M$. Using (31), notice that

$$h'_{M+1} = \frac{1}{2} \ln E_M + \frac{1}{2} \ln (1 - |C_{M+1}|^2) \leq h'_M. \quad (37)$$

Thus, the entropy *decreases* as the operator length *increases*. The bound information (Brilouin, 1956) I_M in the power spectrum is therefore given by

$$I'_M = -h'_M = -\frac{1}{2} \ln E_M, \quad (38)$$

which obviously [since, from (37), we have $-h'_M \leq -h'_{M+1}$] increases monotonically with M as it should.

If the autocorrelation values R_0, \dots, R_{n-1} were known precisely, bound information would continue to increase by using all the estimates and letting $M \rightarrow N - 1$. But the R 's are not precisely known. The finite number of measurements used to compute the R estimates means that only $N - n$ measurements of R_n were made, whereas N measurements of R_0 were made. The quality of information contained in R_0 is correspondingly higher than that in R_n . A quantitative measure of this change is therefore required.

For the moment, take Equation (19) as our estimate of the autocorrelation. Then, assuming that the X_i 's are normally distributed, Parzen (1962) shows that

$$\text{Var}(X_i^* X_{i+n} - R_n) = R_0^2 + R_n^2. \quad (39)$$

Since $|R_0| \geq |R_n|$, for all n , and generally $|R_0| \gg |R_n|$, for large n , the variance (39) can be approximated by the constant R_0^2 . Viewing R_n as a measured quantity (which in fact it usually is not) and using standard arguments from measurement theory, we find that

$$R_n = \frac{1}{N - n} \sum_m X_m^* X_{m+n} \pm 0.67 \frac{R_0}{\sqrt{N - n}}, \quad (40)$$

with fifty percent confidence if R_n is also normally distributed.

The probable error in R_n increases like $(N - n)^{-1/2}$ as $n \rightarrow N - 1$. We imagine that the factor $(N - n)^{-1/2}$ is proportional to the probability p_n that an operator computed from R_0, \dots, R_N is a worse estimate of the true operator than was the operator computed using only R_0, \dots, R_{n-1} . Since we know empirically that the estimate worsens as $M \rightarrow N$ with probability one, the $P(n)$ are normalized by writing:

$$P(n) = \alpha(N - n)^{-\frac{1}{2}}, \quad (41)$$

and

$$1 = \sum_{n=0}^{N-1} P_n \simeq \alpha \int_0^N (N - n)^{-\frac{1}{2}} dn = 2\sqrt{N}\alpha. \quad (42)$$

Equation (42) determines the value of α , for the data that are available.

The average entropy of measurement error associated with an operator of length M is

$$\begin{aligned} h''_M &= -\sum_{n=0}^{M-1} P(n) \ln P(n) \\ &\simeq -\int_0^M P(n) \ln P(n) dn. \end{aligned} \quad (43)$$

The second line of (43) is valid for large N . The value of h''_M increases as M increases in agreement with our intuition. It is well known that the largest average entropy for N probabilities is $\ln N$. Letting $M \rightarrow N$ in (43), we find

$$h''_N = \ln \left(\frac{2N}{e} \right) < \ln N, \quad (44)$$

which is consistent.

Combining (38) and (43), the average information in the power spectrum can be quantitatively estimated using the expression

$$I_M = -(h'_M + h''_M) = -\frac{1}{2} \ln E_M + \int_0^M P(n) \ln P(n) dn. \quad (45)$$

The first term increases while the second term decreases with increasing M . A maximum will occur for some value $1 < M < N$. The spectrum with the maximum information is the optimum spectrum; the value of M that maximizes (45) is the value we are seeking.

The values of (45) can be monitored continuously while the operator is being computed. However, an approximate analytic solution for the maximum can be found without making very restrictive assumptions on the behavior of E_M . Numerical studies of the author on real seismic data have shown that E_M can be represented approximately by

$$E_M \propto M^{-\beta}, \quad (46)$$

where β is a slowly varying function of M . Generally, β is in the range $2 \geq \beta \geq \frac{1}{2}$, with $\beta \simeq 2$ for small M and $\beta \rightarrow \frac{1}{2}$ for large M . Leaving β arbitrary for the moment, substituting (46) into (45), and finding the stationary point, we have

$$\frac{\beta}{2} M^{-1} = -P(M) \ln P(M). \quad (47)$$

Using (42) for α , Equation (47) can be solved graphically for M . The solution for $\beta = 2$ is plotted as the solid line in Figure 1.

An analytic bound on M can be obtained from (47) by noting that the right-hand side of (47) increases with M , so its minimum value occurs when $M = 0$. Thus, M has the very simple bound:

$$M \leq \beta \frac{N}{\ln 2N}. \quad (48)$$

Since we have stated already that $\beta \leq 2$ in general, a useful bound on M for all N appears to be

$$M \leq \frac{2N}{\ln 2N}. \quad (49)$$

Figure 1 compares the values of M obtained from (47), from (48), and from $M = N/2$. The value $\beta = 2$ is chosen because of the empirical evidence mentioned above and also because

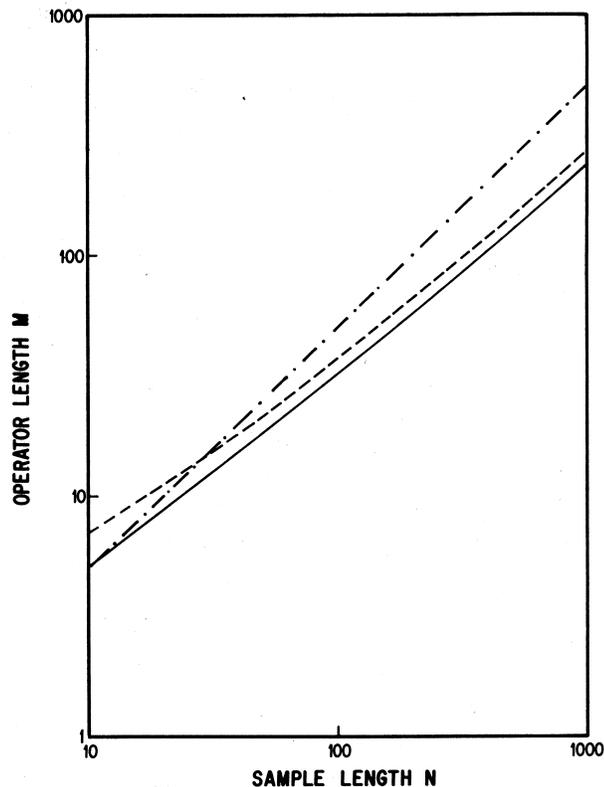
$$|h'_M| \leq \frac{\beta}{2} \ln M \leq \ln N \quad (50)$$

is valid for all $M \leq N$ only for $\beta \leq 2$. The comparison with $M = N/2$ is of interest because various authors (including this one) have often found this value to be satisfactory for small N . The derivation given above is strictly valid only for large N . But the estimate (49) interpolates well between these extremes as is seen in Figure 1.

Because the correspondence between $P(n)$ and $(N - n)^{-\frac{1}{2}}$ has been established by this heuristic argument, the results of this section of the paper should not be interpreted as rigorous estimates of the optimum operator length. Nevertheless, I believe that (47) and (49) are reasonable estimates of the operator length. The derivation was not founded on

any assumptions about the type of stochastic process generating the time series. Hence, these estimates are definitely *not* intended to be an estimate of the order of some underlying autoregressive process. Rather, (49) is an upper bound on the operator length that will extract the most reliable information for a data sample of length N . For example, suppose the time series $\{X_1, \dots, X_N\}$ is a representation of an autoregressive series of order $L \leq M$. Then computing the operator of length L should give the most efficient estimate of the spectrum; but computing the additional $(M - L)$ terms should do little to alter that spectrum. Next, suppose the time series is a representation of an AR series of order $L > M$. The arguments above indicate that we probably cannot obtain a really good estimate of the operator (or the spectrum), because our data sample is simply too small. The best we can hope to do is to compute the operator of length M . In either case, when additional information about the underlying stochastic process is lacking, the best operational decision that can be made appears to be choosing M according to Equations (47) or (49).

Figure 1: Operator length M as a function of data sample length N for three different operator length estimates. The solid line is the solution of Equation (47). The dash line is $M = 2n / \ln(2N)$. The dot-dash line is $M = N/2$. [NR] jim2/. FIG1



EXAMPLES

One good way to study various choices of operator length is to use unexpanded checkshot data. Two traces of this type are shown in Figure 2. By choosing a large window, we obtain a good estimate of the spectrum of the pulse. Then, by choosing smaller and smaller windows which pinch down on the pulse, we should expect the positions of the major peaks to remain unchanged while resolution becomes more difficult.

The first column of Figure 3 gives the ordinary power spectrum of Figure 2(a) com-

puted using five different length windows. In seconds, the windows are from top to bottom (0.0,2.0), (0.1,0.5), (0.15,0.45), (0.15,0.25), and (0.175,0.225). No taper is included in the Fourier transform of the trace. We see that the resolution is quite good for the two second window, but the resolution gets progressively worse until it is almost nonexistent for the 50 ms window.

In contrast, the first column of Figure 4 gives the maximum entropy spectrum of Figure 2(a) with the operator length $M = 2N/\ln 2N$. We see that the two second window MESA spectrum is essentially the same as the ordinary spectrum. However, as the number of data points decreases from 500 (for a two second window) to 13 (for a 50 ms window), we see that MESA is still able to resolve the peaks at 10 Hz and 30 Hz.

The first column of Figures 5–7 give examples of the results obtained from MESA for the trace of Figure 2(a) with other choices of operator length. Choosing $M = N/2$ in Figure 5 gives acceptable results for all but the smallest window where the 10 Hz peak has moved towards 20 Hz. Also, the computation time was increased for the longest three windows. Choosing $M = N$ in Figure 6 demonstrates the fact that choosing a longer operator does *not* lead to improved results. Here the single peak at 10 Hz has been split into two spurious peaks for the shortest two windows. The longest three windows give very spiky spectra and (although they properly indicate where the spectrum lies) they do not give useful power spectra. This Figure shows how the variance in (32) can dominate the bias and produce useless power spectra. Finally, we choose $M = N/\ln N$ in Figure 7 to demonstrate the effect of choosing a slightly different functional form for M . The spectrum of Figure 4 mimics that of Figure 3 better than Figure 7 in all cases except possibly for the 100 ms window where Figure 7 gives a stronger peak at 30 Hz. For the smallest window, Figure 7 does not have its peak at 10 Hz as it should.

Figure 2(b) is also an unexpanded checkshot trace which is translated in time from that in Figure 2(a). The second columns of Figures 3–7 were computed as before, but the input trace was Figure 2(b).

As a final note, we wish to point out that it has been observed in general that $M = 2N/\ln 2N$ is in fact an upper bound on the operator lengths one would obtain from either (35) or (36). For example, using the trace of Figure 2, both the relative error coefficient test and the FPE have a series of minima for $120 < M < 135$ with an absolute minimum at $M = 132$ for both criteria. For comparison, we find $M = 145$ for $N = 500$ using (49) and $M = 127$ using (47). It is encouraging that the arguments of the subsection on **An Information Theory Criterion** give estimates for the operator length so close to those of Equations (35) and (36) without the added complication of monitoring a performance parameter.

DISCUSSION

As with many procedures in seismic data enhancement and analysis, computation of power spectral estimates with MESA can be more art than science. Considerable insight into the stochastic processes involved and experience with choices of operator length is required before MESA can be considered a standard processing tool. The work summarized here was intended to eliminate some of the uncertainty in applying MESA by producing a reasonable upper bound [Equation (49)] on the operator length. In many cases, this upper bound will

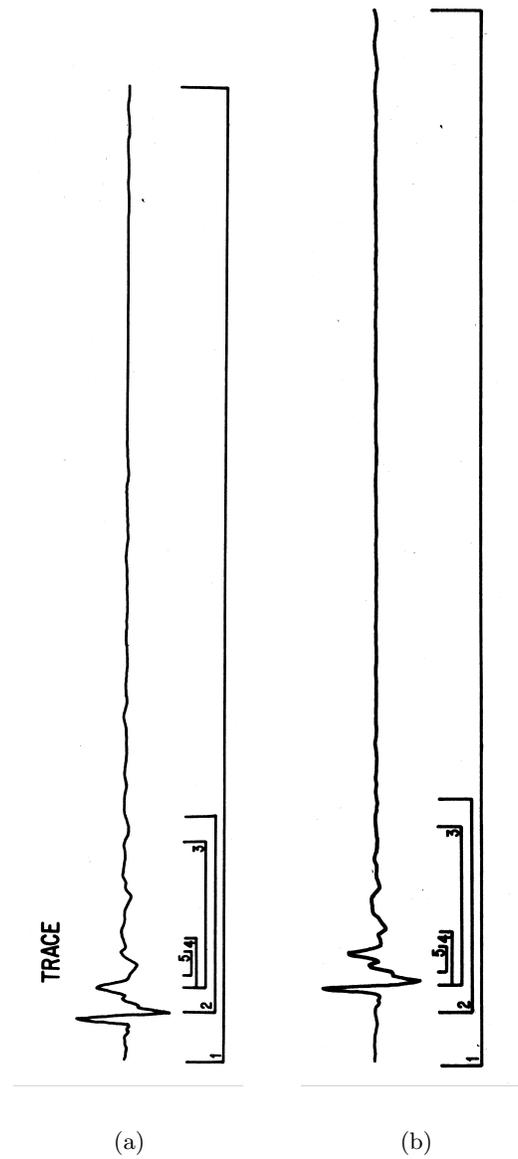


Figure 2: Two checkshot traces: (a): The first two seconds of an unexpanded checkshot trace. The windows indicated in the Figure are respectively: $(0.0, 2.0)$, $(0.1, 0.5)$, $(0.15, 0.45)$, $(0.15, 0.25)$, and $(0.175, 0.225)$ in seconds. (b) Same as previous case for a slightly different trace. [NR] jim2/. FIG2a,FIG2b

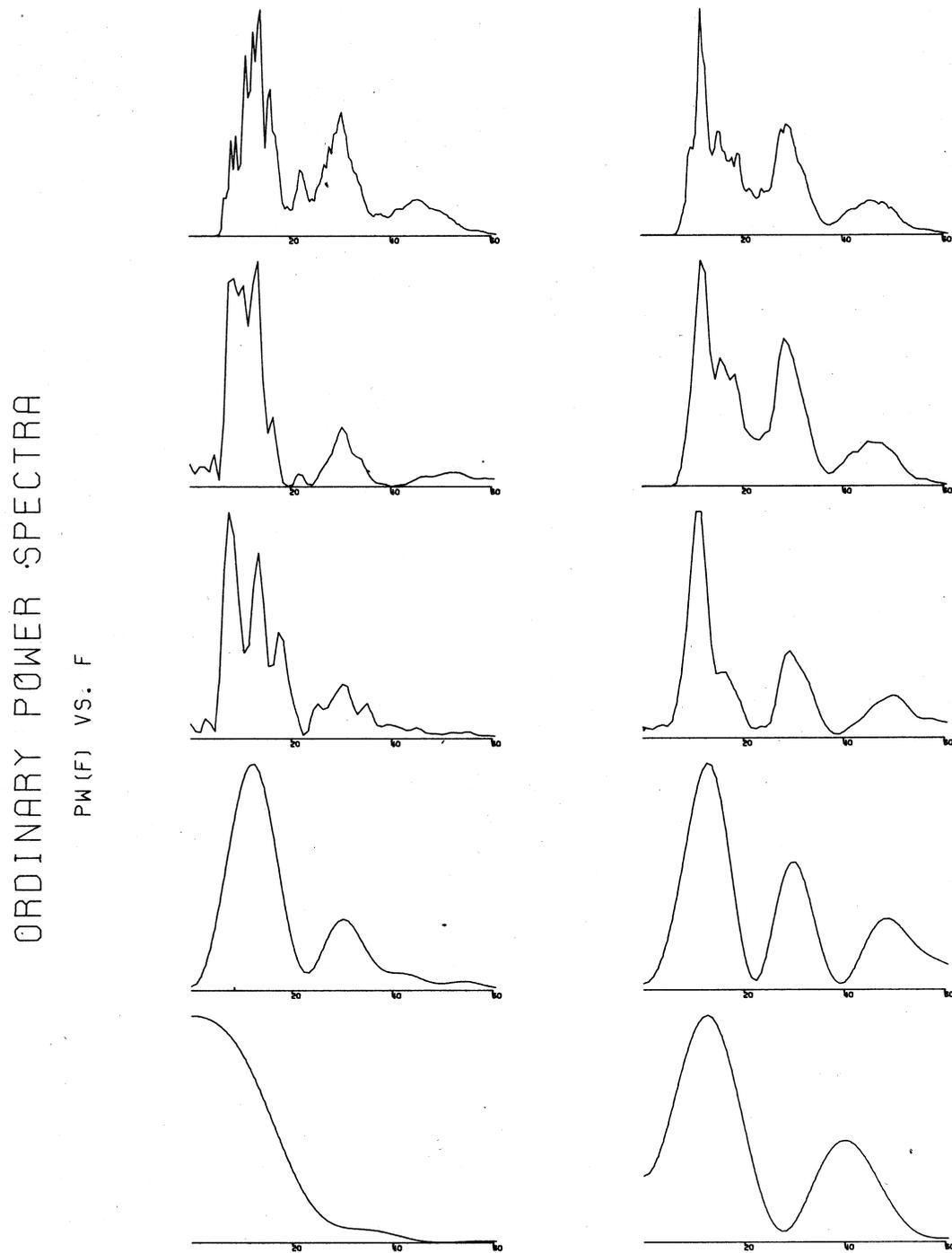


Figure 3: The ordinary power spectrum of the traces in Figure 2 as a function of frequency (0–60 Hz). Windows from top to bottom are (0.0, 2.0), (0.1, 0.5), (0.15, 0.45), (0.15, 0.25), and (0.175, 0.225) in seconds. [NR] jim2/. FIG3

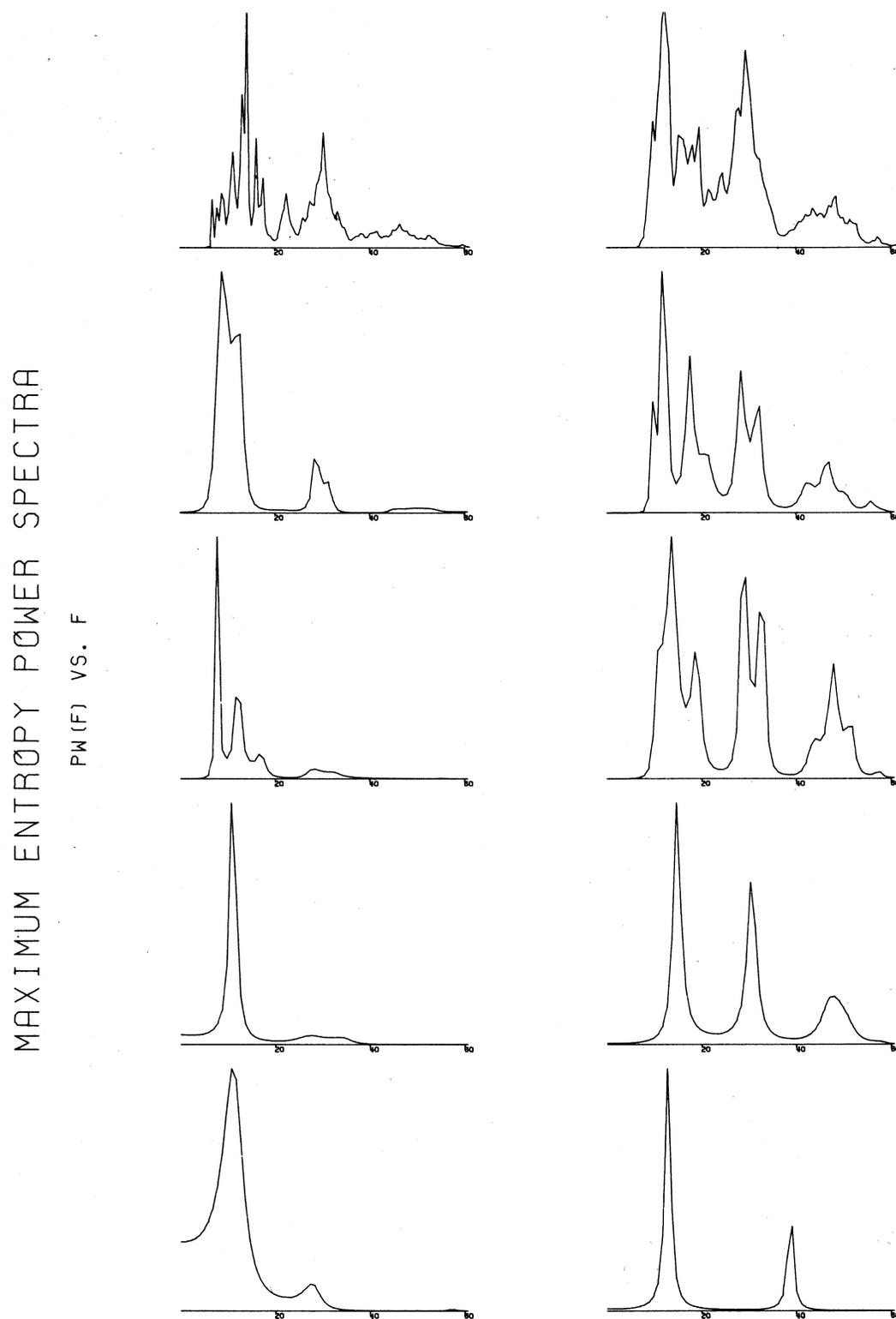


Figure 4: The maximum entropy power spectrum of the traces in Figure 2 as a function of frequency (0–60 Hz) with operator length $M = 2N/\ln(2N)$. Windows same as in Figure 3. Windows from top to bottom are (0.0, 2.0), (0.1, 0.5), (0.15, 0.45), (0.15, 0.25), and (0.175, 0.225) in seconds. [NR] jim2/. FIG4

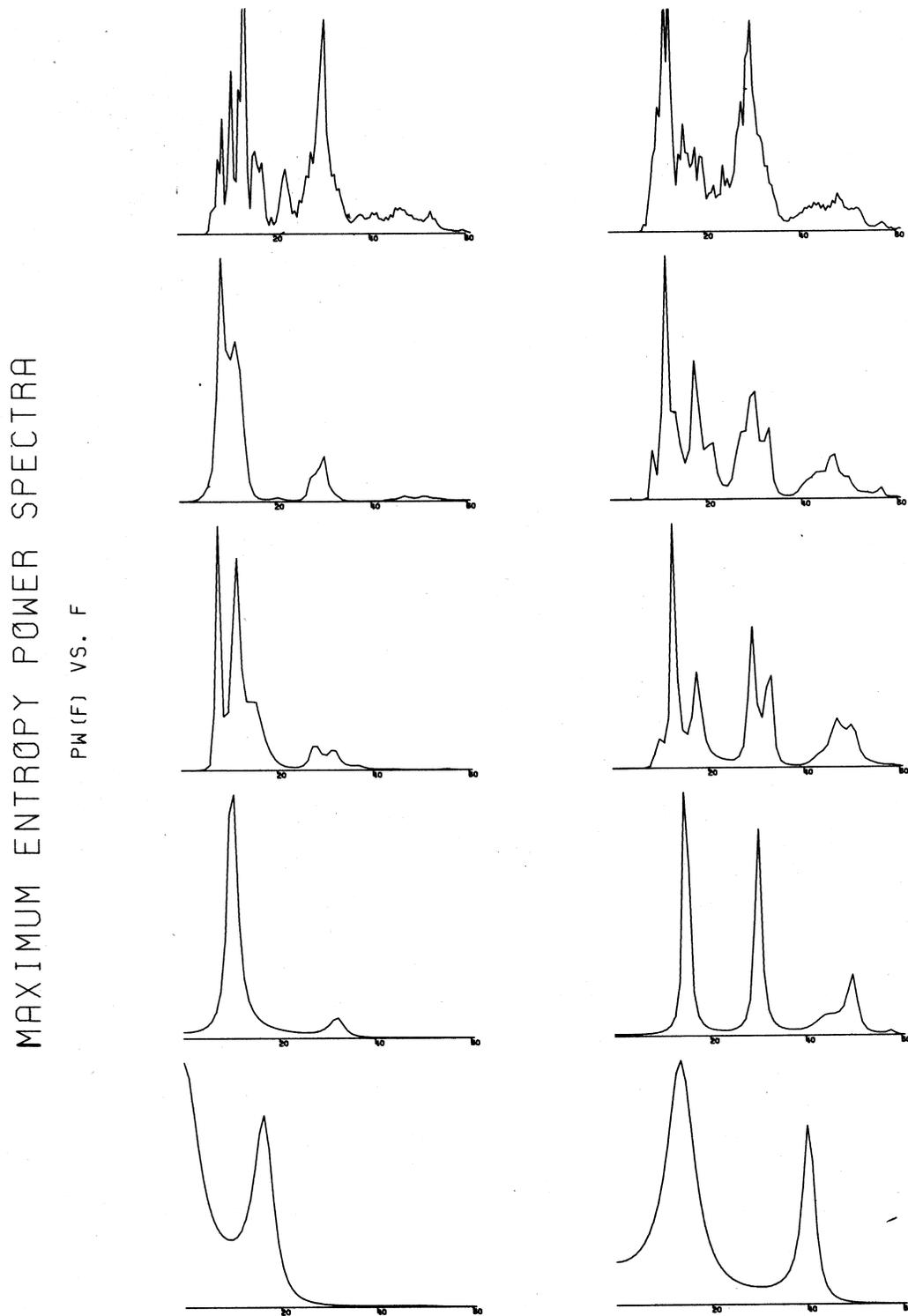


Figure 5: The maximum entropy power spectrum of the traces in Figure 2 as a function of frequency (0–60 Hz) with operator length $M = N/2$. Windows same as in Figure 3. Windows from top to bottom are (0.0, 2.0), (0.1, 0.5), (0.15, 0.45), (0.15, 0.25), and (0.175, 0.225) in seconds. [NR] jim2/. FIG5

MAXIMUM ENTROPY POWER SPECTRA
 PW(F) VS. F

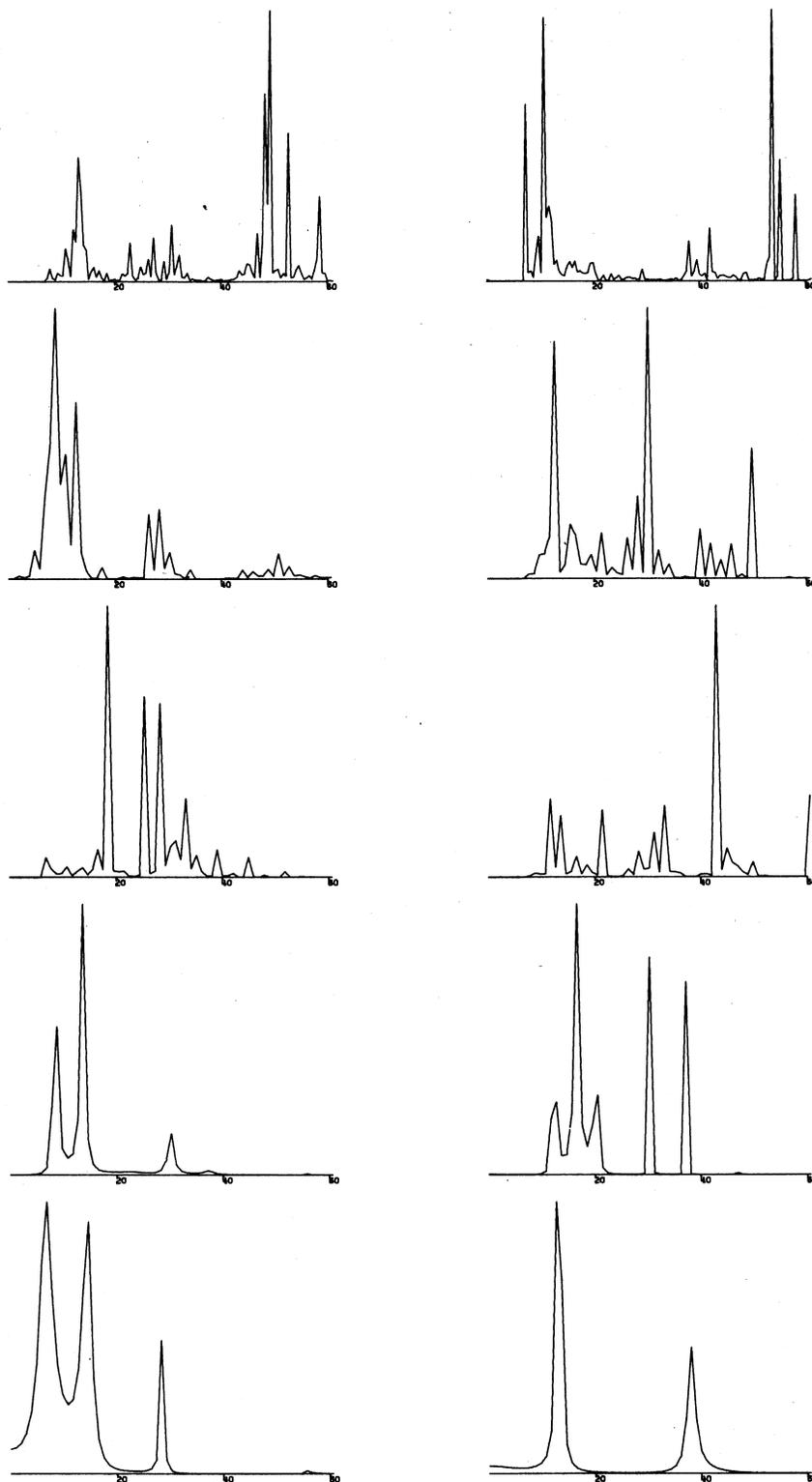


Figure 6: Same as Figure 5 with $M = N$. Windows from top to bottom are (0.0, 2.0), (0.1, 0.5), (0.15, 0.45), (0.15, 0.25), and (0.175, 0.225) in seconds. [NR] jim2/. FIG6

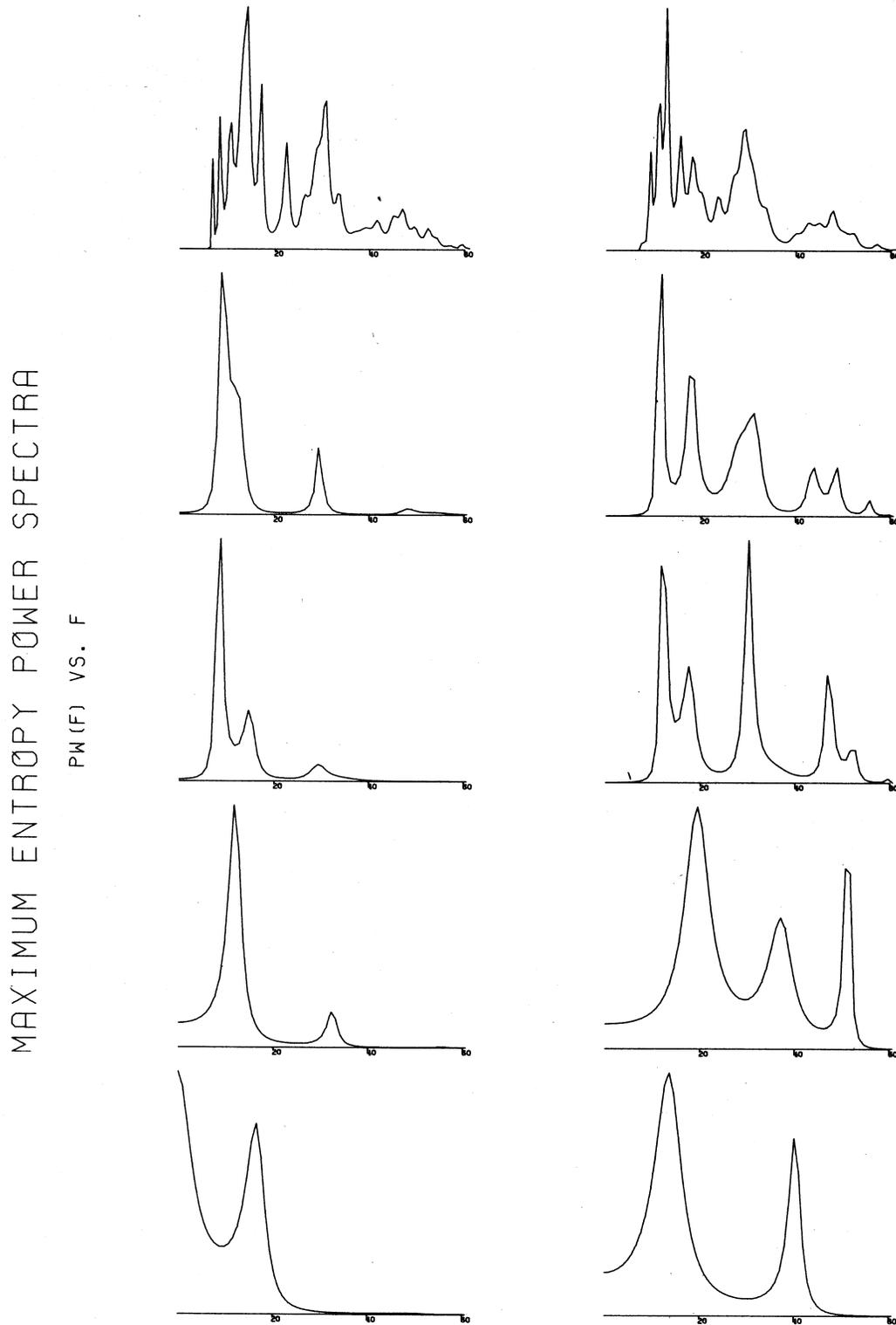


Figure 7: Same as Figure 5 with $M = N/\ln N$. Windows from top to bottom are (0.0, 2.0), (0.1, 0.5), (0.15, 0.45), (0.15, 0.25), and (0.175, 0.225) in seconds. [NR] jim2/. FIG7

itself be a good choice for the operator length, since it often gives values comparable to those of other methods without requiring performance parameter monitoring.

In conclusion, I recommend that anyone wishing to obtain high resolution power spectra for seismic traces should examine both maximum entropy spectra and ordinary spectra. When both methods give peaks approximately at the same frequencies, we may be confident that the maximum entropy method is giving higher resolution of true modes in the spectrum. If the MESA peaks are not reasonably close to the peaks obtained by the better understood Fourier transform method, we should give additional and careful consideration to the proper choice of the operator length for the MESA spectrum.

ACKNOWLEDGMENTS

I thank Robert H. Stolt for helpful discussions and Jerry A. Ware for encouragement. I also thank Bob Clapp for help with the preparation of the final figures.

APPENDIX A

The following derivation of the relationship between entropy and power spectrum is essentially the same as that given by Smylie et al. (1973). The derivation is included here for completeness.

The entropy of N random variables X_1, \dots, X_N is given by

$$H_N = - \int P(X_1, \dots, X_N) \ln[a^{2N} P(X_1, \dots, X_N)] d^N X = -2N \ln a - \int P \ln P d^N X, \quad (\text{A-1})$$

where P is the joint probability density and a is a constant with the same units as X . The power spectrum $P(f)$ computed from the autocorrelation values R_0, \dots, R_{N-1} depends only on the second-order statistics of the time series $\{X_n\}$. Therefore, the given time series cannot be distinguished from a normal (Gaussian) process.

The joint probability density for a normal process with N variables of zero mean is (using matrix notation, where X^T is the transpose of X)

$$P(X_1, \dots, X_N) = \left[(2\pi e)^N \det T_{N-1} \right]^{-\frac{1}{2}} \exp \left(-\frac{1}{2} X^T \cdot T_{N-1}^{-1} \cdot X \right), \quad (\text{A-2})$$

where T_{N-1} is the $N \times N$ Toeplitz matrix [see Claerbout (1976)] given by Equation (11) and X is the N -vector determined by $X^T = (X_1, \dots, X_N)$. Substituting (A-2) into (A-1), we find

$$H_N = \frac{1}{2} \ln \left[(2\pi e)^N \det T_{N-1} \right] - 2N \ln a. \quad (\text{A-3})$$

Setting the arbitrary constant $a = (2\pi e)^{1/4}$ for convenience, Equation (A-3) then becomes

$$H_N = \frac{1}{2} \ln (\det T_{N-1}). \quad (\text{A-4})$$

Since (A-4) necessarily diverges as $N \rightarrow \infty$, a better measure of the information content of the series is the average entropy per variable given by

$$h = \lim_{N \rightarrow \infty} \frac{H_N}{N} = \lim_{N \rightarrow \infty} \ln (\det T_{N-1})^{1/2N}. \quad (\text{A-5})$$

The eigenvalues $\{\lambda_1, \dots, \lambda_N\}$ of T_{N-1} are real and nonnegative since T is Hermitian and nonnegative definite. Furthermore,

$$\det T_{N-1} = \prod_{i=1}^N \lambda_i, \quad (\text{A-6})$$

so

$$h = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{i=1}^N \ln \lambda_i. \quad (\text{A-7})$$

The Szëgo theorem (Grenander and Szëgo, 1958; Smylie et al., 1973) states that, if F is any continuous function, then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N F(\lambda_i) = \frac{1}{2W} \int_{-W}^W F[2WP(f)] df, \quad (\text{A-8})$$

where, as before, W is the Nyquist frequency, $P(f)$ is the power spectrum, and the λ 's are the N eigenvalues of T_{N-1} .

Combining Equations (A-7) and (A-8), we find

$$h = \frac{1}{4W} \int_{-W}^W \ln[2WP(f)] df, \quad (\text{A-9})$$

which is the sought after result.

APPENDIX B

We need to compute the integral

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[\exp(i\theta) - Z_0] d\theta = \frac{P.V.}{2\pi i} \oint_{|Z|=1} \ln(Z - Z_0) \frac{dZ}{Z}, \quad (\text{B-1})$$

where *P.V.* stands for the principal value of the contour (complex) integral when the logarithm's branch cut is taken along the negative real axis.

First, note that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[\exp(i\theta) - Z_0] d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \exp(i\theta) d\theta + \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[1 - \exp(-i\theta)Z_0] d\theta. \quad (\text{B-2})$$

The first integral on the right is just

$$-\frac{1}{2\pi i} \int_{-\pi}^{\pi} \theta d\theta = 0, \quad (\text{B-3})$$

since the integrand is an odd function. When $|Z_0| < 1$, the integrand of the second integral on the right can be expanded in a convergent power series. Integrating term by term, we find that

$$-\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n=1}^{\infty} \frac{Z_0^n}{n} \exp(in\theta) d\theta = 0, \quad (\text{B-4})$$

since $\exp(in\pi) - \exp(-in\pi) = \cos(n\pi) + i \sin(n\pi) - \cos(-n\pi) - i \sin(-n\pi) = 0$ (the two cosines cancel and the two sines both vanish individually for all integer values of n). Thus, we find (B-1) is identically zero for all $|Z_0| < 1$. In particular, it vanishes when $Z_0 = 0$, so

$$\frac{P.V.}{2\pi i} \oint_{|Z|=1} \ln Z \times \frac{dZ}{Z} = 0. \quad (\text{B-5})$$

REFERENCES

- Akaike, H., 1970, Statistical predictor identification: *Ann. Inst. Statist. Math.*, **22**, 203–217.
- Berkhout, A. J. and P. R. Zaanen, 1976, Comparison between wiener filtering, kalman filtering, and deterministic least squares estimation: *Geophys. Prop.*, **24**, 141–197.
- Berryman, J. G., 1978, Choice of operator length for maximum entropy spectral analysis: *Geophysics*, **43**, 1385–1391.
- Burg, J. P., 1972, Relationship between maximum entropy spectra and maximum likelihood spectra: *Geophysics*, **37**, 375–376.
- , 1975, Maximum entropy spectral analysis: Ph.D. thesis, Stanford University, **SEP-6**.
- Chen, W. Y. and G. R. Stegen, 1974, Experiments with maximum entropy power spectra of sinusoids: *J. Geophys. Res.*, **79**, 3019–3022.
- Claerbout, J. F., 1976, *Fundamentals of geophysical data processing: With applications to petroleum prospecting*: McGraw-Hill, New York.
- Grenander, U. and G. Szögo, 1958, *Toeplitz forms and their applications*: University of California Press, Berkeley, CA, Section 5.2.
- Jenkins, G. M. and D. G. Watts, 1968, *Spectral analysis and its applications*: Holden-Day, San Francisco, California.
- Lacoss, R. T., 1971, Data adaptive spectral analysis methods: *Geophysics*, **36**, 661–675.
- Parzen, E., 1962, *Stochastic processes*: Holden-Day, San Francisco, CA.
- , 1964, An approach to empirical time series analysis: *J. Res. National Bureau Standards*, Section D, **86D**, 937–951.
- Smylie, D. E., G. K. C. Clarke, and T. J. Ulrych, 1973, Analysis of irregularities in the earth's rotation: *Methods in Computational Physics*, **13**, 391–430.
- Treitel, S., P. R. Gutowski, and E. A. Robinson, 1977, Empirical spectral analysis revisited: in *Topics in Numerical Analysis III*: Academic Press, New York.
- Ulrych, T. J. and T. N. Bishop, 1975, Maximum entropy spectral analysis and autoregressive decomposition: *Rev. Geophys. Space Phys.*, **13**, 183–200.
- Wiener, N., 1949, *Extrapolation, interpolation, and smoothing of stationary time series*: MIT Press and Wiley, New York.

Seismic anisotropy for polar media and an extended Thomsen formulation for longer offsets

James G. Berryman

ABSTRACT

Crack-influence parameters of Sayers and Kachanov (1991) have been shown to be directly related to Thomsen weak-anisotropy parameters for seismic wave speeds. These results are applied to the problem of seismic propagation in reservoirs having polar (HTI) symmetry due to aligned vertical fractures. To take full advantage of these relationships, it is also helpful to obtain more accurate expressions for seismic wave speeds in polar media at longer offsets than those originally intended for Thomsen's weak anisotropy formulation.

INTRODUCTION

Recently published work of the author (Berryman, 2007, 2008) shows how anisotropy due to cracks or large-scale fractures can be quantified using crack-influence parameters of Sayers and Kachanov (1991). Taking maximal advantage of this work requires use of exact or nearly exact formulas for anisotropic seismic wave speeds in field applications. Thomsen's weak anisotropy formulation (Thomsen, 1986), while accurate for short-offset data, is not adequate at longer offsets in strongly polar media. The point of the work summarized here is to arrive at a formulation almost as simple as Thomsen's, yet accurate enough to be useful at least to angles on the order of $\theta \simeq 45^\circ$.

EXACT RESULTS FOR VTI AND HTI MEDIA

If ρ is inertial mass density, and the coefficients c_{ij} are the usual elastic stiffnesses (for example, in isotropic media, $c_{11} = c_{22} = c_{33} = \lambda + 2\mu$, $c_{44} = c_{55} = c_{66} = \mu$, $c_{12} = c_{13} = c_{23} = \lambda$ with λ and μ being the well-known Lamé elastic parameters), then the exact results for VTI media with x_3 -axis of symmetry (so $c_{11} = c_{22} \neq c_{33}$, $c_{13} = c_{23} \neq c_{12}$, $c_{44} = c_{55} \neq c_{66}$, and $c_{12} = c_{11} - 2c_{66}$) are given exactly by:

$$v_p^2(\theta) = \frac{1}{2\rho} \{c_{44} + [c_{11} \sin^2 \theta + c_{33} \cos^2 \theta] + R(\theta)\} \quad (1)$$

and

$$v_{sv}^2(\theta) = \frac{1}{2\rho} \{c_{44} + [c_{11} \sin^2 \theta + c_{33} \cos^2 \theta] - R(\theta)\}, \quad (2)$$

where

$$R^2(\theta) = \left[(c_{11} - c_{44}) \sin^2 \theta - (c_{33} - c_{44}) \cos^2 \theta \right]^2 + 4(c_{13} + c_{44})^2 \sin^2 \theta \cos^2 \theta, \quad (3)$$

and where polar angle θ is measured from the vertical x_3 -axis directed into the earth. Results for horizontally polarized shear waves in VTI media will not be treated in the following discussion. Results for HTI media (for example, from aligned vertical cracks) are obtained directly by substituting $\theta_H = \theta - \pi/2$ for any vertical plane that is also perpendicular to the vertical plane of symmetry.

Thomsen's anisotropy parameters ϵ , δ , γ can be related directly to the Sayers and Kachanov (1991) crack-influence parameters η_1 and η_2 for horizontal cracks according to $\epsilon_h = \frac{c_{11}-c_{33}}{2c_{33}} = \rho_c[(1+\nu_0)\eta_1 + \eta_2] \frac{E_0}{1-\nu_0^2} \eta_2 G_0 \simeq \frac{2\rho_c \eta_2 G_0}{1-\nu_0} = \delta_h$, and $\gamma_h = \frac{c_{66}-c_{44}}{2c_{44}} = \rho_c \eta_2 G_0$, where Poisson's ratio ν_0 , Young's modulus E_0 , and shear modulus G_0 are the values for the assumed isotropic background medium. For the penny-shaped cracks (having penny radius a) considered (see TABLE 1), Thomsen's parameters ϵ and δ are always found to be equal to each other to lowest order in the crack density $\rho_c = na^3$, with $n = N/V$ being the crack number density.

Crack Parameters	NI Approx.	1st Model $\nu_0 = 0.00$	2nd Model $\nu_0 = 0.4375$
η_1 (GPa ⁻¹)	$-\frac{\nu_0 \eta_2}{2(5-\nu_0)}$	0.0000	-0.0192
η_2 (GPa ⁻¹)	$\frac{8(1-\nu_0)(5-\nu_0)}{15(2-\nu_0)G_0}$	0.1941	0.3994

Table 1: Examples of the first crack-influence parameters (the two for lowest order in powers of ρ_c) estimated by Berryman and Grechka (2006), from simulations of Grechka.

EXTENDING THOMSEN'S FORMULAS TO LARGER OFFSETS FOR VTI AND HTI SYMMETRY

The most obvious problem with Thomsen's approximations to the wave speeds generally occurs in $v_{sv}(\theta)$. As noted in earlier work, the key issue is that Thomsen's approximation for $v_{sv}(\theta)$ is completely symmetric around $\theta = \pi/4 = 45^\circ$, while — unfortunately — this is usually not true of the actual wave speeds $v_{sv}(\theta)$. This inherent error may seem innocuous in itself since it is not immediately clear whether it affects the results for small angles of incidence ($< 15^\circ$) or not, but this inaccuracy clearly does lead to large over- or under-estimates of wave speeds in the neighborhood of both the extreme value (*i.e.*, a peak or a trough) located at $\theta = \theta_{ex}$ and also in the neighborhood of $\theta = 45^\circ \neq \theta_{ex}$. So these discrepancies can certainly become issues at offsets larger than the original design criterion having polar angles $\theta \leq 15^\circ$.

To improve this situation while still making use of a simple and practical approximation to the phase speed, we reconsider an approach originally proposed by Berryman (1979). In particular, notice that the square root of $R^2(\theta)$ in Eq. (3) can be exactly and conveniently rewritten as:

$$R(\theta) = [c_{11} \sin^2 \theta + c_{33} \cos^2 \theta - c_{44}] \sqrt{1 - \zeta(\theta)}, \quad (4)$$

where

$$\zeta(\theta) \equiv \frac{[(c_{11} - c_{44})(c_{33} - c_{44}) - (c_{13} + c_{44})^2] \sin^2 2\theta}{[c_{11} \sin^2 \theta + c_{33} \cos^2 \theta - c_{44}]^2}. \quad (5)$$

To simplify this expression, first notice that $\zeta(\theta)$ has an absolute maximum (or minimum) value, which occurs when θ takes the value θ_m determined by

$$\tan^2 \theta_m = \frac{c_{33} - c_{44}}{c_{11} - c_{44}} \equiv 1 - x_m, \quad (6)$$

where

$$x_m = \frac{c_{11} - c_{33}}{c_{11} - c_{44}} > 0. \quad (7)$$

The inequality in (7) is true for VTI media having horizontal fractures, since for this case the stiffness difference $c_{11} - c_{33}$ is known to be positive (as it is also known that ρ_c , η_2 , and G_0 are all positive, while Poisson's ratio satisfies $\nu_0 \leq \frac{1}{2}$).

Then, the extreme value $\zeta_m \equiv \zeta(\theta_m)$ is itself given by

$$\begin{aligned} \zeta_m &= 1 - \frac{(c_{13} + c_{44})^2}{(c_{11} - c_{44})(c_{33} - c_{44})} \\ &= \frac{2(\epsilon - \delta)c_{33}}{c_{11} - c_{44}} \\ &= \frac{2(\epsilon - \delta)v_p^2(0)}{v_p^2(0)(1 + 2\epsilon) - v_s^2(0)}, \end{aligned} \quad (8)$$

where the second and third expressions in (8) relate ζ_m to the difference between the Thomsen parameters ϵ and δ , and also to $v_p(0)$ and $v_s(0)$. In general, $(\epsilon - \delta)$ — and therefore ζ_m — can take values positive, negative, or zero. Furthermore, $\zeta(\theta)$ may be rewritten as

$$\zeta(\theta) = \frac{2\zeta_m}{1 + \chi(\theta)}, \quad (9)$$

where

$$\chi(\theta) = \frac{1}{2} \left[\frac{\tan^2 \theta}{\tan^2 \theta_m} + \frac{\tan^2 \theta_m}{\tan^2 \theta} \right]. \quad (10)$$

It is always true that $\zeta(\theta) \leq 1$. [Note that $\zeta_m \geq 0$ for all layered media since $\epsilon - \delta \geq 0$ for layered elastic media (Postma, 1955; Backus, 1962; Berryman, 1979). However, such a simple constraint is *not* known for other types of anisotropic systems.] The square root in equation (4), can be expanded to first order as

$$\sqrt{1 - \zeta(\theta)} \simeq 1 - \frac{\zeta(\theta)}{2} = 1 - \frac{\zeta_m}{1 + \chi(\theta)}. \quad (11)$$

Approximate results for $v_p(\theta)$ and $v_{sv}(\theta)$ are therefore:

$$v_p^2(\theta) \simeq \frac{1}{\rho} \left([c_{11} \sin^2 \theta + c_{33} \cos^2 \theta] - \frac{\zeta_m [(c_{11} - c_{44}) \sin^2 \theta + (c_{33} - c_{44}) \cos^2 \theta]}{2[1 + \chi(\theta)]} \right) \quad (12)$$

and

$$v_{sv}^2(\theta) \simeq \frac{1}{\rho} \left(c_{44} + \frac{\zeta_m [(c_{11} - c_{44}) \sin^2 \theta + (c_{33} - c_{44}) \cos^2 \theta]}{2[1 + \chi(\theta)]} \right). \quad (13)$$

The only approximation made in arriving at equations (12) and (13) was the approximation of the square root shown in (11).

Although this simple approach is the one most commonly used, the analysis presented is not really limited to using only the first order Taylor approximation in (11). Other researchers (Fowler, 2003; Pederson et al., 2007) have explored rational approximations to such square roots, but we now choose to take a rather different approach.

Progress is made by noting that the quantity $\frac{1}{2}[1 + \chi(\theta)]$ may be rewritten as:

$$\frac{1}{2}[1 + \chi(\theta)] = \frac{1}{4} \left(\frac{\tan \theta}{\tan \theta_m} + \frac{\tan \theta_m}{\tan \theta} \right)^2 = \frac{(\tan^2 \theta + \tan^2 \theta_m)^2}{4 \tan^2 \theta \tan^2 \theta_m}. \quad (14)$$

To simplify this expression, first multiply numerator and denominator of (14) by $\cos^4 \theta \cos^4 \theta_m$. The denominator of the result is then proportional to $\sin^2 2\theta \sin^2 2\theta_m$, while the numerator is now proportional to the square of the quantity $\cos^2 \theta \cos^2 \theta_m \times (\tan^2 \theta + \tan^2 \theta_m) = \sin^2 \theta \cos^2 \theta_m + \sin^2 \theta_m \cos^2 \theta = \frac{1}{2}(1 - \cos 2\theta \cos 2\theta_m)$. Combining these two results gives

$$\zeta(\theta) = \frac{\zeta_m \sin^2 2\theta_m \sin^2 2\theta}{[1 - \cos 2\theta_m \cos 2\theta]^2}, \quad (15)$$

which (although this may not be immediately obvious) is just a more compact version of (5). Equation (15) is exact; no approximations were made in the transition from (5) to (15).

Dividing the expression in the text before Eq. (15) by $2 \cos^2 \theta_m$, we also have

$$\sin^2 \theta + \tan^2 \theta_m \cos^2 \theta = \frac{[1 - \cos 2\theta_m \cos 2\theta]}{2 \cos^2 \theta_m}. \quad (16)$$

Using the definitions $X_{\pm} \equiv [1 \pm \sqrt{1 - \zeta(\theta)}]$, the exact expression for quasi-SV-wave speed can now be rewritten as

$$2\rho v_{sv}^2 = 2c_{44} + (c_{11} - c_{44})(\sin^2 \theta + \tan^2 \theta_m \cos^2 \theta) X_-. \quad (17)$$

Similarly, the exact equation for quasi-P-wave speed becomes

$$2\rho v_p^2 = 2c_{44} + (c_{11} - c_{44})(\sin^2 \theta + \tan^2 \theta_m \cos^2 \theta) X_+. \quad (18)$$

Both of these expressions are exact rearrangements of the original equations.

These results can be consolidated further by using the result (16), together with the definition

$$c_{11} + c_{33} - 2c_{44} = \frac{c_{11} - c_{44}}{\cos^2 \theta_m} \equiv 2\Delta c. \quad (19)$$

So finally, we have two very compact expressions for the exact wave speeds:

$$\rho v_{sv}^2 = c_{44} + \frac{\Delta c}{2} [1 - \cos 2\theta_m \cos 2\theta] [1 - \sqrt{1 - \zeta(\theta)}] \quad (20)$$

for the quasi-SV-wave speed, and also the corresponding equation which is

$$\rho v_p^2 = c_{44} + \frac{\Delta c}{2} [1 - \cos 2\theta_m \cos 2\theta] [1 + \sqrt{1 - \zeta(\theta)}] \quad (21)$$

for the exact quasi-P-wave speed.

Extended formulas for HTI symmetry

To complete this analysis, we show how the results for HTI symmetry arise when the fractures/cracks are aligned and vertical. These results follow easily from our earlier work:

$$\begin{aligned} \bar{v}_p^2(\theta_H)/v_p^2(0)(1+2\epsilon) &\simeq 1 - \frac{2\epsilon}{1+2\epsilon} \sin^2 \theta_H \\ &- \frac{\epsilon-\delta}{2(1+2\epsilon)} \frac{2 \sin^2 \theta_m \sin^2 2\theta_H}{[1+\cos 2\theta_m \cos 2\theta_H]}, \end{aligned} \quad (22)$$

from which we find:

$$\begin{aligned} \bar{v}_p(\theta_H)/v_p(0)\sqrt{1+2\epsilon} &\approx 1 - \frac{\epsilon}{1+2\epsilon} \sin^2 \theta_H \\ &- \frac{\epsilon-\delta}{4(1+2\epsilon)} \frac{2 \sin^2 \theta_m \sin^2 2\theta_H}{[1+\cos 2\theta_m \cos 2\theta_H]}. \end{aligned} \quad (23)$$

Similarly,

$$\begin{aligned} \bar{v}_{sv}^2(\theta_H)/v_s^2(0) &\simeq 1 + \frac{[v_p^2(0)/v_s^2(0)]}{2} \\ &\times \frac{\epsilon-\delta}{2} \frac{2 \sin^2 \theta_m \sin^2 2\theta_H}{[1+\cos 2\theta_m \cos 2\theta_H]}, \end{aligned} \quad (24)$$

from which follows:

$$\begin{aligned} \bar{v}_{sv}(\theta_H)/v_s(0) &\approx 1 + \frac{[v_p^2(0)/v_s^2(0)]}{4} \\ &\times \frac{\epsilon-\delta}{4} \frac{2 \sin^2 \theta_m \sin^2 2\theta_H}{[1+\cos 2\theta_m \cos 2\theta_H]}. \end{aligned} \quad (25)$$

Again, these formulas reduce exactly to the equivalent Thomsen formulas for HTI symmetry if $\theta_m \rightarrow 45^\circ$. It should always be remembered, however, that these formulas apply only in planes perpendicular to the plane of the aligned fractures. For other angles of propagation, we must also account for the azimuthal dependence on angle ϕ , although this is in fact easy to do. Examples of these results are presented in the three Figures.

DISCUSSION AND CONCLUSIONS

The preceding analysis presented for the phase velocity equations does not depend on the source of the anisotropy, and therefore can be applied to layered media, etc., as well as to fractured media as we have done here.

The Kachanov (1980) and Sayers and Kachanov (1991) crack-influence parameters are ideally suited to analyzing the role of fracture mechanics in producing anisotropic elastic constants for aligned fractures in a reservoir exhibiting VTI or HTI symmetry. When this approach is combined with poroelastic analysis through the use of Skempton's coefficient [Skempton (1954), see also Berryman (2007)], it becomes comparatively easy to analyze a wide range of complicated situations that may raise in reservoir analysis, such as trying to deduce whether the fractures are dry/drained, or fluid-saturated/undrained. Skempton's coefficient B introduces a single parameter that varies from 0 to 1 as fluid properties change from being negligible to being very strong influences on the fracture compliance – and therefore on the Thomsen seismic parameters.

Another important observation from the modeling presented is that the Thomsen weak anisotropy formulation is valid for crack densities up to about $\rho_c \simeq 0.05$, but should be replaced by more accurate approximations, or (better yet) exact calculations whenever possible if the crack density is much above 0.05. When the crack density is $\rho_c \simeq 0.1$, or higher, then higher accuracy approximations are essential. Conversely, if the crack density

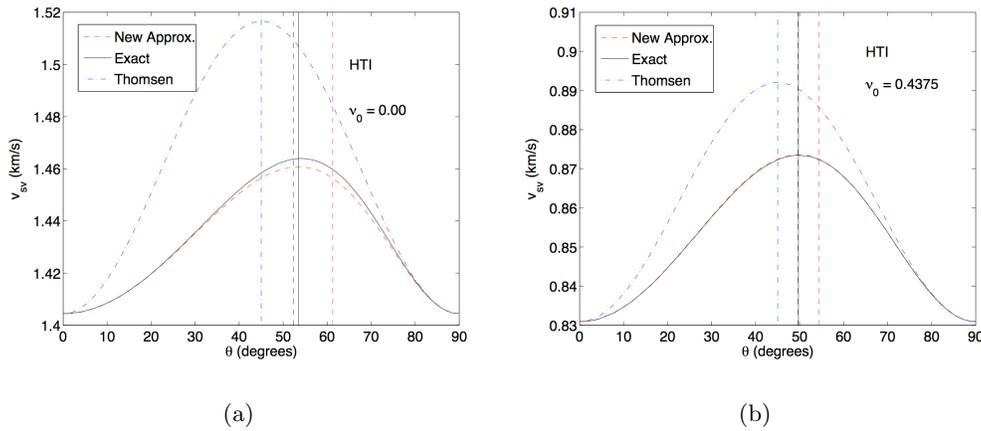


Figure 1: For aligned vertical cracks having crack density $\rho_c = 0.2$ and HTI overall symmetry: two examples of anisotropic quasi-SV shear wave speeds for two values of Poisson's ratio ν_0 of the host medium: top $\nu_0 = 0.00$, bottom $\nu_0 = 0.4375$. Background wave speeds are top $v_s = 1.77$ km/s, bottom $v_s = 1.00$ km/s. Speeds in red are those for the new approximation. The corresponding exact result is then overlain in black. Finally, Thomsen's weak anisotropy curves are overlain in blue. Two of the vertical lines indicate locations of the true peaks of these curves: Thomsen's approximation is the dashed blue line and it is always at 45° . The solid black line is the true peak of the exact expression. The value of θ_m from equation (6) is the vertical line shown in red. The other two dashed black vertical lines are two estimators (not treated here) of the peak of our approximate curve; these two estimators are often nearly indistinguishable. [NR] jim3/. FIG11,FIG12

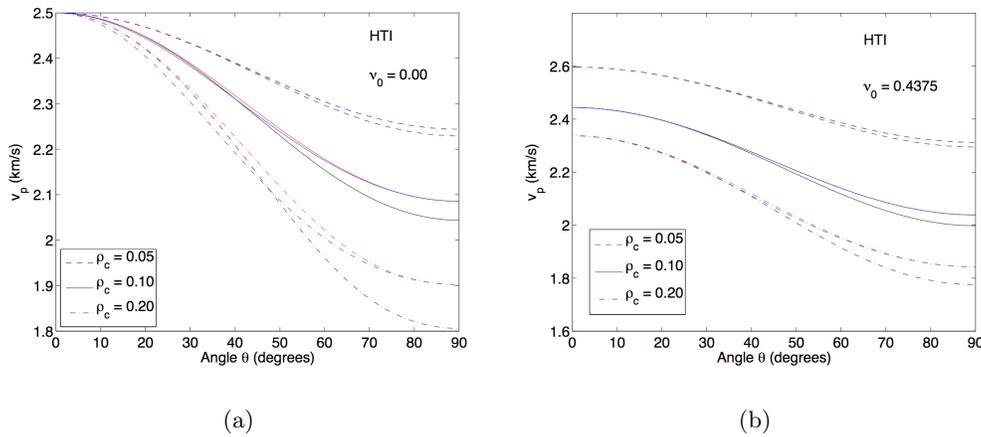


Figure 2: Quasi-P-wave speeds for the same two model reservoirs, for three choices of crack density $\rho_c = 0.05, 0.10, 0.20$. Background P -wave speed for the first model is $v_p = 2.50$ km/s and for the second model $v_p = 3.00$ km/s. As before, speeds in red are those for the new approximation. The corresponding exact result is then overlain in black. Finally, Thomsen's weak anisotropy curves are overlain in blue. [NR] jim3/. FIG7,FIG9

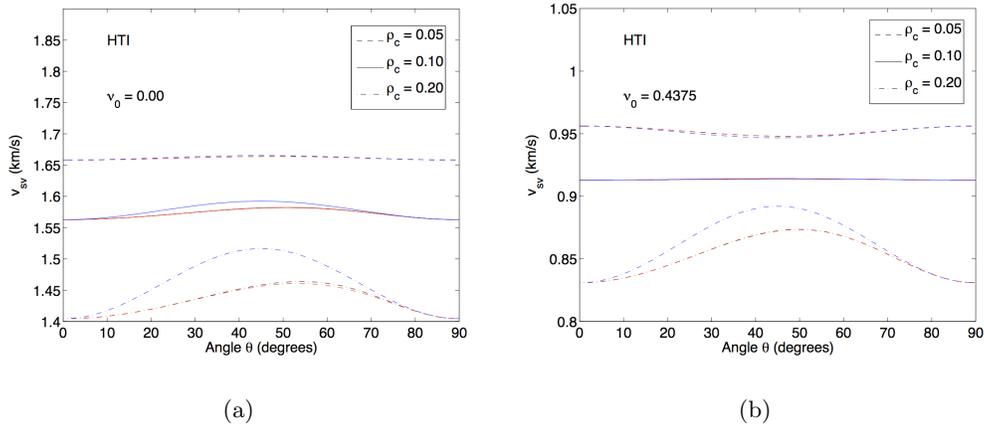


Figure 3: Quasi-SV shear-wave speeds for the same two model reservoirs, for three choices of crack density $\rho_c = 0.05, 0.10, 0.20$. The two values of Poisson's ratio ν_0 of the host medium considered are: top $\nu_0 = 0.00$, bottom $\nu_0 = 0.4375$. Background shear-wave speed for the first model is $v_s = 1.77$ km/s and for the second model $v_s = 1.00$ km/s. As before, speeds in red are those for the new approximation. The corresponding exact result is then overlain in black. Finally, Thomsen's weak anisotropy curves are overlain in blue. Plots for $v_{sh}(\theta)$ are not dependent on the formulation of the new scheme and, therefore, are not displayed. [NR] jim3/. FIG8,FIG10

ρ_c estimated from seismic data using the weak anisotropy formulation is in fact larger than $\rho_c \simeq 0.05$, one conclusion we might reach is that a more accurate method is required both to verify and properly quantify the result.

ACKNOWLEDGMENTS

The author thanks V. Grechka, M. Kachanov, S. Nakagawa, S. R. Pride, and I. Tsvankin for helpful collaborations and conversations.

REFERENCES

- Backus, G. E., 1962, Long-wave elastic anisotropy produced by horizontal layering: *Journal of Geophysical Research*, **67**, 4427–4440.
- Berryman, J. G., 1979, Long-wave elastic anisotropy in transversely isotropic media: *Geophysics*, **44**, 896–917.
- , 2007, Seismic waves in rocks with fluids and fractures: *Geophysical Journal International*, **171**, 954–974.
- , 2008, Exact seismic velocities for transversely isotropic media and extended thomsen formulas for stronger anisotropies: *Geophysics*, **73**, D1–D10.
- Berryman, J. G. and V. Grechka, 2006, Random polycrystals of grains containing cracks: Model of quasistatic elastic behavior for fractured systems: *Journal of Applied Physics*, **100**, 113527.

- Fowler, P., 2003, Practical VTI approximations: A systematic anatomy: *Journal of Applied Geophysics*, **54**, 347–367.
- Kachanov, M., 1980, Continuum model of medium with cracks: *ASCE Journal of Engineering Mech.*, **106**, 1039–1051.
- Pederson, O., B. Ursin, and A. Stovas, 2007, Wide-angle phase-slowness approximations: *Geophysics*, **72**, S177–S185.
- Postma, G. W., 1955, Wave propagation in a stratified medium: *Geophysics*, **20**, 780–806.
- Sayers, C. M. and M. Kachanov, 1991, A simple technique for finding effective elastic constants of cracked solids for arbitrary crack orientation statistics: *International Journal of Solids and Structures*, **27**, 671–680.
- Skempton, A. W., 1954, The pore-pressure coefficients a and b : *Géotechnique*, **4**, 143–147.
- Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966.

Seismic reflector characterization by a multiscale detection-estimation method

Mohammad Maysami

ABSTRACT

Seismic reflector interfaces in the subsurface are typically idealized as zero-order discontinuities. According to this model, the Earth's subsurface is represented by a spiky reflection coefficient sequence for which deconvolution methods have been derived. However, multiscale analysis on sedimentary basins reveals the existence of accumulation of varying-order singularities in the subsurface. Seismic traces with varying-order singularities violate most of the assumptions underlying conventional deconvolution methods. We avoid this problem by proposing a nonlinear parametric method based on a two-step approach that divides the initial problem of delineating and characterizing transitions over the whole seismic signal into two easier subproblems. The waveform characterization implies a representation for the geological record that provide information on the basic subsurface stratigraphy and lithology.

INTRODUCTION

The Earth's subsurface consists of layers of different materials separated by interfaces, also called transitions. Transitions are characteristic of regions where acoustic properties of the Earth vary rapidly compared to the length-scale of the seismic source wavelet. Extracting information about the accurate locations and nature of transitions from seismic data recently has received increasing interest, as it provides geologists with quantitative and complementary information useful in various geophysical applications, ranging from improving geological interpretation to detecting lithology changes.

The aim of seismic deconvolution is to find the locations of seismic reflectors based on certain assumptions on the reflectivity. Robinson (1957) proposed the first method to solve the blind deconvolution problem with an unknown source function. Other approaches are based on the sparsity assumption, where the reflectivity is considered to be given by a sparse spike train (see e.g. Dossal and Mallat (2005) and the references therein).

Multiscale studies of seismic and well data has established the diversity of transitions in the subsurface (Herrmann, 1998; Herrmann et al., 2001). These results reveal the existence of transitions with varying singularity orders in the subsurface. This generalized type of fractional-order singularity is given by fractionally differentiating/integrating a zero-order discontinuity (see Figure 1). Consequently, valuable information is lost when assuming a reflectivity sequence consists of only zero- or first-order singularities. Despite recent developments (Saggaf and Robinson, 2000), the existence of different fractional-order transitions renders conventional deconvolution techniques ineffective. Other methods to extract the singularity orders from seismic data based on wavelet coefficient decay (Liner et al., 2004) yield ambiguous estimates since seismic data is bandlimited. These observations led to a new

type of parametrization, where the observed reflectivity is a superposition of parametrized waveforms. Maysami and Herrmann (2007) proposed a new detection-estimation method (see also Maysami, 2008) to estimate parameters that provide information on the transition sharpness that is related to the lithology (Herrmann et al., 2001; Liner et al., 2004).

The report is organized as follows. First, we present the parametric reflector model. Next we review the individual steps of our new method, i.e., detection, partitioning, and estimation. We also apply this method to a synthetic seismic trace and show the output of each step.

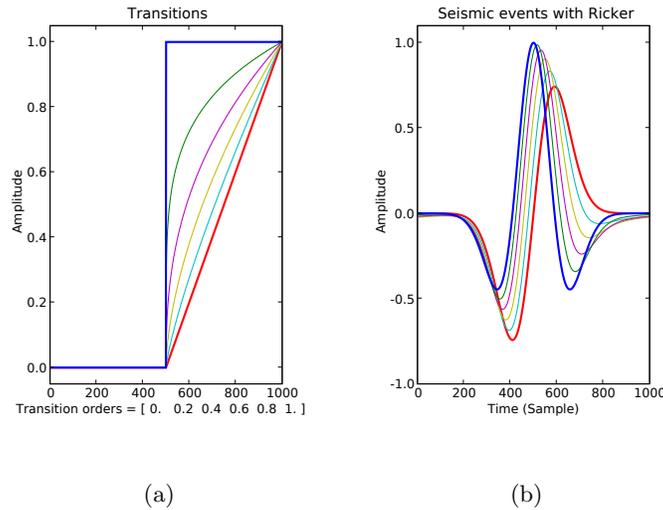


Figure 1: Generalized transition model with singularity orders varying from zero- to first-order. **(a)** Zero- (step), first- (ramp), and fractional-order transitions. **(b)** Corresponding seismic waveforms, yielded by the convolution of the reflectivity with a Ricker wavelet.

[NR] mohammad1/. transitions,waveforms

THE EARTH'S MODEL

We represent a vertical 1-D profile of the Earth, either a time series or depth profile, as a superposition of parametrized waveforms. Without loss of generality, we use the time series notation represented by

$$s(t) = \sum_{i \in I} c_i \psi_{\sigma_i}^{\alpha_i}(t - \tau_i) e^{j\pi\phi_i}, \quad (1)$$

where t represents time, and c_i and ϕ_i ($0 \leq \phi < 2$) are the amplitude and the phase for the i^{th} transition, respectively. Furthermore, $\psi_{\sigma_i}^{\alpha_i}(t - \tau_i)$ is a translated (τ_i) and scaled (σ_i) source wavelet ψ , which is fractionally differentiated ($\alpha < 0$) or integrated ($\alpha > 0$). The parameters of interest in this case, hereafter referred to as attributes, are the amplitude c_i , the location τ_i , the scale σ_i , the singularity order α_i , and the phase ϕ_i . The location (offset) of the waveforms corresponds to subsurface transition depths (stratigraphy), whereas their fractional-order constitutes a measure of transition sharpness (lithology).

THE CHARACTERIZATION PROBLEM

Given the above signal presentation, Maysami and Herrmann (2007) proposed a new method to estimate attributes of major transitions in seismic data (see also Maysami, 2008). They divided the initial problem of delineating and characterizing transitions over the seismic signal into two easier steps: detection and estimation. The first stage locates the major components of the seismic signal and segments it into individual waveforms. The second step estimates attributes of single windowed waveforms by using a nonlinear parametric inversion. Since our problem does not fit into the classical deconvolution framework, we use a multiscale wavelet technique to locate the main events. After segmentation, we input the individual waveforms to a nonlinear inversion procedure to estimate the attributes. This procedure uses rough estimates for the location, scale and phase from the detection stage as initial guesses.

Detection This part aims to approximate stratigraphic information of seismic signal regardless of nature of the transitions. The variety of different orders of transitions in the subsurface calls for a seismic-event-detection technique that does not make any assumptions regarding the transition type. Edge detection based on the multiscale continuous (complex) wavelet transform modulus maxima (Mallat, 1997) offers an approach that is robust for different waveforms reflecting different transition types. First, the method calculates the forward wavelet transform of a seismic trace, s , as a convolution product

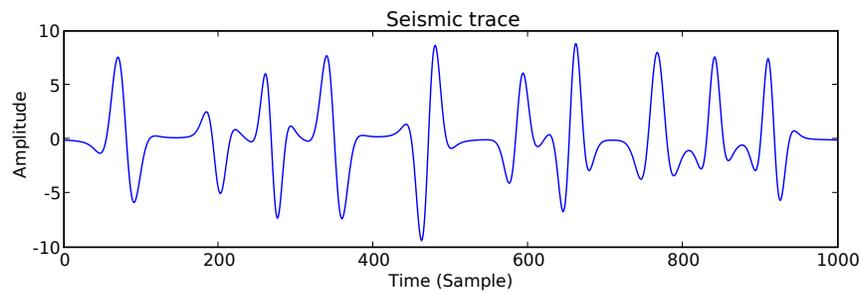
$$\mathcal{W}s(t, \sigma) = (s * \bar{\psi}_\sigma)(t), \quad (2)$$

where $\bar{\psi}_\sigma(t) = \frac{1}{\sqrt{\sigma}}\psi^*\left(\frac{-t}{\sigma}\right)$, and $\sigma \geq 0$ is the scale of wavelet ψ . Here, the symbol $*$ denotes the complex conjugate. The range of scales σ for the wavelet is adapted to the seismic source function. After forming the modulus maxima lines (MML) from the wavelet coefficients (Mallat, 1997), the maximum points along these lines are calculated, yielding rough estimates for the scale (i.e., bandwidth) and position of the reflection events (see Figure 2). The result of this stage is a set of locations, scales, and phase rotations $\{\tau^{(n)}, \sigma^{(n)}, \phi^{(n)}\}$ with $n = 1, N$, and N is the number of detected maxima, which corresponds to location and scale. We use these approximated values as initial guesses in the nonlinear inversion during the estimation stage.

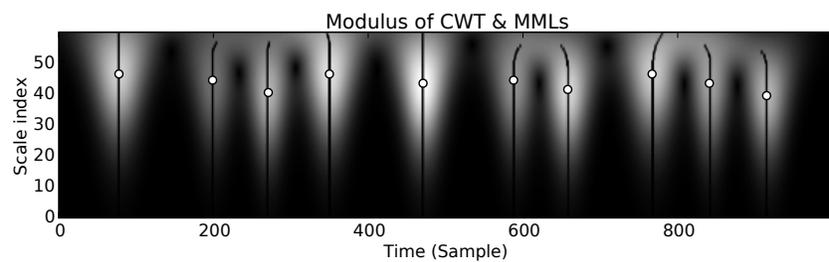
Partitioning Given estimates of the location and scale of the detected events, the trace is segmented into separate events. We extract the n^{th} detected waveform by multiplying the seismic trace by a window function centered at $\tau^{(n)}$ and with a support proportional to $\sigma^{(n)}$ (see Figure 3). The isolated waveforms are given by

$$s^{(n)}(t) = \mathbf{W}[\tau^{(n)}, \sigma^{(n)}]s(t) \quad \text{with } n = 1, N, \quad (3)$$

where $\mathbf{W}[\cdot]$ is the windowing operator. This procedure outputs N vectors containing ‘isolated’ events. Even though this segmentation procedure is somewhat arbitrary, e.g., it depends on a width parameter, we found this method to perform reasonably well for most cases.



(a)



(b)

Figure 2: A typical example for detection stage of a synthetic seismic trace **(a)**. Wavelet coefficients for the signal are plotted in **(b)** with dark colors corresponding to small magnitudes. The vertical and horizontal axes show scale and location, respectively. Modulus maxima lines are shown as dark lines where white circles identify the scale and the location for the corresponding events. [NR] `mohammad1/. synt-trace,mml`

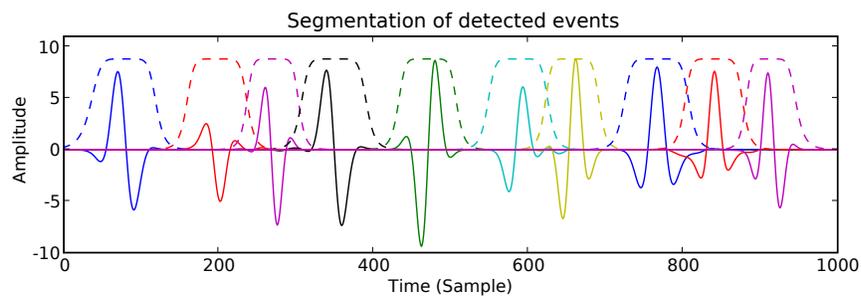


Figure 3: Segmentation of detected events. Each individual event (solid waveform) is extracted by using a window function (dashed line). [NR] `mohammad1/. segment`

Estimation To complete the characterization, we apply a parametric nonlinear inversion procedure to the segmented events. In order to set up this procedure, we first need to refine our mathematical model for the parametrized waveforms in equation 1. We derive our model from a Gaussian bell-shaped waveform. Each element of the parametric family, also known as a manifold, is given by a fractional derivative/integrate of the shifted, scaled, and phase-rotated Gaussian. In the time domain, these waveforms are defined by a nonlinear function, $f_\theta : \mathbb{R}^5 \mapsto \mathbb{R}$, given by

$$f_\theta(t) = D^\alpha \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{(t-\tau)^2/2\sigma^2} \right) e^{j\pi\phi}, \quad (4)$$

with $\theta = [\tau, \sigma, \alpha, \phi]$ the set of parameters and D^α the α -order integration operator. Inspired by the work of Wakin et al. (2005), we can parametrize the windowed signals with the set of $\theta = [\sigma, \tau, \alpha]$ as

$$s^{(n)}(t) = M[\theta](t), \quad (5)$$

where $M[\theta] = \{f_\theta : \theta \in \Theta\}$ and Θ is feasible space for θ .

To estimate the different attributes for individual windowed waveforms, we need to find the optimal θ that minimizes the estimation error $e^{(n)}(\theta)$ given by

$$e^{(n)}(\theta) = \left\| f_\theta - s^{(n)} \right\|_2^2. \quad (6)$$

The minimization problem for isolated event $s^{(n)}$ is then given by

$$\tilde{\theta}^{(n)} = \arg \min_{\theta \in \Theta} \left\| s^{(n)} - M[\theta] \right\|_2^2 \quad \text{with } n = 1, N. \quad (7)$$

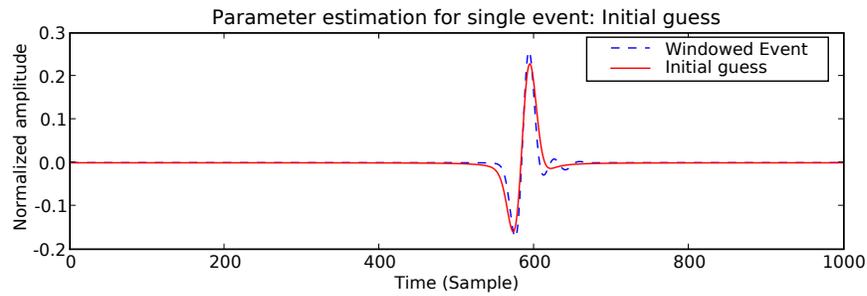
To solve this optimization problem, we employ the BFGS quasi-Newton method (Nocedal and Wright, 1999; Kelley, 1999). In quasi-Newton methods, one need not compute the second derivatives of the objective function for the Hessian matrix. Instead, the Hessian is updated by analyzing successive gradients. Alternatively, a trust region method with the Levenberg-Marquardt parameter (Kelley, 1999) can be used to solve the minimization problem. As with the BFGS method, trust region methods also require the manifold to be smooth. Partial derivatives of our model can be derived analytically since smoothness of Gaussian signals provides differentiability. This suggests we solve the above minimization problem in the frequency domain, where the fractional derivatives are known analytically. The elements of the Gaussian waveform family in the Fourier domain are given by (Blu and Unser, 2003)

$$\hat{f}_\theta(\omega) = (j\omega)^{-\alpha/2+\phi} (-j\omega)^{-\alpha/2-\phi} e^{-\frac{(\sigma^2\omega^2)}{2}} e^{-j\omega\tau}, \quad (8)$$

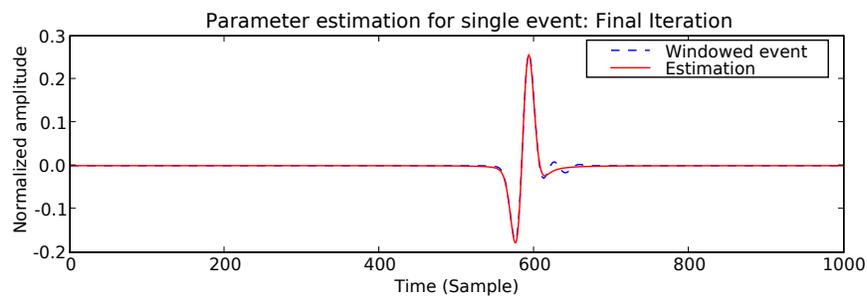
where $\hat{f}_\theta(\omega) = \mathcal{F}(f_\theta(t))$ represents Fourier transform of $f_\theta(t)$. Analytical expressions for partial derivatives of the manifold in the frequency domain are given by

$$\begin{aligned} \frac{\partial}{\partial \tau} \hat{f}_\theta(\omega) &= -j\omega \hat{f}_\theta(\omega), & \frac{\partial}{\partial \alpha} \hat{f}_\theta(\omega) &= -\ln(\omega) \hat{f}_\theta(\omega), \\ \frac{\partial}{\partial \sigma} \hat{f}_\theta(\omega) &= -\sigma\omega^2 \hat{f}_\theta(\omega), & \frac{\partial}{\partial \phi} \hat{f}_\theta(\omega) &= j\pi \hat{f}_\theta(\omega). \end{aligned} \quad (9)$$

Our experiments using the BFGS solver have shown that this optimization method provides acceptable and robust solutions for the estimation problem (see Figure 4). Figure 5 compares the original trace and the reconstructed trace by superposition of the estimated waveforms.



(a)



(b)

Figure 4: Parameter estimation for an individual event in Figure 3. (a) Initial iteration of parameter estimation for the isolated event where dashed line shows windowed event and solid line shows our guess. (b) Final iteration of parameter estimation for the isolated event where the estimated waveform matches the actual event. [NR] mohammad1/. estim-i,estim-f

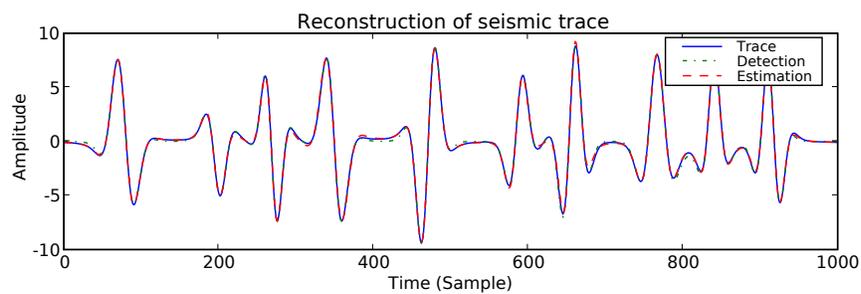


Figure 5: Estimated seismic signal formed by superposition of all characterized events and compared with the original seismic trace. [NR] mohammad1/. trace-match

DISCUSSION AND CONCLUSIONS

In this report, we review a new characterization method allowing for the estimation of fractional-order discontinuities. These scale attributes may lead to an improvement of geological interpretation from seismic traces. The examples we present indicate that the proposed characterization method generates accurate results. In addition, our method is well-suited for estimating scale exponent attributes from bandlimited data. As opposed to wavelet-coefficient-decay-based methods, such as SPICE, our method does not generate possibly ambiguous estimates since we do not rely on ‘infinite’ bandwidth.

ACKNOWLEDGMENTS

Felix J. Herrmann is gratefully thanked for his help throughout this research. The author thanks David Wilkinson for constructive discussions about this research. This work was in part financially supported by NSERC Discovery Grant 22R81254 and CRD Grant DNOISE 334810-05 of Felix J. Herrmann and was carried out as part of the CHARM project with support, secured through Chevron.

REFERENCES

- Blu, T. and M. Unser, 2003, A complete family of scaling functions: The (α, τ) fractional splines: Proceedings of the Twenty-Eighth IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’03), 505 – 508.
- Dossal, C. and S. Mallat, 2005, Sparse spike deconvolution with minimum scale: Proceedings of Signal Processing with Adaptive Sparse Structured Representations, 123 – 126. (<http://spars05.irisa.fr/ACTES/PS2-11.pdf>).
- Herrmann, F. J., 1998, Multiscale analysis of well- and seismic data: Mathematical Methods in Geophysical Imaging V, SPIE, **3453**, 180 – 208.
- Herrmann, F. J., W. J. Lyons, and C. Stark, 2001, Seismic facies characterization by monoscale analysis: Geophysical Research Letters, **28**, 3781 – 3784.
- Kelley, C. T., 1999, Iterative methods for optimization. Frontiers in Applied Mathematics, No. 18: SIAM.
- Liner, C., C.-F. Li, A. Gersztenkorn, and J. Smythe, 2004, SPICE: A new general seismic attribute: 74th Annual International Meeting, SEG, Expanded Abstracts, 433 – 436.
- Mallat, S., 1997, A wavelet tour of signal processing: Academic Press.
- Maysami, M., 2008, Lithology constraints from seismic waveforms: Application to opal-A to opal-CT transition: Master’s thesis, University of British Columbia.
- Maysami, M. and F. Herrmann, 2007, Seismic reflector characterization by a multiscale detection-estimation method: Presented at the EAGE 69th Conference & Exhibition.
- Nocedal, J. and S. J. Wright, 1999, Numerical optimization, 1st ed. Springer Series in Operations Research: Springer-Verlag.
- Robinson, E. A., 1957, Predictive decomposition of seismic traces: Geophysics, **22**, 767 – 778.
- Saggaf, M. M. and E. A. Robinson, 2000, A unified framework for the deconvolution of traces of nonwhite reflectivity: Geophysics, **65**, 1660 – 1676.
- Wakin, M. B., D. L. Donoho, H. Choi, and R. G. Baraniuk, 2005, The multiscale structure of non-differentiable image manifolds: SPIE Wavelets XI, 413 – 429.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Al-Theyab, Abdullah	723-3187	altheyab
Ayeni, Gboyega	723-0463	gayeni
Berryman, James	–	berryman
Biondi, Biondo	723-1319	biondo
Cardoso, Claudio	723-1250	claudio
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
Curry, Bill	723-6006	bill
Guitton, Antoine	–	antoine
Halpert, Adam	723-3187	adam
Lau, Diane	723-1703	diane
Maysami, Mohammad	723-9282	mohammad
de Ridder, Sjoerd	723-1250	sjoerd
Shen, Xukai	723-0463	xukai
Shrage, Jeff	723-0463	jeff
Tang, Yaxun	723-1250	tang

SEP fax number: (650) 723-0683

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, 2007-2008

Name	Company	Telephone	E-Mail
Raymond Abma	BP	(281) 366-4604	abmar1@bp.com
Dimitri Bevc	3DGeo	(408) 450-7840	dimitri@3dgeo.com
Biondo Biondi	SEP	(650) 723-1319	biondo@sep.stanford.edu
Robert Bloor (Co-chair, 1st year)	GX Technology	(713) 789-7250	rib@gxt.com
Luis L. Canales	WesternGeco	(713) 806-5271	lcanales@houston.westerngeco.slb.com
Jon Claerbout	SEP	(650) 723-3717	jon@sep.stanford.edu
Richard Cook	Shell	(713) 245-7195	richard.cook@shell.com
Stewart A. Levin (Co-chair, 2nd year)	Halliburton	(303) 488-3062	salevin@lgc.com

Research Personnel

Abdullah Al Theyab graduated from the University of Tulsa with a double major in Geosciences and Computer Science. In 2005, he joined the Exploration Organization in Saudi Aramco working in seismic data processing and interpretation. Currently, he is pursuing a M.S. degree in Stanford Exploration Project.



Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell scholar at University of Leeds, where he obtained an M.Sc with Distinction in Exploration Geophysics. Gboyega joined SEP in September 2006 to work towards his Ph.D in Geophysics. He is a member of SEG, EAGE, AGU and AAPG.



James G. Berryman received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on seismic waves in rocks and sediments – at AT&T Bell Laboratories (1978-81) and at Lawrence Livermore National Laboratory (1981-), where he is currently a physicist in the Energy and Environment Directorate. He received the Maurice Anthony Biot Medal of the ASCE in May, 2005, for his work in the mechanics and acoustics of porous media containing fluids. Continuing research interests include acoustic, seismic, and electrical methods of geophysical imaging and studies of waves in porous media. He is a member of APS, AGU, ASA, and SEG.



Biondo L. Biondi graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987, he worked as a Research Geophysicist for TOTAL, Compagnie Francaise des Petroles in Paris. After his Ph.D. at Stanford, Biondo worked for three years with Thinking Machines Co. on the applications of massively parallel computers to seismic processing. After leaving Thinking Machines, Biondo started 3DGeo Development, a software and service company devoted to high-end seismic imaging. Biondo is now Associate Professor (Research) of Geophysics and leads SEP efforts in 3-D imaging. He is a member of SEG and EAGE.



Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. 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.



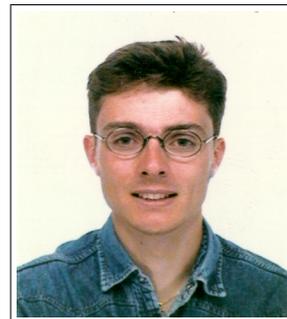
William Curry graduated in 2000 from the University of Alberta with a B.Sc. (Hons.) in Geophysics, and received his M.S. from SEP in 2002. From 1998-1999, he worked as a seismic interpreter with Jet Energy Corporation in Calgary, Canada. During the summer of 2002 he interned with Marathon Oil in Houston, in 2005 he interned with ExxonMobil Exploration Co., and in 2006 interned with 3DGeo Development. Bill is a student member of the EAGE, IEEE, SEG, and SIAM.



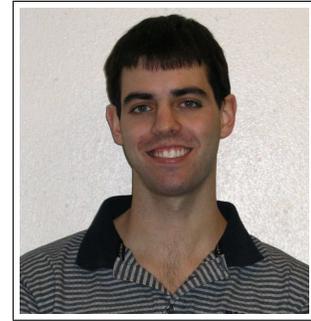
Claudio Guerra received his B.Sc. in Geology from Federal University of Rio de Janeiro, Brazil in 1988 and a M.Sc. from State University of Campinas, Brazil in 1999. Since 1989, he has been working for Petrobras, Brazil. He joined SEP in 2006 and is currently pursuing a Ph.D. in geophysics at Stanford University. He is member of SEG and SBGf.



Antoine Guitton received a M.Sc. in geophysics from Université de Strasbourg, France in 1996 and from Stanford University in 2000. He received his Ph.D. from Stanford University in 2005. He was awarded Best Student Paper from the SEG in 1999 and received the EAGE Van Weelden award in 2004. He was research assistant at the Institut Français du Pétrole (Paris-1996/97) working on well seismic imaging and research geophysicist at CGG Houston (1997-98) working on multiples attenuation. He is now Senior Geophysicist for 3DGeo Inc. in Santa Clara, CA. His research interests include interpolation, noise attenuation, inversion and novel seismic-interpretation techniques. He is a member of the SEG and EAGE.



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.



Mohammad Maysami graduated in 2005 with two B.Sc. in Electrical Engineering and Petroleum Engineering from Sharif University of Technology, Tehran, Iran. After completing a M.A.Sc. degree in January 2008 in Geophysics at University of British Columbia in Vancouver, Canada, he joined SEP. He is working toward a Ph.D. in Earth, Energy, and Environmental Sciences (EEES) which is jointly supervised by SEP and SCRF (Stanford Center for Reservoir Forecasting). He is a student member of SEG, AGU, and SPE.



Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences at Utrecht University, the Netherlands. He was a graduate exchange student at the Department of Geophysics, Colorado School of Mines, USA, in the fall of 2004 and the spring of 2005. He obtained an M.Sc. (Aug. 2007) at Delft University of Technology, the Netherlands, for his thesis 'Simulation of interferometric seismoelectric Green's function recovery' in Applied Earth Sciences with specialisation in applied geophysics. He joined SEP in the fall of 2007 and is currently pursuing a Ph.D. in Geophysics at Stanford University.



Guojian Shan received his B.Sc. in Mathematics School of Peking University in July, 1998. From 1998 to 2001, he studied in Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences (CAS), and received his M.S. in Applied Mathematics in July, 2001. He joined SEP in 2001 and received his Ph.D. in April, 2008. He is now working for Chevron as a research geophysicist. He is a member of the SEG.



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.



Jeff Shragge graduated in 1998 with a BScH in Honours Physics from Queen's University in Kingston, Canada. After completing a MSc degree in 2001 in teleseismic imaging at the UBC in Vancouver, Canada, he spent 2001 and 2002 working for a small geophysical consulting company based out of Woburn, MA. He joined SEP in 2002, and is working towards a Ph.D. in Geophysics. His main research interest is migration and wavefield inversion. He is a member of SEG and AGU.



Yaxun Tang received his B.Sc. (Jul. 2003) and M.Sc. (Aug. 2005) in Geophysics from School of Ocean and Earth Science, Tongji University, Shanghai. He joined SEP in 2005 and is currently working towards a Ph.D. in Geophysics at Stanford University. He is a member of SEG.



Alejandro A. Valenciano received a B.Sc. degree in Physics from Havana University (Cuba) in 1994, and a M.Sc. in Physics from Simon Bolivar University (Venezuela) in 1998. He worked in the Earth Science Department of PDVSA-INTEVEP from 1995 to 2001. He joined SEP to work towards a Ph.D in geophysics in the Fall of 2001.



SEP ARTICLES PUBLISHED OR IN PRESS

- Artman, B., 2006, Imaging passive seismic data: Geophysics.
- Artman, B., G. Alvarez, and K. Matson, 2007, Image-space surface-related multiple prediction: Geophysics.
- Biondi, B., 2007a, Angle-domain common image gathers from anisotropic migration: Geophysics, **72**, S81–S91.
- , 2007b, Residual moveout in anisotropic angle-domain common-image gathers: Geophysics, **72**, S93–S103.
- Biondi, B. L., 2006, 3D Seismic Imaging: Society of Exploration Geophysicists.
- , 2007c, Concepts and Applications in 3D Seismic Imaging: Society of Exploration Geophysicists, in press.
- Gray, S., D. Trad, B. Biondi, and L. Lines, 2006, Towards wave-equation imaging and velocity estimation: Recorder, **31**, 47–53.
- Guitton, A., B. Kaelin, and B. Biondi, 2007, Least-squares attenuation of reverse-time-migration artifacts: Geophysics, **72**, G19–G23.
- Gunther, R. and M. Reshef, 2007, Dip corrections for velocity analysis in super-gathers: J. Seis. Expl., **16**.
- Haines, S., A. Guitton, and B. Biondi, 2007a, Seismoelectric data processing for surface surveys of shallow targets: Geophysics, **72**, G1–G8.
- Haines, S., S. Pride, S. Klemperer, and B. Biondi, 2007b, Seismoelectric imaging of shallow targets: Geophysics, **72**, G9–G20.
- Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3D salt boundaries: Geophysics, in press.
- Rosales, D. and B. Biondi, 2006, Converted-wave azimuth moveout: Geophysics, **71**, S99–S110.
- Rosales, D. A., S. Fomel, and B. Biondi, 2007, Angle-domain common-image gathers for converted-wave: Geophysics, in press.
- Shragge, J., 2007, Waveform inversion by one-way wavefield extrapolation: SEG Technical Program Expanded Abstracts, **26**, 2040–2044.
- , 2008, Riemannian wavefield extrapolation: Nonorthogonal coordinate systems: Geophysics, **73**, T11–T21.
- Shragge, J. and G. Shan, 2007, Prestack depth migration in elliptic coordinates: SEG Technical Program Expanded Abstracts, **26**, 2185–2189.
- Valenciano, A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35.

