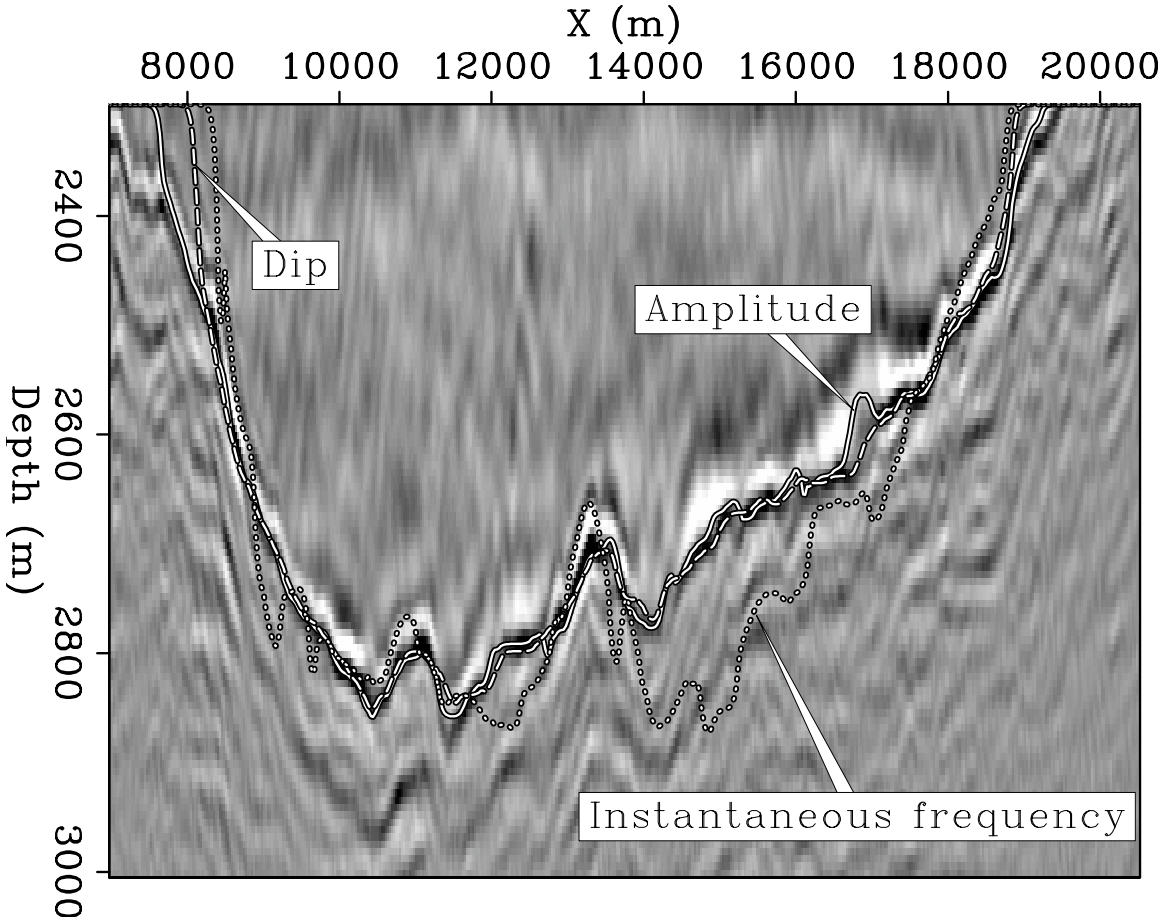
STANFORD EXPLORATION PROJECT

Abdullah Al Theyab, Gboyega Ayeni, Biondo Biondi, Jon Claerbout, Robert Clapp, Claudio Guerra, Seth Haines, Adam Halpert, Ching-Bih Liaw, Mohammad Maysami, Francis Muir, Nelson Nagales, Sjoerd de Ridder, Xukai Shen, Jeff Shragge, and Yaxun Tang

Report Number 136, October 2008



Preface

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

ER denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library[2]. 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 online[3].

[1] http://sepwww.stanford.edu/private/docs/sep136
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP136 — TABLE OF CONTENTS

Velocity Analysis

Yaxun Tang and Claudio Guerra and Biondo Biondi, Image-space wave-equation tomography in the generalized source domain 1

Claudio Guerra and Biondo Biondi, Phase encoding with Gold codes for wave-equation migration 23

Biondo Biondi, An image-focusing semblance functional for velocity analysis 43

Mohammad Maysami, Migration velocity analysis with cross-gradient constraint 55

Abdullah Al Theyab and Biondo Biondi, Transmission effects of localized variations of Earth's visco-acoustic parameters 61

Imaging Hessian

Gboyega Ayeni and Biondo Biondi, Joint wave-equation inversion of time-lapse seismic data 71

Yaxun Tang, Modeling, migration, and inversion in the generalized source and receiver domain 97

Computation and Interpretation

Adam Halpert and Robert G. Clapp, Salt body segmentation with dip and frequency attributes 113

Robert G. Clapp and Nelson Nagales, Hyercube viewer: New displays and new data-types 125

Ching-Bih Liaw and Robert G. Clapp, Many-core and PSPI: Mixing fine-grain and coarse-grain parallelism 131

Robert G. Clapp, Reverse time migration: Saving the boundaries 137

Imaging with non-standard coordinates and sources

Jeff Shragge, Angle-domain common-image gathers in generalized coordinates 145

Sjoerd de Ridder, Seismic interferometry versus spatial auto-correlation method on the regional coda of the NPE 165

Interpolation and Inversion

Xukai Shen, 3D pyramid interpolation 179

Jon Claerbout, Anti-crosstalk 189

Miscellaneous

Sjoerd de Ridder and Seth S. Haines, Seismic investigation of natural coal fires: A pre-fieldwork synthetic feasibility study..... 195

Jon Claerbout and Francis Muir, Hubbert math 203

Jon Claerbout, Tar sands: Reprieve or apocalypse? 207

SEP phone directory 213

('SEP article published or in press, 2008',)..... 219

Image-space wave-equation tomography in the generalized source domain

Yaxun Tang, Claudio Guerra, and Biondo Biondi

ABSTRACT

We extend the theory of image-space wave-equation tomography to the generalized source domain, where a smaller number of synthesized shot gathers are generated either by data-space phase encoding or image-space phase encoding. We demonstrate how to evaluate the wave-equation forward tomographic operator and its adjoint in this new domain. We compare the gradients of the tomography objective functional obtained using both data-space and image-space encoded gathers with that obtained using the original shot gathers. We show that with those encoded shot gathers we can obtain a gradient similar to that computed in the original shot-profile domain, but at lower computational cost. The saving in cost is important for putting this theory into practical applications. We illustrate our examples on a simple model with Gaussian anomalies in the subsurface.

INTRODUCTION

Wave-equation tomography has the potential to accurately estimate the velocity model in complex geological scenarios where ray-based traveltime tomography is prone to fail. Wave-equation-based tomography uses band-limited wavefields instead of infinite-frequency rays as carriers of information, thus it is robust even in the presence of strong velocity contrasts and immune from multi-pathing issues. Generally speaking, wave-equation tomography can be classified into two different categories based on the domain where it minimizes the residual. The domain can be either the data space or the image space. The data-space approach directly compares the modeled waveform with the recorded waveform, and is widely known as waveform inversion, or data-space wave-equation tomography (Tarantola, 1987; Mora, 1989; Woodward, 1992; Pratt, 1999). The main disadvantage of the data-space approach is that in complex areas, the recorded waveforms can be very complicated and are usually of low signal-to-noise ratio (S/N), so matching the full waveform might be extremely difficult. On the other hand, the image-space approach, also known as image-space wave-equation tomography, minimizes the residual in the image domain obtained after migration. The migrated image is often much simpler than the original data, because even with a relatively inaccurate velocity, migration is able to (partially) collapse diffractions and enhance the S/N; thus the image-space wave-equation tomography has the potential to mitigate some of the difficulties that we encounter in the data-space approach. Another advantage of the image-space approach is that the more efficient one-way wave-equation extrapolator can be used. In waveform inversion, however, the one-way propagator is difficult (if not impossible) to use because of its inability to model the multiple arrivals, although some tweaks can be employed so that the one-way propagator can be applied to turning-wave tomography (Shragge, 2007).

However, despite its theoretical advantages, image-space wave-equation tomography is still computationally challenging. Each iteration of tomographic velocity updating is computationally expensive and often converges slowly. Practical applications are still rare and small in scale (Biondi and Sava, 1999; Shen et al., 2005; Albertin et al., 2006). The goal of this paper is to extend the theory of image-space wave-equation tomography from the conventional shot-profile domain (Shen, 2004; Shen et al., 2005) to the generalized source domain, where a smaller number of synthesized shot gathers make the tomographic velocity update substantially faster.

The generalized source domain can be obtained either by data-space phase encoding or image-space phase encoding. For the data-space phase encoding, the synthesized shot gathers are obtained by linear combination of the original shot gathers after some kind of phase encoding; in particular, here we mainly consider plane-wave phase encoding (Whitmore, 1995; Zhang et al., 2005; Duquet and Lailly, 2006; Liu et al., 2006) and random phase encoding (Romero et al., 2000). As the encoding process is done in the data space, we call it data-space phase encoding. For the image-space phase encoding, the synthesized gathers are obtained by prestack exploding-reflector modeling (Biondi, 2006, 2007; Guerra and Biondi, 2008b), where several subsurface-offset-domain common-image gathers (SODCIGs) and several reflectors are simultaneously demigrated to generate areal source and areal receiver gathers. To attenuate the cross-talk, the SODCIGs and the reflectors have to be encoded, e.g., by random phase encoding. Because the encoding process is done in the image space, we call it image-space phase encoding. We show that in these generalized source domains, we can obtain gradients, which are used for updating the velocity model, similar to that obtained in the original shot-profile domain, but with less computational cost.

This paper is organized as follows: We first briefly review the theory of image-space wave-equation tomography. Then we discuss how to evaluate the forward tomographic operator and its adjoint in the original shot-profile domain. The latter is an important component in computing the gradient of the tomography objective functional. We then extend the theory to the generalized source domain. Finally, we show examples on a simple synthetic model.

IMAGE-SPACE WAVE-EQUATION TOMOGRAPHY

Image-space wave-equation tomography is a non-linear inverse problem that tries to find an optimal background slowness that minimizes the residual field, $\mathbf{\Delta I}$, defined in the image space. The residual field is derived from the background image, $\mathbf{I}$, which is computed with a background slowness (or the current estimate of the slowness). The residual field measures the correctness of the background slowness; its minimum (under some norm, e.g. ℓ_2) is achieved when a correct background slowness has been used for migration. There are many choices of the residual field, such as residual moveout in the Angle-Domain Common-Image Gathers (ADCIGs), differential semblance in the ADCIGs, reflection-angle stacking power (in which case we have to maximize the residual field, or minimize the negative stacking power), etc.. Here we follow a definition similar to that in Biondi (2008), and define a general form of the residual field as follows:

$$\mathbf{\Delta I} = \mathbf{I} - \mathbf{F}(\mathbf{I}), \tag{1}$$

where $\mathbf{F}$ is a focusing operator, which measures the focusing of the migrated image. For example, in the Differential Semblance Optimization (DSO) method (Shen, 2004), the focusing operator takes the following form:

$$\mathbf{F}(\mathbf{I}) = (\mathbf{1} - \mathbf{O})\,\mathbf{I}, \tag{2}$$

where $\mathbf{1}$ is the identity operator and $\mathbf{O}$ is the DSO operator either in the subsurface offset domain or in the angle domain (Shen, 2004). The subsurface-offset-domain DSO focuses the energy at zero offset, whereas the angle-domain DSO flattens the ADCIGs.

In the wave-equation migration velocity analysis (WEMVA) method (Sava, 2004), the focusing operator is the linearized residual migration operator defined as follows:

$$\mathbf{F}(\mathbf{I}) = \mathbf{R}[\rho]\mathbf{I} \approx \mathbf{I} + \mathbf{K}[\Delta\rho]\mathbf{I}, \tag{3}$$

where ρ is the ratio between the background slowness $\widehat{\mathbf{s}}$ and the true slowness $\mathbf{s}$, and $\Delta\rho = 1 - \rho = 1 - \frac{\widehat{\mathbf{s}}}{\mathbf{s}}$; $\mathbf{R}[\rho]$ is the residual migration operator (Sava, 2003), and $\mathbf{K}[\Delta\rho]$ is the differential residual migration operator defined as follows (Sava and Biondi, 2004a,b):

$$\mathbf{K}[\Delta\rho] = \Delta\rho\,\frac{\partial \mathbf{R}[\rho]}{\partial \rho}\bigg|_{\rho=1}. \tag{4}$$

The linear operator $\mathbf{K}[\Delta\rho]$ applies different phase rotations to the image for different reflection angles and geological dips (Biondi, 2008).

In general, if we choose ℓ_2 norm, the tomography objective function to minimize can be written as follows:

$$J = \frac{1}{2}||\mathbf{\Delta I}||_2 = \frac{1}{2}||\mathbf{I} - \mathbf{F}(\mathbf{I})||_2, \tag{5}$$

where $||\cdot||_2$ stands for the ℓ_2 norm. Gradient-based optimization techniques such as the quasi-Newton method and the conjugate gradient method can be used to minimize the objective function J. The gradient of J with respect to the slowness $\mathbf{s}$ reads as follows:

$$\nabla J = \Re\left(\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}} - \frac{\partial \mathbf{F}(\mathbf{I})}{\partial \mathbf{s}}\right)'(\mathbf{I} - \mathbf{F}(\mathbf{I}))\right), \tag{6}$$

where $\Re$ denotes taking the real part of a complex value and $'$ denotes the adjoint. For the DSO method, the linear operator $\mathbf{O}$ is independent of the slowness, so we have

$$\frac{\partial \mathbf{F}(\mathbf{I})}{\partial \mathbf{s}} = (\mathbf{1} - \mathbf{O})\frac{\partial \mathbf{I}}{\partial \mathbf{s}}. \tag{7}$$

Substituting Equations 2 and 7 into Equation 6 and evaluating the gradient at a background slowness yields

$$\nabla J_{\mathrm{DSO}} = \Re\left(\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\bigg|_{\mathbf{s}=\widehat{\mathbf{s}}}\right)'\mathbf{O}'\mathbf{O}\widehat{\mathbf{I}}\right), \tag{8}$$

where $\widehat{\mathbf{I}}$ is the background image computed using the background slowness $\widehat{\mathbf{s}}$.

For the WEMVA method, the gradient is slightly more complicated, because in this case, the focusing operator is also dependent on the slowness $\mathbf{s}$. However, one can simplify

it by assuming that the focusing operator is applied on the background image $\widehat{\mathbf{I}}$ instead of $\mathbf{I}$, and $\widehat{\Delta\rho}$ is also picked from the background image $\widehat{\mathbf{I}}$, that is

$$\mathbf{F}(\widehat{\mathbf{I}}) = \widehat{\mathbf{I}} + \mathbf{K}[\widehat{\Delta\rho}]\widehat{\mathbf{I}}. \tag{9}$$

With these assumptions, we get the "classic" WEMVA gradient as follows:

$$\nabla J_{\text{WEMVA}} = \Re\left(-\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\Big|_{\mathbf{s}=\widehat{\mathbf{s}}}\right)' \mathbf{K}[\widehat{\Delta\rho}]\widehat{\mathbf{I}}\right). \tag{10}$$

The complete WEMVA gradient without the above assumptions can also be derived following the method described by Biondi (2008).

No matter which gradient we choose to back-project the slowness perturbation, we have to evaluate the adjoint of the linear operator $\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\big|_{\mathbf{s}=\widehat{\mathbf{s}}}$, which defines a linear mapping from the slowness perturbation $\mathbf{\Delta s}$ to the image perturbation $\mathbf{\Delta I}$. This is easy to see by expanding the image $\mathbf{I}$ around the background slowness $\widehat{\mathbf{s}}$ as follows:

$$\mathbf{I} = \widehat{\mathbf{I}} + \frac{\partial \mathbf{I}}{\partial \mathbf{s}}\Big|_{\mathbf{s}=\widehat{\mathbf{s}}} (\mathbf{s} - \widehat{\mathbf{s}}) + \cdots. \tag{11}$$

Keeping only the zero and first order terms, we get the linear operator $\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\big|_{\mathbf{s}=\widehat{\mathbf{s}}}$ as follows:

$$\mathbf{\Delta I} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}\Big|_{\mathbf{s}=\widehat{\mathbf{s}}} \mathbf{\Delta s} = \mathbf{T}\mathbf{\Delta s}, \tag{12}$$

where $\mathbf{\Delta I} = \mathbf{I} - \widehat{\mathbf{I}}$ and $\mathbf{\Delta s} = \mathbf{s} - \widehat{\mathbf{s}}$. $\mathbf{T} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}\big|_{\mathbf{s}=\widehat{\mathbf{s}}}$ is the wave-equation tomographic operator. The tomographic operator can be evaluated either in the source and receiver domain (Sava, 2004) or in the shot-profile domain (Shen, 2004). In next section we follow an approach similar to that discussed by Shen (2004) and review the forward and adjoint tomographic operator in the shot-profile domain. In the subsequent sections, we generalize the expression of the tomographic operator to the generalized source domain.

THE TOMOGRAPHIC OPERATOR IN THE SHOT-PROFILE DOMAIN

For the conventional shot-profile migration, both source and receiver wavefields are downward continued with the following one-way wave equations (Claerbout, 1971):

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) D(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ D(x, y, z = 0, \mathbf{x}_s, \omega) = \overline{f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s)} \end{cases}, \tag{13}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) U(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ U(x, y, z = 0, \mathbf{x}_s, \omega) = Q(x, y, z = 0, \mathbf{x}_s, \omega) \end{cases}, \tag{14}$$

where the overline stands for complex conjugate; $D(\mathbf{x}, \mathbf{x}_s, \omega)$ is the source wavefield for a single frequency ω at image point $\mathbf{x} = (x, y, z)$ with the source located at $\mathbf{x}_s = (x_s, y_s, 0)$; $U(\mathbf{x}, \mathbf{x}_s, \omega)$ is the receiver wavefield for a single frequency ω at image point $\mathbf{x}$ for the source

located at $\mathbf{x}_s$; $s(\mathbf{x})$ is the slowness at $\mathbf{x}$; $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector; $f_s(\omega)$ is the frequency dependent source signature, and $\overline{f_s(\omega)\delta(\mathbf{x}-\mathbf{x}_s)}$ defines the point source function at $\mathbf{x}_s$, which serves as the boundary condition of Equation 13. $Q(x, y, z = 0, \mathbf{x}_s, \omega)$ is the recorded shot gather for the shot located at $\mathbf{x}_s$, which serves as the boundary condition of Equation 14. To produce the image, the following cross-correlation imaging condition is used:

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega), \tag{15}$$

where $\mathbf{h} = (h_x, h_y, h_z)$ is the subsurface half offset.

The perturbed image can be derived by a simple application of the chain rule to Equation 15:

$$\begin{aligned} \Delta I(\mathbf{x}, \mathbf{h}) &= \sum_{\mathbf{x}_s} \sum_{\omega} \Big(\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) + \\ &\quad \widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) \Big), \end{aligned} \tag{16}$$

where $\widehat{D}(\mathbf{x}-\mathbf{h}, \mathbf{x}_s, \omega)$ and $\widehat{U}(\mathbf{x}+\mathbf{h}, \mathbf{x}_s, \omega)$ are the background source and receiver wavefields computed with the background slowness $\widehat{s}(\mathbf{x})$; $\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)$ and $\Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega)$ are the perturbed source wavefield and perturbed receiver wavefield, which are the results of the slowness perturbation $\Delta s(\mathbf{x})$. The perturbed source and receiver wavefields satisfy the following one-way wave equations, which are linearized with respect to slowness (see Appendix A for derivations):

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2} \right) \Delta D(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{-i\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{s}^2(\mathbf{x})}}} \widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega) \\ \Delta D(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}, \tag{17}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2} \right) \Delta U(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{-i\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{s}^2(\mathbf{x})}}} \widehat{U}(\mathbf{x}, \mathbf{x}_s, \omega) \\ \Delta U(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}. \tag{18}$$

Recursively solving Equations 17 and 18 gives us the perturbed source and receiver wavefields. The perturbed source and receiver wavefields are then used in Equation 16 to generate the perturbed image $\Delta I(\mathbf{x}, \mathbf{h})$, where the background source and receiver wavefields are precomputed by recursively solving Equations 13 and 14 with a background slowness $\widehat{s}(\mathbf{x})$. Appendix B gives a more detailed matrix representation of how to evaluate the forward tomographic operator $\mathbf{T}$.

To evaluate the adjoint tomographic operator $\mathbf{T}'$, we first apply the adjoint of the imaging condition in Equation 16 to get the perturbed source and receiver wavefields $\Delta D(\mathbf{x}, \mathbf{x}_s, \omega)$ and $\Delta U(\mathbf{x}, \mathbf{x}_s, \omega)$ as follows:

$$\Delta D(\mathbf{x}, \mathbf{x}_s, \omega) = \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \overline{\widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega)}, \tag{19}$$

$$\Delta U(\mathbf{x}, \mathbf{x}_s, \omega) = \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \overline{\widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)}. \tag{20}$$

Then we solve the adjoint equations of Equations 17 and 18 to get the slowness perturbation $\Delta s(\mathbf{x})$. Again, in order to solve the adjoint equations of Equations 17 and 18, the background source wavefield $\widehat{D}(\mathbf{x},\mathbf{x}_s,\omega)$ and the background receiver wavefield $\widehat{U}(\mathbf{x},\mathbf{x}_s,\omega)$ have to be computed in advance. Appendix C gives a more detailed matrix representation of how to evaluate the adjoint tomographic operator $\mathbf{T}'$.

TOMOGRAPHY WITH THE ENCODED WAVEFIELDS

It is clear from previous sections that the cost for computing the gradient of the objective function J in the original shot-profile domain is at least twice the cost of a shot-profile migration, because to compute the perturbed wavefields, the background wavefields are required. Because minimizing the objective function J requires a considerable number of gradient and function evaluations, image-space wave-equation tomography in the conventional shot-profile domain seems to be infeasible for large-scale 3-D applications, even with modern computer resources. To reduce the cost and make this powerful method more practical, we extend the theory of image-space wave-equation tomography to the generalized source domain, where a smaller number of synthesized shot gathers are used for computing the gradient. We discuss two different strategies to generate the generalized shot gathers, i.e., the data-space phase-encoding method and the image-space phase-encoding method, both of which can achieve considerable data reduction while still keeping the necessary kinematic information for velocity analysis.

Data-space encoded wavefields

The data-space encoded shot gathers are obtained by linear combination of the original shot gathers after phase encoding. For simplicity, we mainly consider plane-wave phase-encoding (Whitmore, 1995; Zhang et al., 2005; Duquet and Lailly, 2006; Liu et al., 2006) and random phase-encoding (Romero et al., 2000). Because of the linearity of the one-way wave equation with respect to the wavefield, the encoded source and receiver wavefields also satisfy the same one-way wave equations defined by Equations 13 and 14, but with different boundary conditions:

$$\begin{cases} \left(\frac{\partial}{\partial z}+i\sqrt{\omega^2 s^2(\mathbf{x})-|\mathbf{k}|^2}\right)\widetilde{D}(\mathbf{x},\mathbf{p}_s,\omega)=0 \\ \widetilde{D}(x,y,z=0,\mathbf{p}_s,\omega)=\sum_{\mathbf{x}_s}\overline{f_s(\omega)\delta(\mathbf{x}-\mathbf{x}_s)\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)} \end{cases}, \tag{21}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z}+i\sqrt{\omega^2 s^2(\mathbf{x})-|\mathbf{k}|^2}\right)\widetilde{U}(\mathbf{x},\mathbf{p}_s,\omega)=0 \\ \widetilde{U}(x,y,z=0,\mathbf{x}_s,\omega)=\sum_{\mathbf{x}_s}Q(x,y,z=0,\mathbf{x}_s,\omega)\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega) \end{cases}, \tag{22}$$

where $\widetilde{D}(\mathbf{x},\mathbf{p}_s,\omega)$ and $\widetilde{U}(\mathbf{x},\mathbf{p}_s,\omega)$ are the encoded source and receiver wavefields respectively, and $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)$ is the phase-encoding function. In the case of plane-wave phase encoding, $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)$ is defined as

$$\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)=e^{i\omega\mathbf{p}_s\mathbf{x}_s}, \tag{23}$$

where $\mathbf{p}_s$ is the ray parameter for the source plane waves on the surface. In the case of random phase encoding, the phase function is

$$\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega) = e^{i\gamma(\mathbf{x}_s, \mathbf{p}_s, \omega)}, \tag{24}$$

where $\gamma(\mathbf{x}_s, \mathbf{p}_s, \omega)$ is a random sequence in $\mathbf{x}_s$ and ω. The parameter $\mathbf{p}_s$ defines the index of different realizations of the random sequence (Tang, 2008). The final image is obtained by applying the cross-correlation imaging condition and summing the images for all $\mathbf{p}_s$'s:

$$I_{\rm de}(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{p}_s} \sum_{\omega} |c|^2 \widetilde{D}(\mathbf{x} - \mathbf{h}, \mathbf{p}_s, \omega) \widetilde{U}(\mathbf{x} + \mathbf{h}, \mathbf{p}_s, \omega), \tag{25}$$

where $c = \omega$ for plane-wave phase encoding and $c = 1$ for random phase encoding (Tang, 2008). It has been shown by Etgen (2005) and Liu et al. (2006) that plane-wave phase-encoding migration, by stacking a considerable number of $\mathbf{p}_s$, produces a migrated image almost identical to the shot-profile migrated image. If the original shots are well sampled, the number of plane waves required for migration is generally much smaller than the number of the original shot gathers (Etgen, 2005). Therefore plane-wave source migration is widely used in practice. Random-phase encoding migration is also an efficient tool, but the random phase function is not very effective in attenuating the crosstalk, especially when many sources are simultaneously encoded (Romero et al., 2000; Tang, 2008). Nevertheless, if many realizations of the random sequences are used, the final stacked image would also be approximately the same as the shot-profile migrated image. Therefore, the following relation approximately holds:

$$I(\mathbf{x}, \mathbf{h}) \approx I_{\rm de}(\mathbf{x}, \mathbf{h}). \tag{26}$$

That is, with the data-space encoded gathers, we obtain an image similar to that computed by the more expensive shot-profile migration. From Equation 25, the perturbed image can be easily obtained as follows:

$$\begin{aligned} \Delta I_{\rm de}(\mathbf{x}, \mathbf{h}) \;=\; & \sum_{\mathbf{p}_s} \sum_{\omega} |c|^2 \Big(\Delta\widetilde{D}(\mathbf{x} - \mathbf{h}, \mathbf{p}_s, \omega) \widehat{\widetilde{U}}(\mathbf{x} + \mathbf{h}, \mathbf{p}_s, \omega) + \\ & \widehat{\widetilde{D}}(\mathbf{x} - \mathbf{h}, \mathbf{p}_s, \omega) \Delta\widetilde{U}(\mathbf{x} + \mathbf{h}, \mathbf{p}_s, \omega) \Big), \end{aligned} \tag{27}$$

where $\widehat{\widetilde{D}}(\mathbf{x}, \mathbf{p}_s, \omega)$ and $\widehat{\widetilde{U}}(\mathbf{x}, \mathbf{p}_s, \omega)$ are the data-space encoded background source and receiver wavefields; $\Delta\widetilde{D}(\mathbf{x}, \mathbf{p}_s, \omega)$ and $\Delta\widetilde{U}(\mathbf{x}, \mathbf{p}_s, \omega)$ are the perturbed source and receiver wavefields in the data-space phase-encoding domain, which satisfy the perturbed one-way wave equations defined by Equations 17 and 18. The tomographic operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$ can be implemented in a manner similar to that discussed in Appendices B and C by replacing the original wavefields with the data-space phase encoded wavefields.

Image-space encoded wavefields

The image-space encoded gathers are obtained using the prestack exploding-reflector modeling method introduced by Biondi (2006) and Biondi (2007). The general idea of this method is to model the data and the corresponding source function that are related to only

one event in the subsurface, where a single unfocused SODCIG (obtained with an inaccurate velocity model) is used as the initial condition for the recursive upward continuation with the following one-way wave equations:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_D(\mathbf{x},\omega;x_m,y_m) = I_D(\mathbf{x},\mathbf{h};x_m,y_m) \\ Q_D(x,y,z=z_{\max},\omega;x_m,y_m) = 0 \end{cases}, \tag{28}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_U(\mathbf{x},\omega;x_m,y_m) = I_U(\mathbf{x},\mathbf{h};x_m,y_m) \\ Q_U(x,y,z=z_{\max},\omega;x_m,y_m) = 0 \end{cases}, \tag{29}$$

where $I_D(\mathbf{x},\mathbf{h};x_m,y_m)$ and $I_U(\mathbf{x},\mathbf{h};x_m,y_m)$ are the isolated SODCIGs at the horizontal location (x_m,y_m) for a single reflector, and are suitable for the initial conditions for the source and receiver wavefields, respectively. They are obtained by rotating the original unfocused SODCIGs according to the apparent geological dip of the reflector. This rotation maintains the velocity information needed for migration velocity analysis, especially for dipping reflectors (Biondi, 2007). By collecting the wavefields at the surface, we obtain the areal source data $Q_D(x,y,z=0,\omega;x_m,y_m)$ and the areal receiver data $Q_U(x,y,z=0,\omega;x_m,y_m)$ for a single reflector and a single SODCIG located at (x_m,y_m).

Since the size of the migrated image volume can be very big in practice and there are usually many reflectors in the subsurface, modeling each reflector and each SODCIG one by one may generate a data set even bigger than the original data set. One strategy to reduce the cost is to model several reflectors and several SODCIGs simultaneously (Biondi, 2006); however, this process generates unwanted crosstalk. As discussed by Guerra and Biondi (2008b,a), random phase encoding could be used to attenuate the crosstalk. The randomly encoded areal source and areal receiver wavefields can be computed as follows:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_D(\mathbf{x},\mathbf{p}_m,\omega) = \widetilde{I}_D(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega) \\ Q_D(x,y,z=z_{\max},\mathbf{p}_m,\omega) = 0 \end{cases}, \tag{30}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_U(\mathbf{x},\mathbf{p}_m,\omega) = \widetilde{I}_U(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega) \\ Q_U(x,y,z=z_{\max},\mathbf{p}_m,\omega) = 0 \end{cases}, \tag{31}$$

where $\widetilde{I}_D(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega)$ and $\widetilde{I}_U(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega)$ are the encoded SODCIGs after rotations. They are defined as follows:

$$\widetilde{I}_D(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega) = \sum_{x_m}\sum_{y_m} I_D(\mathbf{x},\mathbf{h},x_m,y_m)\beta(\mathbf{x},x_m,y_m,\mathbf{p}_m,\omega), \tag{32}$$

$$\widetilde{I}_U(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega) = \sum_{x_m}\sum_{y_m} I_U(\mathbf{x},\mathbf{h},x_m,y_m)\beta(\mathbf{x},x_m,y_m,\mathbf{p}_m,\omega), \tag{33}$$

where $\beta(\mathbf{x},x_m,y_m,\mathbf{p}_m,\omega) = e^{i\gamma(\mathbf{x},x_m,y_m,\mathbf{p}_m,\omega)}$ is chosen to be the random phase-encoding function, with $\gamma(\mathbf{x},x_m,y_m,\mathbf{p}_m,\omega)$ being a uniformly distributed random sequence in $\mathbf{x}$, x_m, y_m and ω; the variable $\mathbf{p}_m$ is the index of different realizations of the random sequence. Recursively solving Equations 30 and 31 gives us the encoded areal source data $Q_D(x,y,z=$

$0, \mathbf{p}_m, \omega)$ and areal receiver data $Q_U(x, y, z = 0, \mathbf{p}_m, \omega)$, which can be collected on the surface.

The synthesized new data sets are downward continued using the same one-way wave equation defined by Equations 13 and 14 (with different boundary conditions) as follows:

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{D}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widetilde{D}(x, y, z = 0, \mathbf{p}_m, \omega) = Q_D(x, y, z = 0, \mathbf{p}_m, \omega) \end{array} \right. , \tag{34}$$

and

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{U}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widetilde{U}(x, y, z = 0, \mathbf{x}_s, \omega) = Q_U(x, y, z = 0, \mathbf{p}_m, \omega) \end{array} \right. , \tag{35}$$

where $\widetilde{D}(\mathbf{x}, \mathbf{p}_m, \omega)$ and $\widetilde{U}(\mathbf{x}, \mathbf{p}_m, \omega)$ are the downward continued areal source and areal receiver wavefields for realization $\mathbf{p}_m$. The image is produced by cross-correlating the two wavefields and summing images for all realization $\mathbf{p}_m$ as follows:

$$I_{\rm me}(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{p}_m} \sum_{\omega} \widetilde{D}(\mathbf{x}, \mathbf{p}_m, \omega) \widetilde{U}(\mathbf{x}, \mathbf{p}_m, \omega). \tag{36}$$

The crosstalk artifacts can be further attenuated if the number of $\mathbf{p}_m$ is large; therefore, approximately, the image obtained by migrating the image-space encoded gathers is kinematically equivalent to the image obtained in the shot-profile domain.

From Equation 36, the perturbed image is easily obtained as follows:

$$\begin{aligned} \Delta I_{\rm me}(\mathbf{x}, \mathbf{h}) &= \sum_{\mathbf{p}_m} \sum_{\omega} \Big(\Delta\widetilde{D}(\mathbf{x} - \mathbf{h}, \mathbf{p}_m, \omega) \widehat{\widetilde{U}}(\mathbf{x} + \mathbf{h}, \mathbf{p}_m, \omega) + \\ &\quad \widehat{\widetilde{D}}(\mathbf{x} - \mathbf{h}, \mathbf{p}_m, \omega) \Delta\widetilde{U}(\mathbf{x} + \mathbf{h}, \mathbf{p}_m, \omega) \Big) , \end{aligned} \tag{37}$$

where $\widehat{\widetilde{D}}(\mathbf{x}, \mathbf{p}_m, \omega)$ and $\widehat{\widetilde{U}}(\mathbf{x}, \mathbf{p}_m, \omega)$ are the image-space encoded background source and receiver wavefields; $\Delta\widetilde{D}(\mathbf{x}, \mathbf{p}_m, \omega)$ and $\Delta\widetilde{U}(\mathbf{x}, \mathbf{p}_m, \omega)$ are the perturbed source and receiver wavefields in the image-space phase-encoding domain, which satisfy the perturbed one-way wave equations defined by Equations 17 and 18. The tomographic operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$ can be implemented in a manner similar to that discussed in Appendices B and C, by replacing the original wavefields with the image-space phase-encoded wavefields.

NUMERICAL EXAMPLES

We test the image-space wave-equation tomography in the generalized source domain on a simple model which contains only one reflector located at $z = 1500$ m. Figure 1 shows the correct slowness model. The slowness model consists of a constant background slowness 1/2000 s/m and two Gaussian anomalies located at $(x = -800, z = 800)$ and $(x = 800, z = 800)$ respectively. The left anomaly has 5% higher slowness, while the right one has 5% lower slowness. We modeled 401 shots ranging from -4000 m to 4000 m, with a shot interval 20 m. The receiver locations also range from -4000 m to 4000 m, but with a 10 m interval. The receivers are fixed for all shots to mimic a land acquisition geometry.

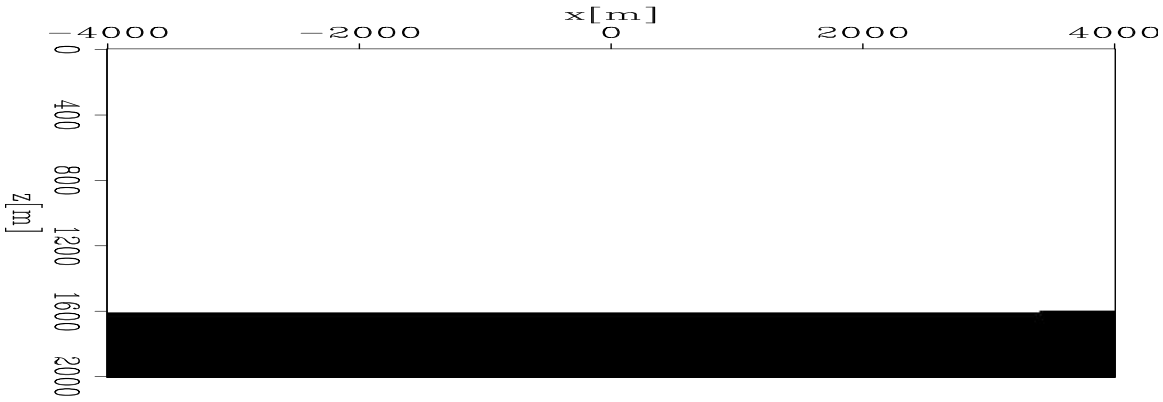


Figure 1: The correct slowness model. The slowness model consists of a constant background slowness (1/2000 s/m) and two 5% Gaussian anomalies. [**ER**] yaxun1/. twin-slow

Figure 2 shows the migrated images in different domains computed with a background slowness $\widehat{s} = 1/2000$ s/m. Figure 2(a) is obtained by migrating the original 401 shot gathers. Because of the inaccuracy of the slowness model, we can identify the mispositioning of the reflectors, especially beneath the Gaussian anomalies. Figure 2(b) is obtained by migrating the data-space plane-wave encoded gathers, where 61 plane waves are migrated; the result is almost identical to that in Figure 2(a); Figure 2(c) is obtained by migrating the image-space encoded gathers. The image-space encoded areal source and receiver data are generated by simultaneously modeling 100 randomly encoded unfocused SODCIGs, and 4 realizations of the random sequence are used; hence we have 40 image-space encoded areal gathers (each realization contains 10 areal shots). The kinematics of the result look almost the same as those in Figure 2(a). However, notice the wavelet squeezing effect and the random noise in the background caused by the random phase encoding.

Figure 3 shows the image perturbations obtained by applying the forward tomographic operator $\mathbf{T}$ in different domains. For this example, we assume that we know the correct slowness perturbation $\Delta\mathbf{s}$, which is obtained by subtracting the background slowness $\widehat{\mathbf{s}}$ from the correct slowness $\mathbf{s}$. Figure 3(a) shows the image perturbation computed with the original 401 shot gathers; notice the relative 90 degree phase rotation compared to the background image shown in Figure 2(a). Figure 3(b) is the result obtained by using 61 data-space plane-wave encoded gathers; the result is almost identical to Figure 3(a). Figure 3(c) shows the result computed with 40 image-space encoded gathers; the kinematics are also similar to those in Figure 3(a).

Figure 4 illustrates the predicted slowness perturbations by applying the adjoint tomographic operator $\mathbf{T}'$ to the image perturbations obtained in Figure 3. For comparison, Figure 4(a) shows the correct slowness perturbation, i.e., $\Delta\mathbf{s} = \mathbf{s} - \widehat{\mathbf{s}}$; Figure 4(b) is the predicted slowness perturbation by back-projecting Figure 3(a) using all 401 shot gathers; Figure 4(c) is the result by back-projecting Figure 3(b) using all 61 data-space plane-wave encoded gathers and is almost identical to Figure 4(b); Figure 4(d) shows the result by back-projecting Figure 3(c) using all 40 image-space encoded gathers. The result is also similar to Figure 4(b). However, notice that Figure 4(d) shows a slightly less focused result than Figure 4(b) and (c), which might be caused by the unattenuated crosstalk and the

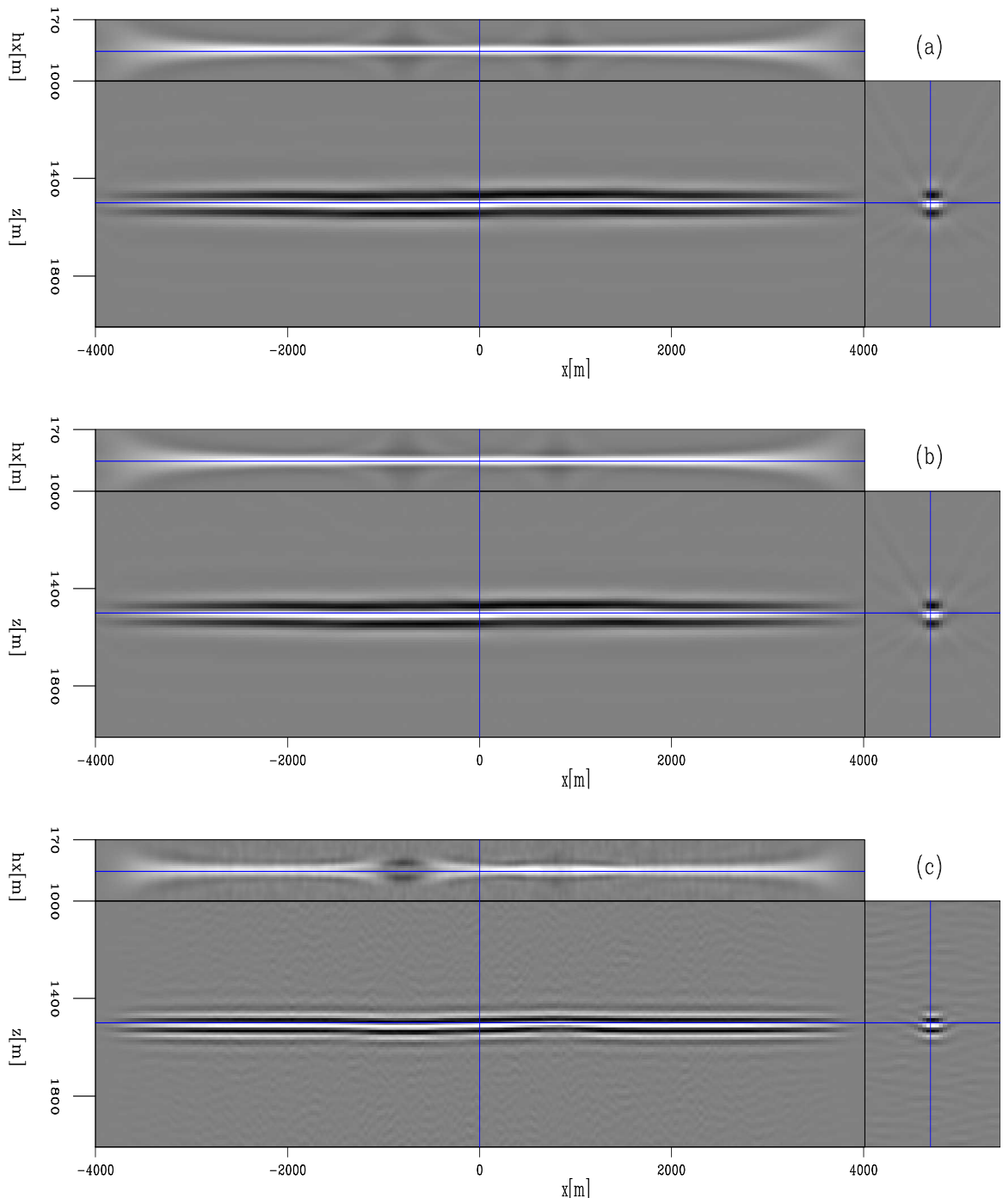


Figure 2: Migrated image cubes with a constant background slowness ($\widehat{s} = 1/2000$ s/m). Panel (a) is the result obtained in the original shot-profile domain; Panel (b) is the result obtained by migrating 61 plane waves, while panel (c) is obtained by migrating 40 image-space encoded areal gathers. [**CR**] yaxun1/. twin-bimg-all

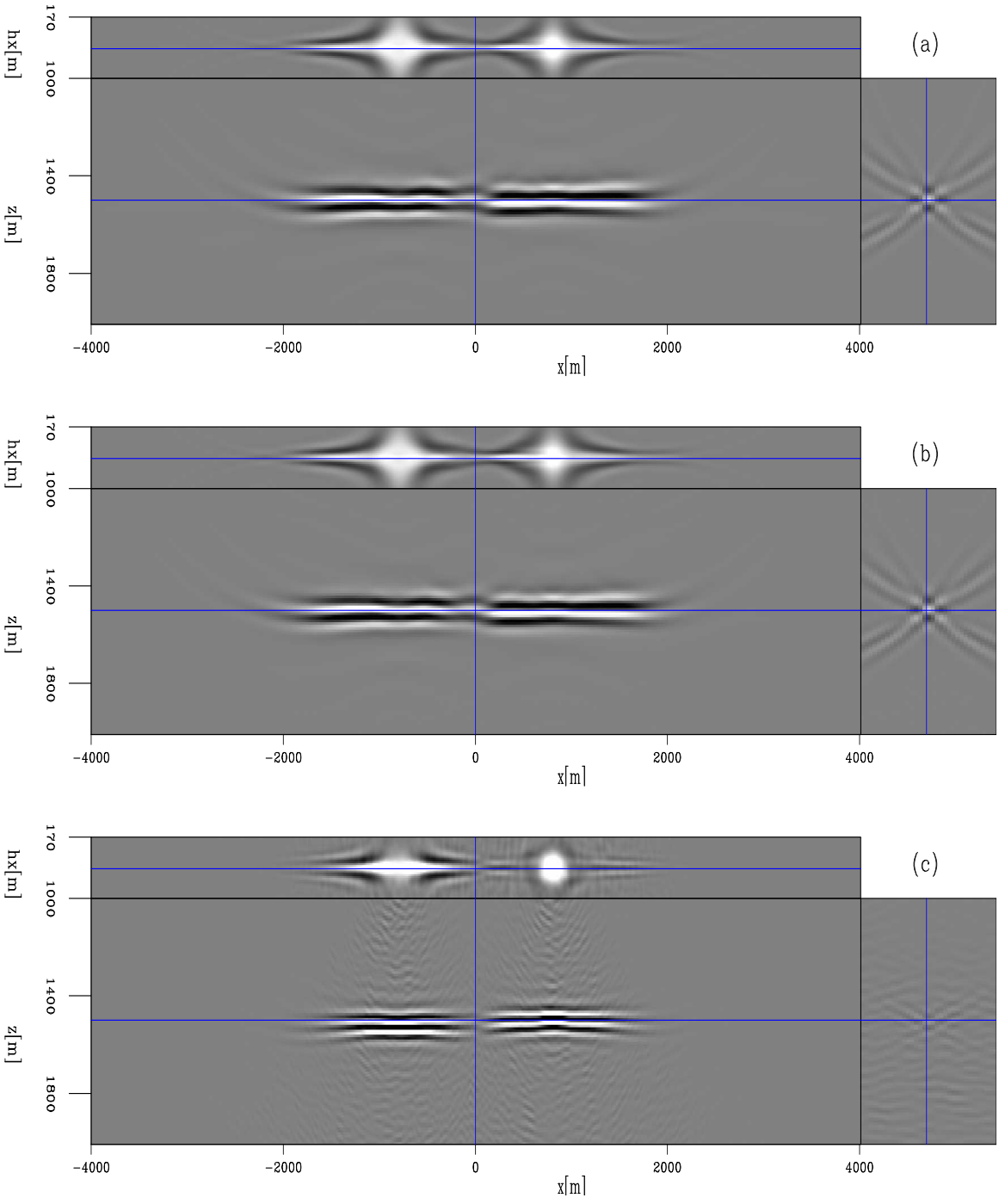


Figure 3: The image perturbations obtained by applying the forward tomographic operator **T** to the correct slowness perturbations in different domains. Panel (a) shows the image perturbation obtained using the original shot gathers, while panels (b) and (c) are obtained using the data-space encoded gathers and image-space encoded gathers, respectively. [**CR**] yaxun1/. twin-dimg-all

pseudo-random noise presented in Figure 3(c).

The final example we show is the comparison among the gradients of the objective functional obtained in different domains. For simplicity, here we compare only the negative DSO gradients ($-\nabla J_{\rm DSO}$) defined by Equation 8 (we compare $-\nabla J_{\rm DSO}$ instead of $\nabla J_{\rm DSO}$, because $-\nabla J_{\rm DSO}$ determines the search direction in a gradient-based nonlinear optimization algorithm). Figure 5 shows the DSO image perturbations computed as follows:

$$\Delta I(\mathbf{x},\mathbf{h}) = |\mathbf{h}|^2 \widehat{I}(\mathbf{x},\mathbf{h}), \tag{38}$$

or in matrix form:

$$\Delta \mathbf{I} = \mathbf{O}'\mathbf{O}\widehat{\mathbf{I}}, \tag{39}$$

where $\mathbf{O}$ is the DSO operator. Figure 5(a) is the result obtained in the original shot-profile domain, whereas Figure 5(b) and (c) are obtained in the data-space phase-encoding domain and the image-space phase-encoding domain, respectively. The coherent energy at non-zero offests are indicators of velocity errors.

Figure 6 shows the negative gradients of the DSO objective functional ($-\nabla J_{\rm DSO}$) obtained by back-projecting the DSO image perturbations shown in Figure 5. For comparison, Figure 6(a) shows the exact slowness perturbation, which is the same as Figure 4(a); Figure 6(b) shows the result obtained in the original shot-profile domain; Figure 6(c) shows the result obtained in the data-space phase-encoding domain, which is almost identical to Figure 6(b); Figure 6(d) shows the result obtained in the image-space phase-encoding domain. The result is also similar to Figure 6(b), though the unattenuated crosstalk and the random noise make the gradient less well behaved than those in Figure 6(b) and (c). Most important, the gradient in Figure 6(d) is pointing towards the correct direction, which is crucial for a gradient-based optimization algorithm to converge to the correct solution.

CONCLUSIONS

We extend the theory of image-space wave-equation tomography to the generalized source domain. One important advantage of this new domain is that we are able to synthesize a much smaller data set while still keeping necessary velocity information for migration velocity analysis; hence the computational cost of performing image-space wave-equation tomography can be significantly reduced. We demonstrate how these new data sets can be generated by using both the data-space phase encoding method and the image-space phase encoding method. Our preliminary tests on a simple synthetic model show that with the synthesized gathers, we are able to obtain a gradient of the tomography objective functional similar to that computed using the original shot gathers, but at significantly lower cost. The correct gradient is thus important for the gradient-based optimization algorithm to converge to the correct solution.

REFERENCES

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: 76th Ann. Internat. Mtg., Expanded Abstracts, 3345–3349, Soc. of Expl. Geophys.

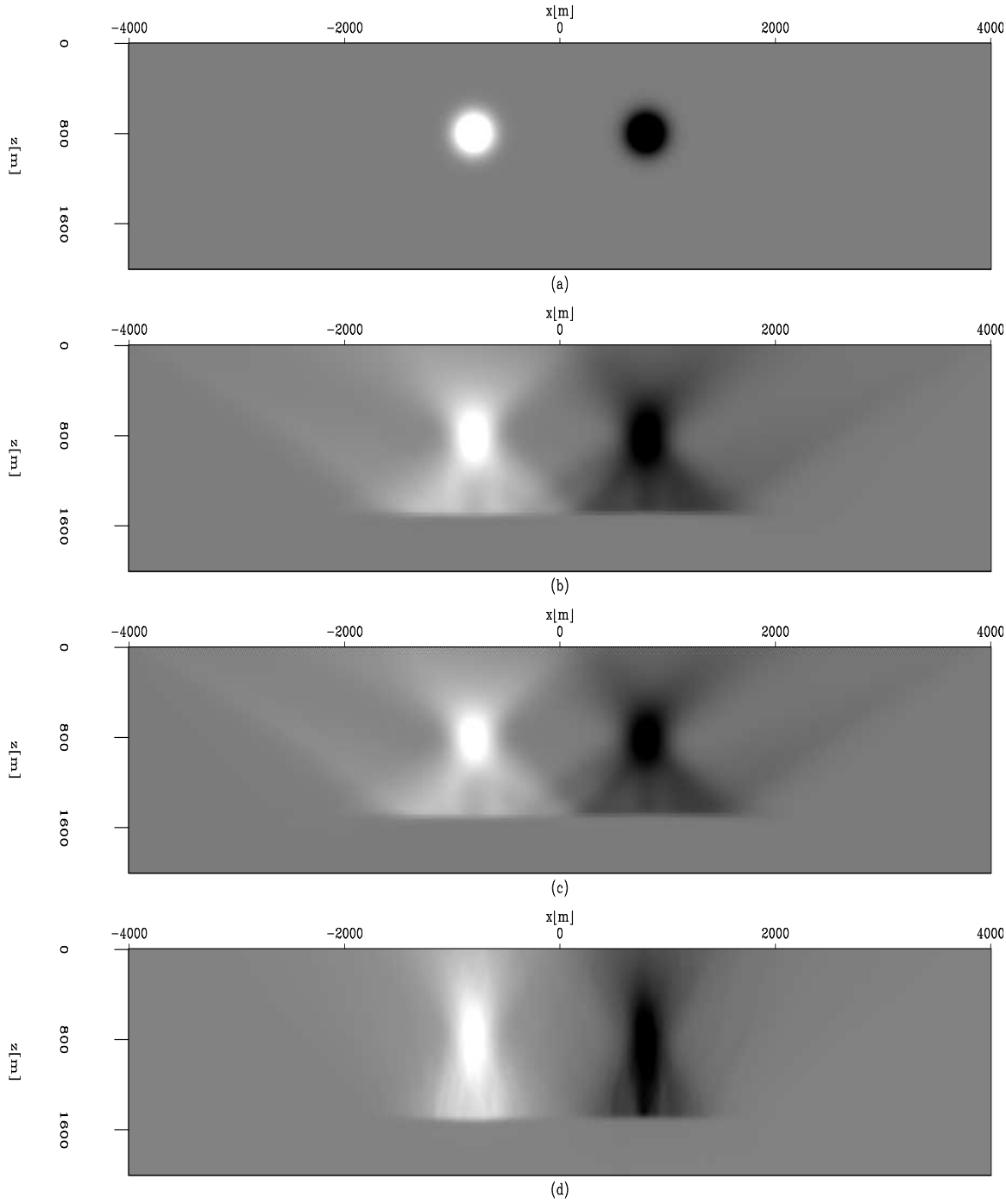


Figure 4: The slowness perturbation obtained by applying the adjoint tomographic operator $\mathbf{T}'$ on the image perturbations in Figure 3. Panel (a) shows the exact slowness perturbation; Panel (b) shows the slowness perturbation estimated by back-projecting the image perturbation shown in Figure 3(a); Panel (c) shows the result obtained using the data-space plane-wave encoded gathers by back-projecting Figure 3(b) and Panel (d) shows the result obtained using the image-space encoded gathers by back-projecting Figure 3(c). [**CR**] yaxun1/. twin-dslw-all

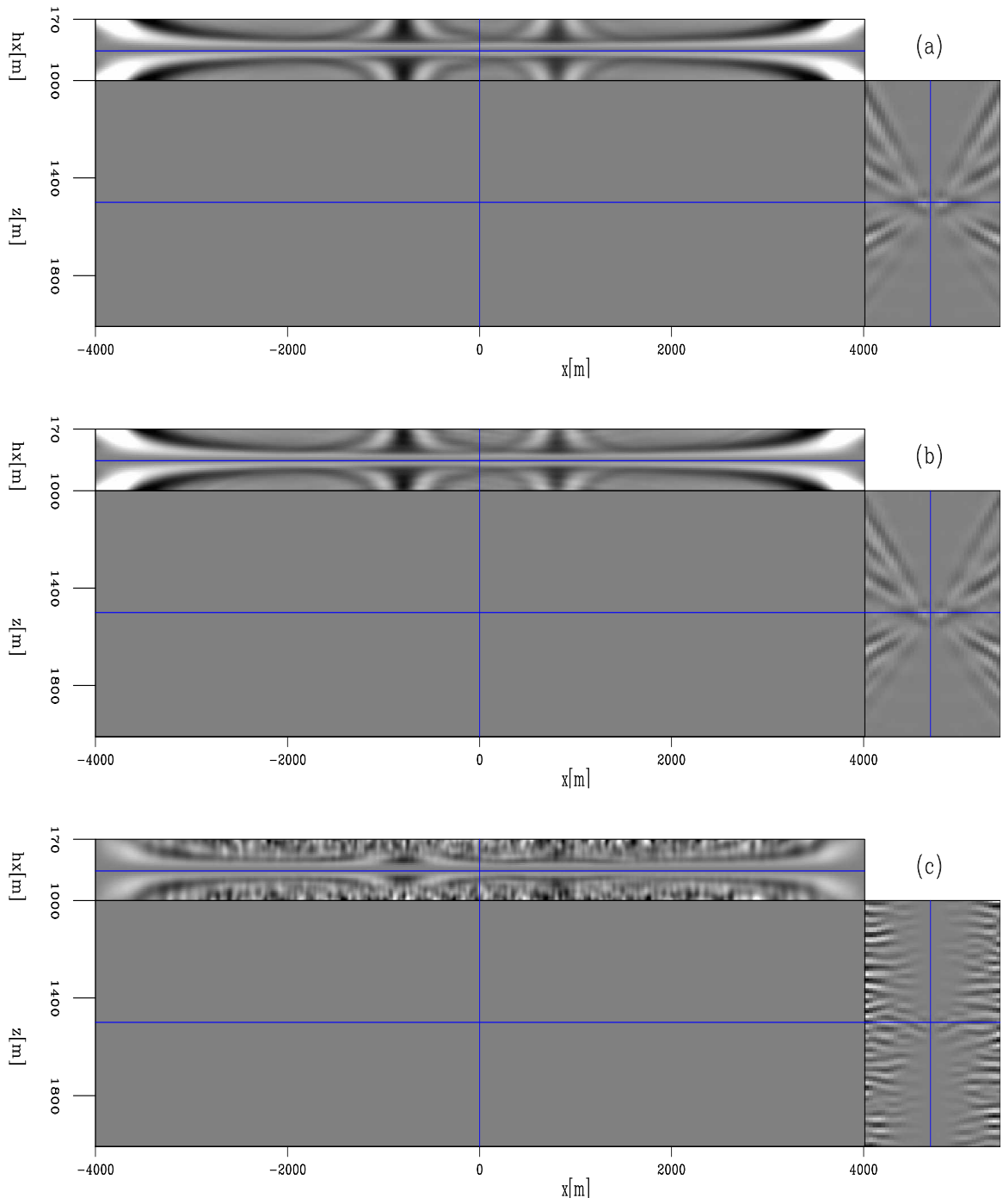


Figure 5: The DSO image perturbations. The coherent energy at non-zero offsets indicates velocity errors. Panel (a) is obtained using the original shot gathers; Panels (b) and (c) are obtained using the data-space encoded gathers and the image-space encoded gathers, respectively. [**CR**] yaxun1/. twin-dimg-offdso-all

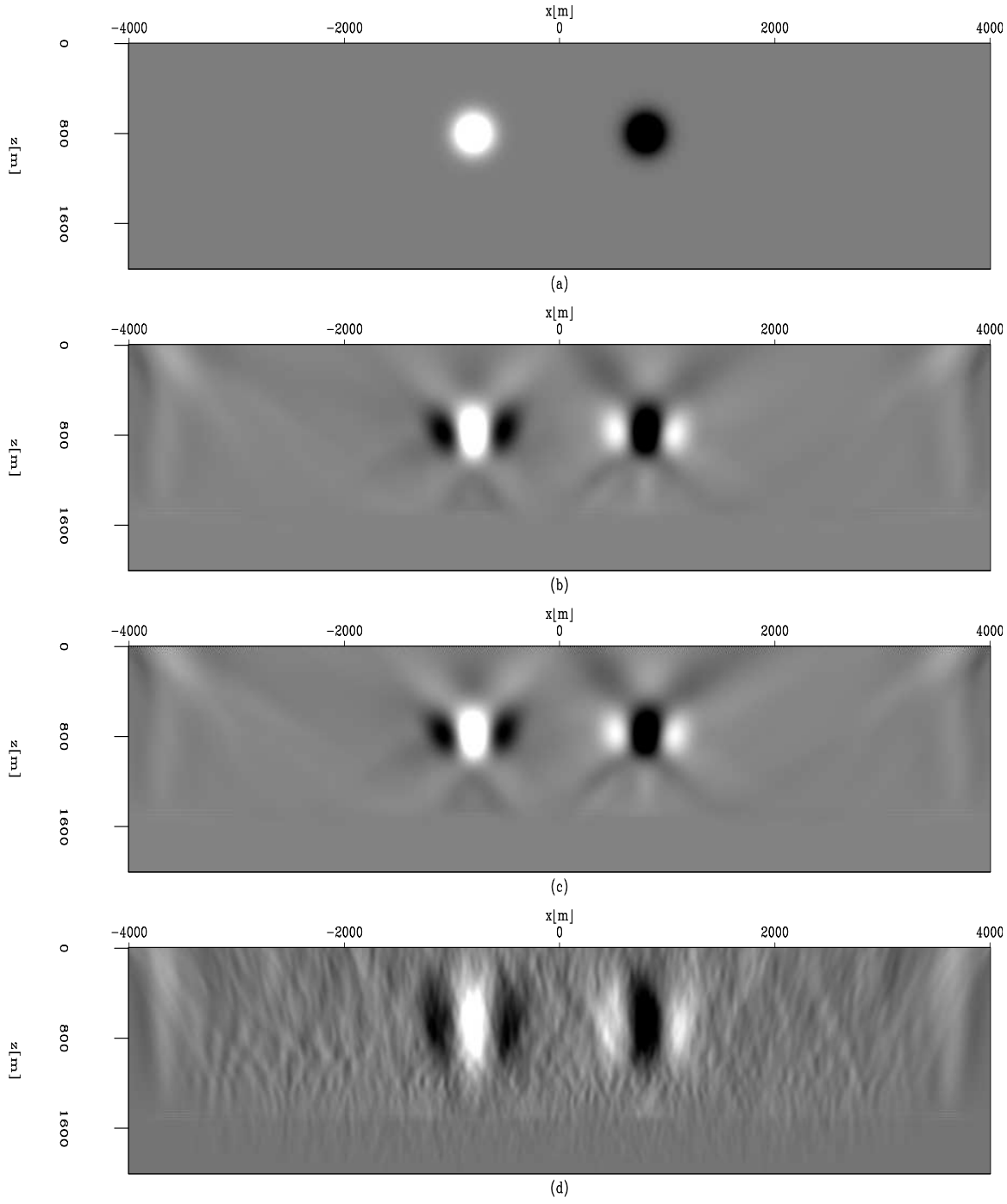


Figure 6: The negative DSO gradients obtained using different methods. Panel (a) shows the exact slowness perurbation; Panel (b) shows the result obtained using the original shot gathers; Panels (c) and (d) show the results obtained using the data-space phase encoded gathers and the image-space phase encoded gathers, respectively. [**CR**] yaxun1/. twin-dslw-offdso-all

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: 76th Ann. Internat. Mtg., Expanded Abstracts, 3056–3060, Soc. of Expl. Geophys.

———, 2007, Prestack modeling of image events for migration velocity analysis: **SEP-131**, 101–118.

———, 2008, Automatic wave-equation migration velocity analysis: **SEP-134**, 65–78.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: 69th Ann. Internat. Mtg., Expanded Abstracts, 1723–1726, Soc. of Expl. Geophys.

Claerbout, J. F., 1971, Towards a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Duquet, B. and P. Lailly, 2006, Efficient 3D wave-equation migration using virtual planar sources: Geophysics, **71**, S185–S197.

Etgen, J. T., 2005, How many angles do we really need for delayed-shot migration: 75th Ann. Internat. Mtg., Expanded Abstracts, 1985–1988, Soc. of Expl. Geophys.

Guerra, C. and B. Biondi, 2008a, Phase-encoding with Gold codes for wave-equation migration: **SEP-136**.

———, 2008b, Prestack exploding reflector modeling: The crosstalk problem: **SEP-134**, 79–92.

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.

Mora, P., 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

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.

Sava, P., 2003, Prestack residual migration in frequency domain: Geophysics, **68**, 634–640.

———, 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**, 593–606.

———, 2004b, Wave-equation migration velocity analysis-II: Examples: Geophysical Prospecting, **52**, 607–623.

Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.

Shen, P., W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot-profile migration: 75th Ann. Internat. Mtg., Expanded Abstracts, 2249–2252, Soc. of Expl. Geophys.

Shragge, J., 2007, Waveform inversion by one-way wavefield extrapolation: Geophysics, **72**, A47–A50.

Tang, Y., 2008, Modeling, migration and inversion in the generalized source and receiver domain: **SEP-136**.

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

Whitmore, N. D., 1995, An Imaging Hierarchy for Common Angle Plane Wave Seismogram: PhD thesis, University of Tulsa.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

Zhang, Y., J. Sun, C. Notfors, S. Grey, L. Chemis, and J. Young, 2005, Delayed-shot 3D depth migration: Geophysics, **70**, E21–E28.

APPENDIX A

This appendix derives the perturbed one-way wave equation with respect to the slowness perturbation. Let us start with the one-way wave equation for the source wavefield as follows:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) D(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ D(x, y, z = 0, \mathbf{x}_s, \omega) = f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s) \end{cases}, \tag{A-1}$$

We can rewrite the slowness and the source wavefield as follows:

$$s(\mathbf{x}) = \widehat{s}(\mathbf{x}) + \Delta s(\mathbf{x}) \tag{A-2}$$

$$D(\mathbf{x}, \mathbf{x}_s, \omega) = \widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega) + \Delta D(\mathbf{x}, \mathbf{x}_s, \omega), \tag{A-3}$$

where $\widehat{s}(\mathbf{x})$ and $\widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega)$ are the background slowness and background wavefield, and $\Delta s(\mathbf{x})$ and $\Delta D(\mathbf{x}, \mathbf{x}_s, \omega)$ are small perturbations in slowness and source wavefield, respectively. If $\Delta s(\mathbf{x})$ is small, then the square root in the first equation of A-1 can be approximated using Taylor expansion as follows:

$$\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2} \approx \sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2} + \frac{\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{\mathbf{s}}^2(\mathbf{x})}}}. \tag{A-4}$$

Substituting Equations A-2, A-3 and A-4 into Equation A-1 and ignoring the second-order terms yield the following linearized one-way wave equation for the perturbed source wavefield:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \Delta D(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{-i\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{s}^2(\mathbf{x})}}} \widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega) \\ \Delta D(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}. \tag{A-5}$$

Similarly, we can also obtain the linearized one-way wave equation for the perturbed receiver wavefield as follows:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \Delta U(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{-i\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{s}^2(\mathbf{x})}}} \widehat{U}(\mathbf{x}, \mathbf{x}_s, \omega) \\ \Delta U(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}. \tag{A-6}$$

APPENDIX B

This appendix demonstrates a matrix representation of the forward tomographic operator $\mathbf{T}$. Let us start with the source wavefield, where the source wavefield $\mathbf{D}_z$ at depth z is downward continued to depth $z + \Delta z$ by the one-way extrapolator $\mathbf{E}_z(\mathbf{s}_z)$ as follows:

$$\mathbf{D}_{z+\Delta z} = \mathbf{E}_z(\mathbf{s}_z)\mathbf{D}_z, \tag{B-1}$$

where the one-way extrapolator is defined as follows:

$$\mathbf{E}_z(\mathbf{s}_z) = e^{-ik_z(\mathbf{s}_z)\Delta z} = e^{-i\sqrt{\omega^2 \mathbf{s}_z^2 - |\mathbf{k}|^2}} \tag{B-2}$$

The perturbed source wavefield at some depth level can be derived from the background wavefield by a simple application of the chain rule to equation B-1:

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{\Delta E}_z(\widehat{\mathbf{s}}_z)\widehat{\mathbf{D}}_z, \tag{B-3}$$

where $\widehat{\mathbf{D}}_z$ is the background source wavefield and $\mathbf{\Delta E}_z$ represents the perturbed extrapolator, which can be obtained by a formal linearization with respect to slowness of the extrapolator defined in Equation B-2:

$$\begin{aligned}\mathbf{E}_z(\mathbf{s}_z) = e^{-ik_z(\mathbf{s}_z)\Delta z} &\approx e^{-i\Delta z\widehat{k}_z} + e^{-i\Delta z\widehat{k}_z}\left(-i\Delta z\left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z \\ &= \mathbf{E}_z(\widehat{\mathbf{s}}_z) + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z\left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z,\end{aligned} \tag{B-4}$$

where $\widehat{k}_z = k_z(\widehat{\mathbf{s}}_z)$ and $\widehat{\mathbf{s}}_z$ is the background slowness at depth z. From Equation B-4, the perturbed extrapolator reads as follows:

$$\mathbf{\Delta E}_z(\widehat{\mathbf{s}}_z) = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z\left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z. \tag{B-5}$$

Substituting Equation B-5 into B-3 yields

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z\left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{D}}_z\mathbf{\Delta s}_z. \tag{B-6}$$

Let us define a scattering operator $\mathbf{G}_z$ that interacts with the background wavefield as follows:

$$\mathbf{G}_z(\widehat{\mathbf{D}}_z, \widehat{\mathbf{s}}_z) = \left(-i\Delta z\left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{D}}_z = \frac{-i\omega\Delta z}{\sqrt{1-\frac{|\mathbf{k}|^2}{\omega^2\widehat{\mathbf{s}}_z^2}}}\widehat{\mathbf{D}}_z. \tag{B-7}$$

Then the perturbed source wavefield for depth $z+\Delta z$ can be rewritten as follows:

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{G}_z(\widehat{\mathbf{D}}_z, \widehat{\mathbf{s}}_z)\mathbf{\Delta s}_z. \tag{B-8}$$

We can further write out the recursive Equation B-8 for all depths in the following matrix form:

$$\begin{pmatrix}\mathbf{\Delta D}_0\\ \mathbf{\Delta D}_1\\ \mathbf{\Delta D}_2\\ \vdots\\ \mathbf{\Delta D}_n\end{pmatrix} = \begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{\Delta D}_0\\ \mathbf{\Delta D}_1\\ \mathbf{\Delta D}_2\\ \vdots\\ \mathbf{\Delta D}_n\end{pmatrix} + \begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{G}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_1 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \cdots & \mathbf{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_n\end{pmatrix}\begin{pmatrix}\mathbf{\Delta s}_0\\ \mathbf{\Delta s}_1\\ \mathbf{\Delta s}_2\\ \vdots\\ \mathbf{\Delta s}_n\end{pmatrix},$$

or in a more compact notation,

$$\mathbf{\Delta D} = \mathbf{E}(\widehat{\mathbf{s}})\mathbf{\Delta D} + \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-9}$$

The solution of Equation B-9 can be formally written as follows:

$$\mathbf{\Delta D} = (\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-10}$$

Similarly, the perturbed receiver wavefield satisfies the following recursive relation:

$$\mathbf{\Delta U}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta U}_z + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{G}_z(\widehat{\mathbf{U}}_z, \widehat{\mathbf{s}}_z)\mathbf{\Delta s}_z, \tag{B-11}$$

where $\mathbf{G}_z(\widehat{\mathbf{U}}_z, \widehat{\mathbf{s}}_z)$ is the scattering operator, which interacts with the background receiver wavefield as follows:

$$\mathbf{G}_z(\widehat{\mathbf{U}}_z, \widehat{\mathbf{s}}_z) = \left(-i\Delta z \left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{U}}_z = \frac{-i\omega\Delta z}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{\mathbf{s}}_z^2}}}\widehat{\mathbf{U}}_z. \tag{B-12}$$

We can also write out the recursive Equation B-12 for all depth levels in the following matrix form:

$$\begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{G}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta s}_0 \\ \mathbf{\Delta s}_1 \\ \mathbf{\Delta s}_2 \\ \vdots \\ \mathbf{\Delta s}_n \end{pmatrix},$$

or in a more compact notation,

$$\mathbf{\Delta U} = \mathbf{E}(\widehat{\mathbf{s}})\mathbf{\Delta U} + \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-13}$$

The solution of Equation B-13 can be formally written as follows:

$$\mathbf{\Delta U} = (\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-14}$$

With the background wavefields and the perturbed wavefields, the perturbed image can be obtained as follows:

$$\begin{pmatrix} \mathbf{\Delta I}_0 \\ \mathbf{\Delta I}_1 \\ \mathbf{\Delta I}_2 \\ \vdots \\ \mathbf{\Delta I}_n \end{pmatrix} = \begin{pmatrix} \widehat{\mathbf{U}}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{U}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{U}}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \widehat{\mathbf{U}}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta D}_0 \\ \mathbf{\Delta D}_1 \\ \mathbf{\Delta D}_2 \\ \vdots \\ \mathbf{\Delta D}_n \end{pmatrix} + \begin{pmatrix} \widehat{\mathbf{D}}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{D}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{D}}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \widehat{\mathbf{D}}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix},$$

or in a more compact notation,

$$\mathbf{\Delta I} = \mathrm{diag}\left(\widehat{\mathbf{U}}\right)\mathbf{\Delta D} + \mathrm{diag}\left(\widehat{\mathbf{D}}\right)\mathbf{\Delta U}. \tag{B-15}$$

Substituting Equations B-10 and B-14 into Equation B-15 yields

$$\begin{aligned} \mathbf{\Delta I} \;=\; & \Big(\mathrm{diag}\left(\widehat{\mathbf{U}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}}) + \\ & \mathrm{diag}\left(\widehat{\mathbf{D}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\Big)\mathbf{\Delta s}, \end{aligned} \tag{B-16}$$

from which we can read the forward tomographic operator $\mathbf{T}$ as follows:

$$\begin{aligned} \mathbf{T} \;=\; & \mathrm{diag}\left(\widehat{\mathbf{U}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}}) + \\ & \mathrm{diag}\left(\widehat{\mathbf{D}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}}). \end{aligned} \tag{B-17}$$

APPENDIX C

This appendix demonstrates a matrix representation of the adjoint tomographic operator $\mathbf{T}'$. Since the slowness perturbation $\mathbf{\Delta s}$ is linearly related to the perturbed wavefields, $\mathbf{\Delta D}$ and $\mathbf{\Delta U}$, to obtain the back-projected slowness perturbation, we first must get the back-projected perturbed wavefields from the perturbed image $\mathbf{\Delta I}$. From Equation B-15, the back-projected perturbed source and receiver wavefields are obtained as follows:

$$\mathbf{\Delta D} = \overline{\text{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I} \tag{C-1}$$

and

$$\mathbf{\Delta U} = \overline{\text{diag}\left(\widehat{\mathbf{D}}\right)}\mathbf{\Delta I}. \tag{C-2}$$

Then the adjoint equations of Equations B-10 and B-14 are used to get the back-projected slowness perturbation $\mathbf{\Delta s}$. Let us first look at the adjoint equation of Equation B-10, which can be written as follows:

$$\mathbf{\Delta s}_D = \mathbf{G}'(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{E}'(\widehat{\mathbf{s}})\left(\mathbf{1} - \mathbf{E}'(\widehat{\mathbf{s}})\right)^{-1}\mathbf{\Delta D}. \tag{C-3}$$

We can define a temporary wavefield $\mathbf{\Delta P}_D$ that satisfies the following equation:

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\left(\mathbf{1} - \mathbf{E}'(\widehat{\mathbf{s}})\right)^{-1}\mathbf{\Delta D}. \tag{C-4}$$

After some simple algebra, the above equation can be rewritten as follows:

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta P}_D + \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta D}. \tag{C-5}$$

Substituting Equation C-1 into equation C-5 yields

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta P}_D + \mathbf{E}'(\widehat{\mathbf{s}})\overline{\text{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I}. \tag{C-6}$$

Therefore, $\mathbf{\Delta P}_D$ can be obtained by recursive upward continuation, where $\mathbf{\Delta D} = \overline{\text{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I}$ serves as the initial condition. The back-projected slowness perturbation from the perturbed source wavefield is then obtained by applying the adjoint of the scattering operator $\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})$ to the wavefield $\mathbf{\Delta P}_D$ as follows:

$$\mathbf{\Delta s}_D = \mathbf{G}'(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{\Delta P}_D. \tag{C-7}$$

Similarly, the adjoint equation of Equation B-14 reads as follows:

$$\mathbf{\Delta s}_U = \mathbf{G}'(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{E}(\widehat{\mathbf{s}})'\left(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}})'\right)^{-1}\mathbf{\Delta U}. \tag{C-8}$$

We can also define a temporary wavefield $\mathbf{\Delta P}_U$ that satisfies the following equation:

$$\mathbf{\Delta P}_U = \mathbf{E}(\widehat{\mathbf{s}})'\left(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}})'\right)^{-1}\mathbf{\Delta U}. \tag{C-9}$$

After rewriting it, we get the following recursive form:

$$\begin{aligned}\mathbf{\Delta P}_U &= \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta P}_U + \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta U} \\ &= \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta P}_U + \mathbf{E}(\widehat{\mathbf{s}})'\overline{\text{diag}\left(\widehat{\mathbf{D}}\right)}\mathbf{\Delta I}.\end{aligned} \tag{C-10}$$

The back-projected slowness perturbation from the perturbed receiver wavefield is then obtained by applying the adjoint of the scattering operator $\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})$ to the wavefield $\mathbf{\Delta P}_U$ as follows:

$$\mathbf{\Delta s}_U = \mathbf{G}'(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta P}_U. \tag{C-11}$$

The total back-projected slowness perturbation is obtained by adding $\mathbf{\Delta s}_D$ and $\mathbf{\Delta s}_U$ together:

$$\mathbf{\Delta s} = \mathbf{\Delta s}_D + \mathbf{\Delta s}_U. \tag{C-12}$$

Phase encoding with Gold codes for wave-equation migration

Claudio Guerra and Biondo Biondi

ABSTRACT

Prestack exploding-reflector modeling aims to synthesize a small dataset comprised of areal shots, while preserving the correct kinematics to be used in iterations of migration velocity analysis. To achieve this goal, the amount of data is reduced by combining the modeled areal data into sets, we call super-areal data. However, crosstalk arises during migration due to the correlation of wavefields resulting from different modeling experiments. Phase encoding the modeling experiments can attenuate crosstalk during migration. In the geophysical community, the most used phase-encoding schemes are plane-wave-phase encoding and random-phase encoding. Here, we exploit the application of Gold codes commonly used in wireless communication, radar and medical imaging communities to phase encode data. We show that adequately selecting the Gold codes can potentially shift the crosstalk out of the migration domain, or the region of interest if a target-oriented approach is used, yielding an image free of crosstalk.

INTRODUCTION

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 computed with wave-equation migration. For migration velocity analysis, the aim is to generate a considerably smaller dataset than the one used in the initial migration, while maintaining the necessary kinematic information to update the velocity.

Conceptually, the synthesized areal data are computed by upward propagating source and receiver wavefields using subsurface-offset-domain common-image gathers (SODCIGs) as initial conditions. To decrease the number of experiments to migrate, we take advantage of the linearity of the wave propagation to combine several experiments into a set of composite records. Combining several experiments gives rise to crosstalk during imaging (Biondi, 2006; Guerra and Biondi, 2008). Guerra and Biondi (2008) use pseudo-random-phase encoding (Romero et al., 2000) during the modeling step to attenuate crosstalk.

It is common, in the exploration geophysics community, to employ pseudo-random codes using intrinsic functions specific to the programing language. These pseudo-random codes present, generally, a uniform distribution. Their autocorrelation and cross-correlation functions have no special properties. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of the same order of magnitude as those of the non-zero lags of the autocorrelation function. Herein, theses codes are called conventional random codes.

In wireless communication, especially for systems using Code Division Multiple Access (CDMA), a class of different pseudo-random codes have been widely used (Shi and Schelgel, 2003). These codes are binary sequences and have unique autocorrelation and cross-correlations properties which make them more suited to achieve the above-mentioned objectives with minimal crosstalk. The autocorrelation function is represented by a large peak, whose amplitude equals the number of samples in the code, and the cross-correlation peaks, at non-zero lags, with the same amplitudes as that of the autocorrelation. Examples of binary pseudo-random codes used by these communities are Golay (Golay, 1961; Tseng, 1972), Kasami (Kasami, 1966) and Gold codes (Gold, 1967). Medical imaging (Gran, 2005) and radar communities (Levanon and Mozeson, 2004) also exploit the statistical properties of these pseudo-random codes to increase bandwidth, signal-to-noise ratio and pulse compression.

Quan and Harris (1991) analyze orthogonal codes to encode simultaneous source signatures for cross-well surveys, and conclude that m-sequences and Gold codes provide the best results on the separation of the seismograms. Here, we exploit the properties of the Gold codes to encode the prestack exploding reflector modeling experiments.

In the next section we give a brief description of the prestack exploding-reflector modeling. Then we discuss how to compute the Gold codes. To illustrate the effectiveness of phase encoding with Gold codes, we compare the migration of prestack-exploding-reflector modeled data encoded with conventional random codes and Gold codes.

PRESTACK EXPLODING-REFLECTOR MODELING

Starting from a prestack image obtained by wave-equation migration represented by a single SODCIG, areal source and receiver wavefields are modeled at the surface by

$$\begin{aligned} S(x,y,z=0,\omega;\mathbf{x}_m) &= G(\mathbf{x}_m-\mathbf{h};x,y,z=0,\omega) * I_s(\mathbf{x},\mathbf{h};\mathbf{x}_m), \\ R(x,y,z=0,\omega;\mathbf{x}_m) &= G(\mathbf{x}_m+\mathbf{h};x,y,z=0,\omega) * I_r(\mathbf{x},\mathbf{h};\mathbf{x}_m), \end{aligned} \tag{1}$$

where $S(x,y,z=0,\omega;\mathbf{x}_m)$ is the source wavefield and $R(x,y,z=0,\omega;\mathbf{x}_m)$ is the receiver wavefield. $I_s(\mathbf{x},\mathbf{h};\mathbf{x}_m)$ and $I_r(\mathbf{x},\mathbf{h};\mathbf{x}_m)$ are the prestack images used as initial conditions for the source and receiver wavefield extrapolation, respectively, at a selected position, $\mathbf{x}_m$. These prestack images should be dip-independent gathers. They are computed by remapping the dip along the offset direction according to the apparent geological dip (Biondi, 2007). $G(\mathbf{x}_m \pm \mathbf{h};x,z=0,\omega)$ represents the operator that extrapolates the wavefields from the subsurface to the surface; $\mathbf{h}$ is the subsurface offset; ω is the temporal frequency; and $\mathbf{x}$ is the vector of spatial coordinates.

In the case of using an one-way extrapolator, the source and receiver wavefields are upward continued according to the one-way wave equations

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) S(\mathbf{x},\omega;\mathbf{x}_m) = I_s(\mathbf{x},\mathbf{h};\mathbf{x}_m) \\ S(x,y,z=z_{\max},\omega;\mathbf{x}_m) = 0 \end{cases}, \tag{2}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) R(\mathbf{x},\omega;\mathbf{x}_m) = I_r(\mathbf{x},\mathbf{h};\mathbf{x}_m) \\ R(x,y,z=z_{\max},\omega;\mathbf{x}_m) = 0 \end{cases}, \tag{3}$$

where $s(\mathbf{x})$ is the slowness at $\mathbf{x}$; $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector.

Using the linearity of the wave propagation, sets of individual modeling experiments can be combined into the same areal data, such that the amount of data input into migration can be significantly decreased, reducing its cost. However, this procedure generates crosstalk when applying the imaging condition during migration.

Guerra and Biondi (2008) introduce strategies to attenuate the crosstalk. Migration of $(\mathbf{x}, \omega)$-random-phase encoded data disperses the crosstalk energy throughout the image as a pseudo-random background noise. By adding more realizations of random-phase encoded areal data, the speckled noise can be further attenuated. The encoded source wavefield, $\widetilde{S}(\mathbf{x}, \mathbf{p}_m, \omega)$, and the encoded receiver wavefield, $\widetilde{R}(\mathbf{x}, \mathbf{p}_m, \omega)$, are synthesized according to

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{S}(\mathbf{x}, \mathbf{p}_m, \omega) = \widetilde{I}_s(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) \\ \widetilde{S}(x, y, z = z_{\max}, \mathbf{p}_m, \omega) = 0 \end{cases}, \tag{4}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{R}(\mathbf{x}, \mathbf{p}_m, \omega) = \widetilde{I}_r(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) \\ \widetilde{R}(x, y, z = z_{\max}, \mathbf{p}_m, \omega) = 0 \end{cases}, \tag{5}$$

where $\widetilde{I}_s(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega)$ and $\widetilde{I}_r(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega)$ are the encoded SODCIGs after rotations. They are defined as follows:

$$\widetilde{I}_s(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) = \sum_{\mathbf{x}_m} I_s(\mathbf{x}, \mathbf{h}, \mathbf{x}_m) \beta(\mathbf{x}, \mathbf{x}_m, \mathbf{p}_m, \omega), \tag{6}$$

$$\widetilde{I}_r(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) = \sum_{\mathbf{x}_m} I_r(\mathbf{x}, \mathbf{h}, \mathbf{x}_m) \beta(\mathbf{x}, \mathbf{x}_m, \mathbf{p}_m, \omega), \tag{7}$$

where $\beta(\mathbf{x}, \mathbf{x}_m, \mathbf{p}_m, \omega) = e^{i\gamma(\mathbf{x}, \mathbf{x}_m, \mathbf{p}_m, \omega)}$ is the phase-encoding function; the variable $\mathbf{p}_m$ is the index of different realizations of phase encoding.

The areal shot migration is performed by downward continuation of the areal source and receiver wavefields according to the following one-way wave equations

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widehat{S}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widehat{S}(x, y, z = 0, \mathbf{p}_m, \omega) = \widetilde{S}(x, y, z = 0, \mathbf{p}_m, \omega) \end{cases}, \tag{8}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widehat{R}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widehat{R}(x, y, z = 0, \mathbf{p}_m, \omega) = \widetilde{R}(x, y, z = 0, \mathbf{p}_m, \omega) \end{cases}, \tag{9}$$

where the encoded source wavefield, $\widetilde{S}(\mathbf{x}, \mathbf{p}_m, \omega)$, and the encoded receiver wavefield, $\widetilde{R}(\mathbf{x}, \mathbf{p}_m, \omega)$, are used as boundary conditions.

The image, $\widehat{I}(\mathbf{x}, \mathbf{h})$, is obtained by cross-correlation of the source wavefield, $\widehat{S}(\mathbf{x}, \mathbf{p}_m, \omega)$, with the receiver wavefield, $\widehat{R}(\mathbf{x}, \mathbf{p}_m, \omega)$

$$\widehat{I}(\mathbf{x}, \mathbf{h}) = \sum_{\omega} \sum_{\mathbf{p}_m} \widehat{S}^{\star}(\mathbf{x} - \mathbf{h}, \mathbf{p}_m, \omega) \widehat{R}(\mathbf{x} + \mathbf{h}, \mathbf{p}_m, \omega), \tag{10}$$

where * represents complex conjugation.

GOLD CODES

Before describing Gold codes it is useful to define maximum length sequences.

Linear feedback shift registers (LFSR) (Moon and Stirling, 2000) are called state machines, whose components and functions are:

- the **shift register** – shifts the bit pattern and registers the output bit; and
- the **feedback function** – computes the input bit according to the tap sequence and inserts the computed bit into the input bit position.

The output sequence of bits form pseudo-random binary sequences, which are completely controlled by the tap sequence. A tap sequence defines which bits in the current state will be combined to determine the input bit for the next state. The combination is generally performed using module-2 addition (*exclusive or* – XOR). This means that adding the selected bit values defined by the tap sequence, if the sum is odd the output of the function is one; otherwise the output is zero. Table 1 shows the internal states and the output sequence of a 4-bit LFSR with tap sequence $[4, 1]$. For the current state, the input bit (bit 1) is computed by the sum module–2 of the bits defined by the tap sequence(bits 1 and 4). The rest of the bits in the register (bits 2, 3 and 4) are obtained by shifting the bit values in the previous state to the right. For example, bit 2 of the current state is bit 1 of the previous state. The output bit of the current state is the last bit (bit 4) in the register from the previous state.

Register States				
bit1 (Tap)	bit2	bit3	bit4 (Tap)	**Output Sequence**
1	1	0	1	
0	1	1	0	1
0	0	1	1	0
1	0	0	1	1
0	1	0	0	1
0	0	1	0	0
0	0	0	1	0
1	0	0	0	1
1	1	0	0	0
1	1	1	0	0
1	1	1	1	0
0	1	1	1	1
1	0	1	1	1
0	1	0	1	1
1	0	1	0	1
1	1	0	1	0

Table 1: Internal states and the output sequence of a 4-bit LFSR with tap sequence [4,1].

In number theory, Galois fields (**GF**) are finite fields in which all operations result in an element of the field (Lidl and Niederreiter, 1994). Addition, subtraction and multiplication

of polynomials are defined in a finite field. Module 2 arithmetic forms the basis of **GF**(2) (Galois field of order 2). Addition and multiplication operations in GF(2) can be represented by bitwise operators XOR and AND, respectively. Table 2 synthesizes the possible output values of addition and multiplication over **GF**(2).

+	0	1		×	0	1
0	0	1		0	0	0
1	1	0		1	0	1

Table 2: **GF**(2) addition and multiplication possible outcomes.

If a polynomial can not be represented as the product of two or more polynomials, it is called an irreducible polynomial. For instance, x^2+x+1 is irreducible over **GF**(2) because it can not be factored. However, x^2+1 is not irreducible over **GF**(2) because, using normal algebra, $(x+1)(x+1) = x^2+2x+1$, and after reduction module–2, is x^2+1 (the term $2x$ is dropped). So, in **GF**(2), $x^2+1 \equiv x^2+2x+1$.

The importance of studying irreducible polynomials **GF**(2) is that they are used to represent tap sequences. Considering an irreducible polynomial, the corresponding tap sequence is given by the exponents of the terms with coefficients of 1 (Dinan and Jabbari, 1998). Special tap sequences can be used to generate particular pseudo-random binary sequences. They are called maximum length sequences (m-sequences) and, by definition, are the largest codes that can be generated by a LFSR for a given tap sequence. Their length is (b^n-1), where n is the number of elements of the tap sequence, and $b = 2, 3$ *or* 5.

The autocorrelation function of an m-sequence, $\Phi_{mls}(k)$, is given by

$$\Phi_{mls}(k) = \begin{cases} b^n - 1 & \text{for} \quad k = 0, \\ -1 & \text{for} \quad k \neq 0, \end{cases} \tag{11}$$

where k is the lag of correlation. In spite of the good autocorrelation properties, m-sequences, in general, are not immune to cross-correlation problems, and they may have large and unpredictable cross-correlation values. However, the so-called preferred pairs of m-sequences have cross-correlation functions which might assume the predicted values, -1, $-1+p$, and $-1-p$, where $p = 2^{(n+1)/2}$ for n odd or $p = 2^{(n+2)/2}$ for n even. Given a (2^n-1)–length m-sequence, $a(k)$, and $gcd\{n, 4\} = 1$ (greatest common divisor of n and 4), its preferred pair is the result of decimation computed by applying on $a(k)$ a circular shift of q samples, where $q = 2m+1$ and $gcd\{m, n\} = 1$. Figure 1 shows the cross-correlation of m-sequences that form a preferred pair computed with $m = 5$.

The number of possible preferred pairs of m-sequences is limited, when compared to the requirements of practical applications of wireless communication. Preferred pairs of m-sequences, however, are used to generate Gold codes (Dinan and Jabbari, 1998).

In CDMA, Gold codes are used as chipping sequences that allow several callers to use the same frequency, resulting in less interference and better utilization of the available bandwidth. Originally proposed by Gold (1967), Gold codes can be computed by module-2 addition (*exclusive or*) of circularly shifted preferred pairs of m-sequences of length 2^n-1. The autocorrelation function of a Gold code, $\Phi_{gc}(k)$, is given by

$$\Phi_{gc}(k) = \begin{cases} \pm 2^n - 1 & \text{for} \quad k = 0, \\ \pm 1 & \text{for} \quad k \neq 0. \end{cases} \tag{12}$$

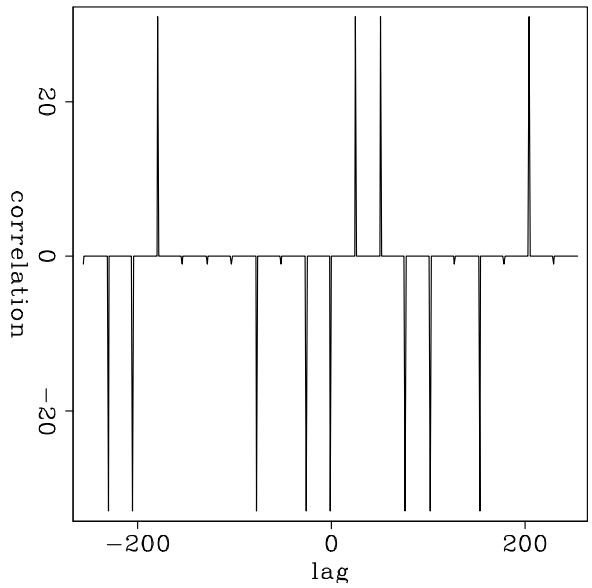


Figure 1: Cross-correlation of a m-sequences that form a preferred pair.[**ER**] claudio1/. prefpairs

More interestingly, the two valued cross-correlation function of Gold sequences, $\Psi_{gc}(k)$, is given by

$$\Psi_{gc}(k) = \begin{cases} \pm(2^n - 1) & \text{for} \quad k = \lambda, \\ \mp 1 & \text{for} \quad k \neq \lambda. \end{cases} \tag{13}$$

where the correlation lag λ is given by the difference between the number of circular shifts applied to the m-sequence to compute the Gold codes.

Figure 2 illustrates the correlation properties of the Gold codes. The left part shows the autocorrelation of the Gold code generated with one circular shift of the preferred pair of m-sequence. The right part shows the cross-correlation of the Gold code generated with one circular shift of the preferred pair of m-sequence with the Gold code generated with 84 circular shifts. Note that the peak of the cross-correlation occurs at lag 84. For comparison, we show the correlation functions of conventional random codes in Figure 3. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of the same order of magnitude as those of the non-zero lags of the autocorrelation function.

After computing Gold codes, we use their phase information to encode the modeling experiments.

EXAMPLES

We illustrate the use of the encoding methods on a simple model of a flat reflector, 0.5 km deep, embedded in a medium with a constant velocity of 2 km/s. The original data is migrated with a 5% slower velocity. We used the same slower velocity to perform the

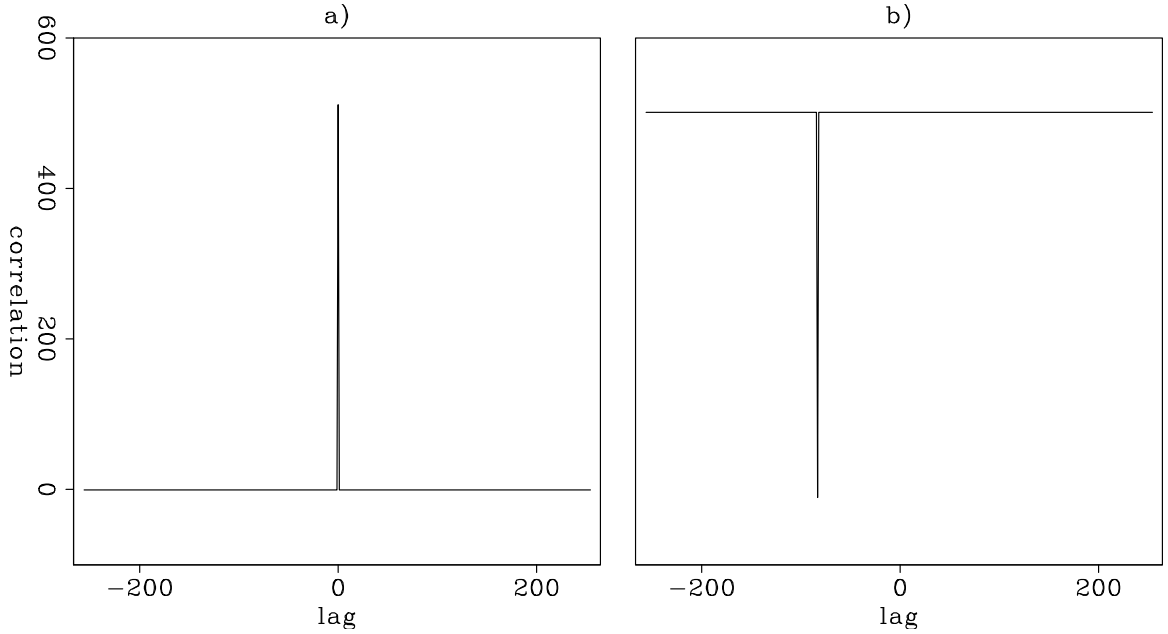


Figure 2: Correlation functions of Gold codes. a) Autocorrelation of Gold code generated with one circular shift of the preferred pair of m-sequence. b) Cross-correlation of the Gold code generated with one circular shift of the preferred pair of m-sequence with that generated with 84 circular shifts. The peak of the cross-correlation occurs at lag -84.[**ER**] claudio1/. gold184

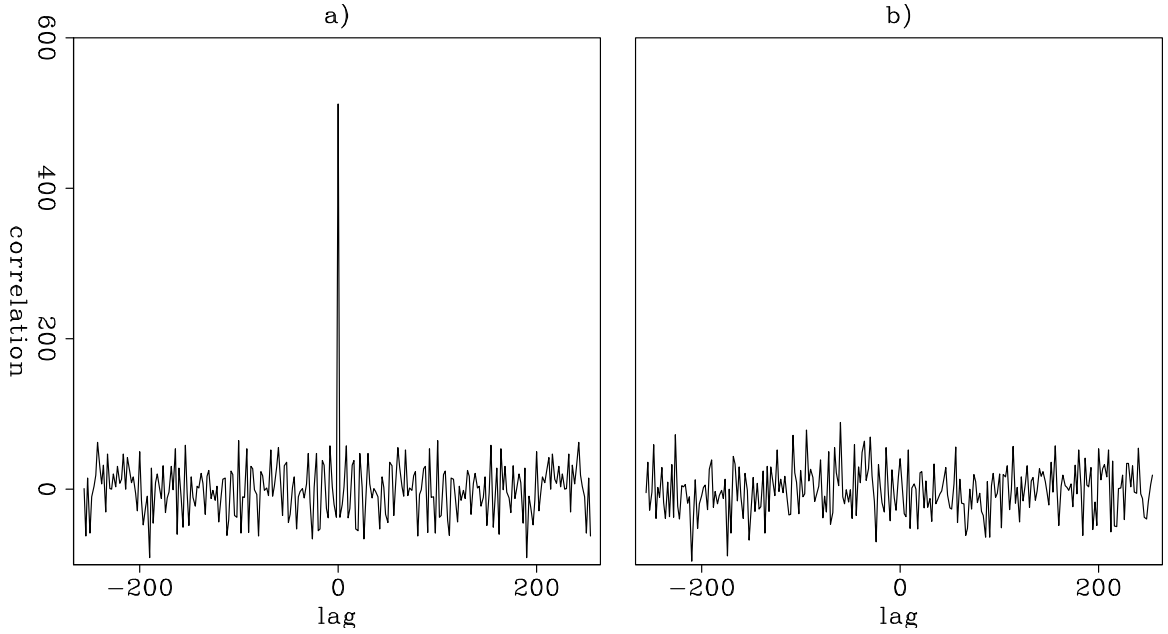


Figure 3: Correlation functions of conventional random codes. a) Autocorrelation of the conventional random code. b) A cross-correlation.[**ER**] claudio1/. rand

modeling. For the areal shot migration, we show examples of migrating Gold code phase-encoded data with the slower velocity. For comparison of the crosstalk behavior, when migrating with a velocity different from that used to model, we also show images migrated with a 5% faster velocity. Super-areal data are comprised of the collection of 10 modeling experiments initiated at every 10th CMP coordinate. We computed the prestack migrated image with 61 subsurface offsets to observe how crosstalk is shifted when using Gold codes. It should have been reasonable to use a much smaller number of offsets, as the moveout information is restricted to the 21 central offsets.

Figure 4 shows the areal shot migration of data generated by the prestack exploding-reflector modeling without combining the modeling experiments into super-areal shots. The panel on the left is the zero-subsurface-offset section, and the panel on the right is a SODCIG. This result represents the ideal image we would like to obtain if the crosstalk could be eliminated. Our objective in phase encoding the modeling experiments is to achieve satisfactory crosstalk attenuation in such a way that the moveout information is not altered.

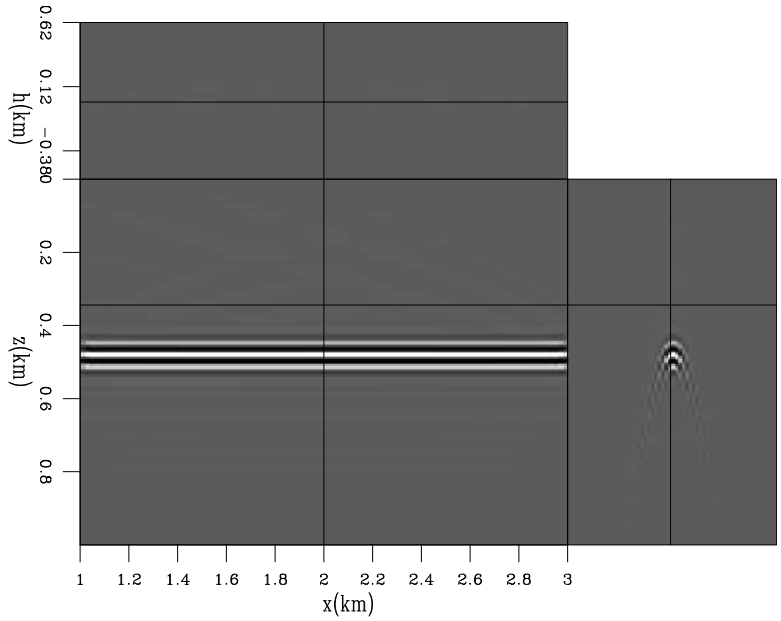


Figure 4: Areal shot migration of synthesized data with no combination of the modeling experiments into super-areal data.[**CR**] claudio1/. perm0

Figure 5 shows the areal shot migration of data generated by the prestack exploding-reflector modeling with no phase encoding applied. The super-areal data, input to areal shot migration, are comprised of modeling experiments initiated at every tenth SODCIG. The SODCIGs resulting from the areal shot migration, show strong crosstalk at subsurface offsets different from zero. The crosstalk is periodic with a period half of the spacing of the modeling experiments in a super-areal shot.

Figure 6 shows areal shot migration of one realization of phase encoding modeling with conventional random codes. The strong crosstalk observed in Figure 4 is now dispersed throughout the image. The dispersed crosstalk can be further attenuated by migrating more random realizations, but this increases the cost of migration. Figure 7 shows the

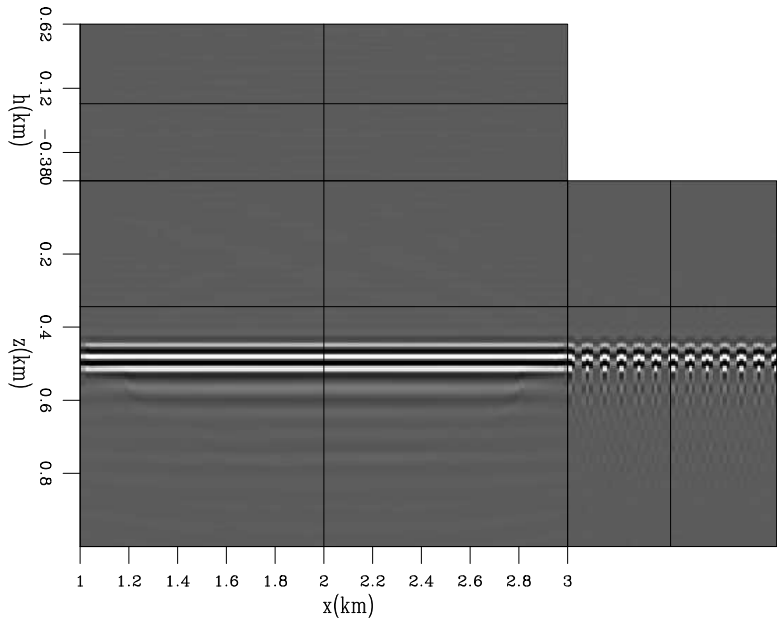


Figure 5: Areal shot migration of synthesized data with no phase encoding applied. The super-areal data comprises 10 modeling experiments. Notice the crosstalk in the SODCIG.[**CR**] claudio1/. perma

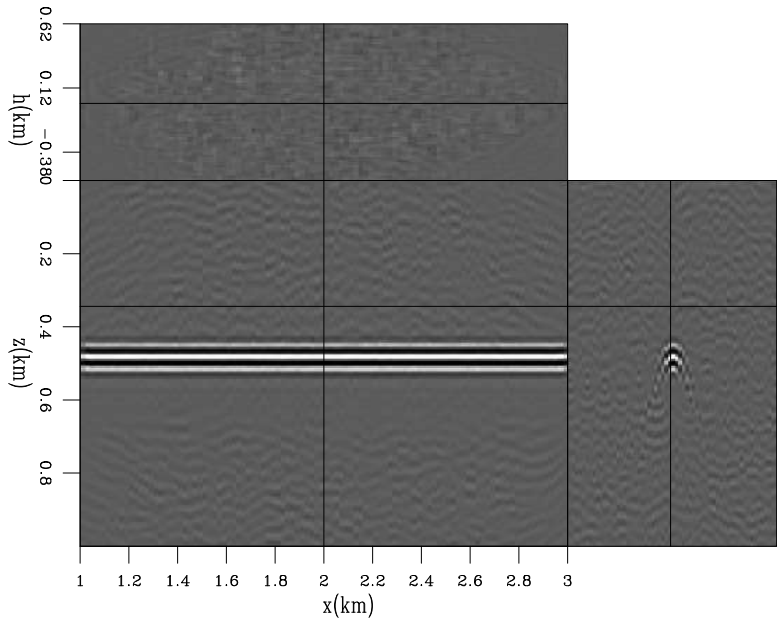


Figure 6: Areal shot migration of one realization of synthesized data with conventional random-phase encoding.[**CR**] claudio1/. conv1r

migration of 4 realizations of conventional random encoding modeling.

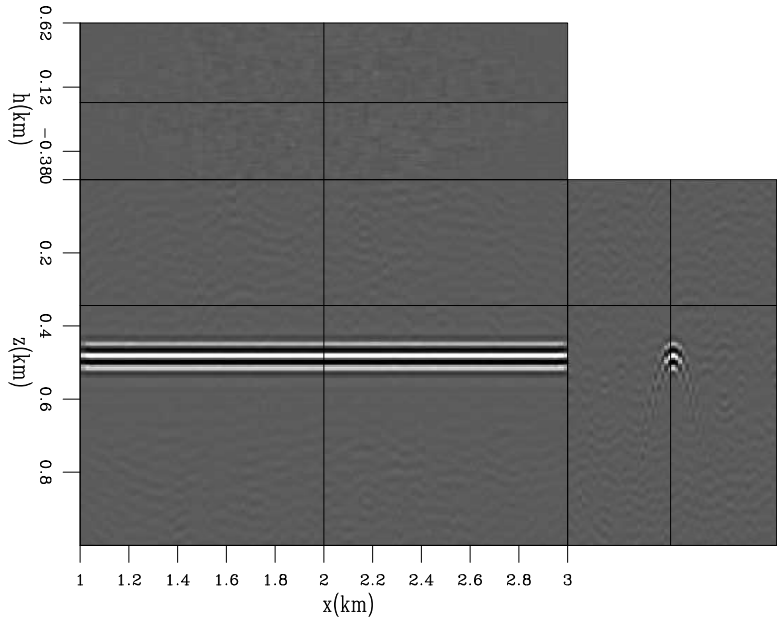


Figure 7: Areal shot migration of four realizations of synthesized data with conventional random-phase encoding.[**CR**] claudio1/. conv5r

Crosstalk attenuation is incomplete when using conventional random codes because their autocorrelations are not a perfect spike, nor are their cross-correlations zero. Gold codes partially satisfy these requirements: the autocorrelation is almost a perfect spike, except for -1's at lags different from zero, and, similarly, the cross-correlations are -1 everywhere except where they peak. Therefore, to obtain good results when using Gold codes, it is critical to select the codes which provide the best crosstalk attenuation. That is because the cross-correlation functions have peaks with the same magnitude as those of the autocorrelation function.

In Figure 8, the areal shot migration was performed on encoded data with Gold codes that have cross-correlation peaks at every 5th (Figure 8a), 10th (Figure 8b) and, 20th lags (Figure 8c). This means that when applying the multi-offset imaging condition, unrelated wavefields, delayed in time by the phase functions, will cross-correlate at depths different from that of the related wavefields, which were encoded with the same phase. The crosstalk of the Figure 5 is now shifted in depth according to the selected set of Gold codes. In Figure 8a, the apexes of the crosstalk are displaced in depth at a constant spacing of, approximately, 0.1 km; in Figure 8b the spacing is approximately 0.2 km, and in Figure 8c, 0.4 km. Notice that in spite of the crosstalk is still present in Figure 8c, its complete elimination can be achieved with a simple depth-windowing of the image.

The amount of depth shift of the crosstalk is defined by the lag where the cross-correlation of the Gold codes peaks. For the simple case of constant velocity, the depth shifts, δz, are given by

$$\delta z = \frac{n_\lambda v}{2 n_\omega d_\omega} \tag{14}$$

where n_ω is the number of frequencies, d_ω is the frequency interval and n_λ is the lag where

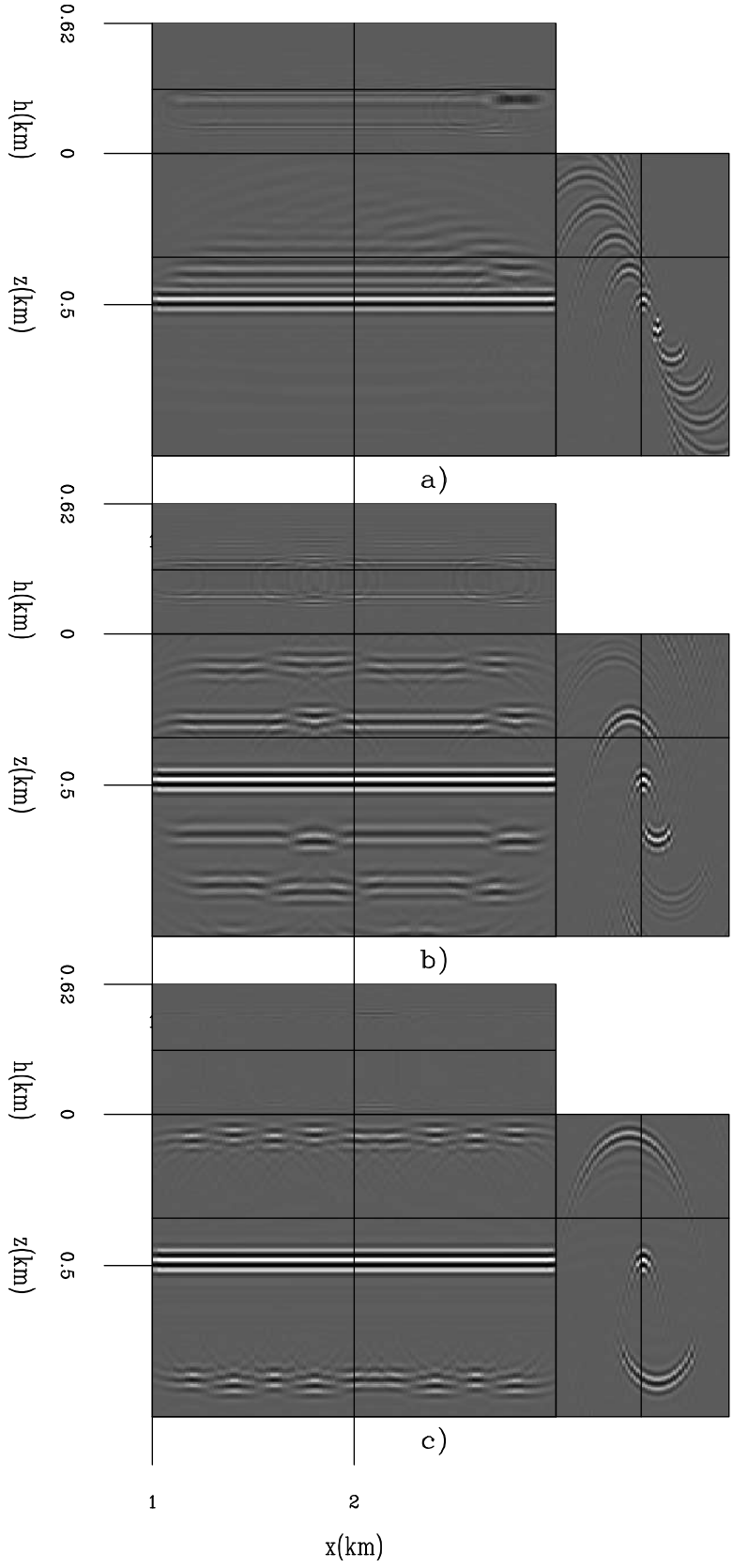


Figure 8: Areal shot migration of one realization of synthesized data with Gold code phase encoding. Gold codes have cross-correlation peaks at every 5th (a), 10th (b) and, 20th lag (c), and the depth shifts are, respectively, 0.1 km, 0.2 km, and 0.4 km.[**CR**] claudio1/. gold1x

the cross-correlation of the Gold codes peaks. For the present example, this amounts to $\delta z = 0.021 * n_\lambda$ km.

One possibility to statistically attenuate the crosstalk is to randomly select the Gold codes. Figure 9 shows the areal shot migration of one realization of encoded data with randomly selected Gold codes. The crosstalk shows different patterns than that of the sequentially selected Gold codes.

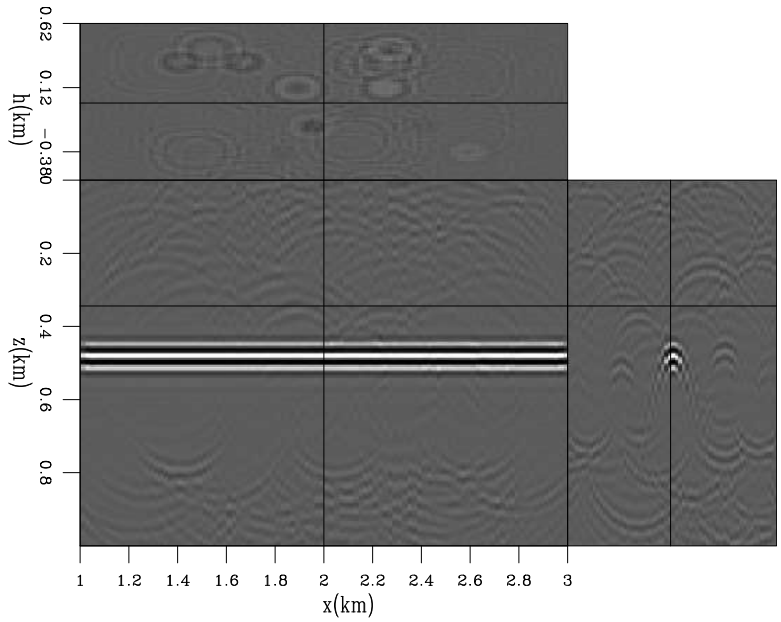


Figure 9: Areal shot migration of one realization of synthesized data with Gold code phase encoding. Gold codes are randomly selected.[**CR**] claudio1/. gold1r

As before, migrating more realizations of randomly selected Gold-encoded data further attenuates crosstalk. Figure 10 shows the migration of 4 realizations of randomly selected Gold-encoded data. Comparison with Figure 9 shows that much of the remaining crosstalk energy has been attenuated. However, this approach does not exploit the statistical properties of the Gold phase functions and the crosstalk shows up locally coherent, at random positions. This strategy is definitely not suited to provide a kinematically reliable image.

Considering that the crosstalk is shifted in depth, as observed in Figure 8, and that the amount of shift is determined by the lag where the cross-correlation of the Gold codes peaks (equation 14), one can choose the Gold codes according to a suitable interval that completely shifts the crosstalk away from the zone of interest. This strategy shares similar idea as the linear-phase encoding Romero et al. (2000), which aims to shift the crosstalk out of the migration domain by using a linear function of frequency. In the Appendix, we show that phase encoding with Gold codes is equivalent to linear phase encoding. Figure 11 shows the areal shot migration of data encoded by selecting every 50th Gold code, meaning that the depth shifts are multiples of 1 km, as predicted by equation 14, which should be adequate to completely push the crosstalk away from the SODCIG. Contrary to what was expected, the crosstalk is still present. Of course, its complete elimination can be achieved by windowing around the central traces (Figure 12).

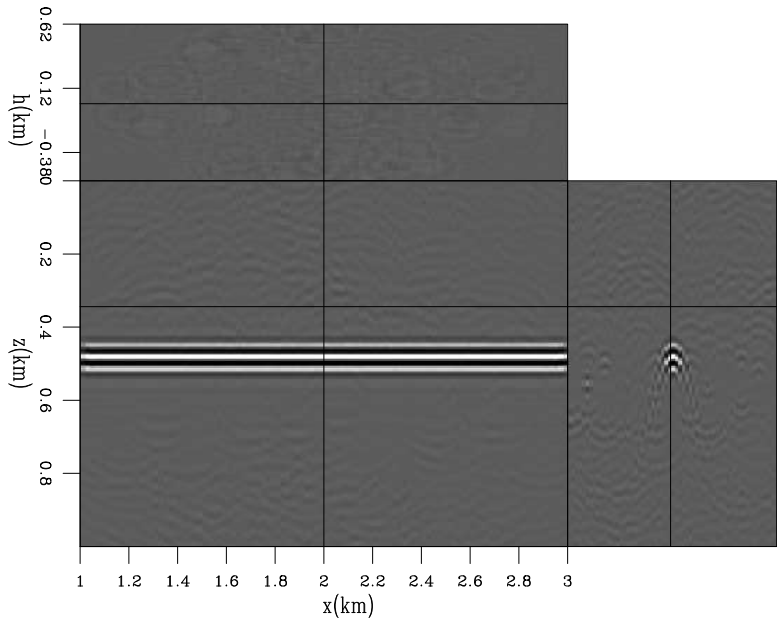


Figure 10: Areal shot migration of four realizations of synthesized data with Gold code phase encoding. Gold codes are randomly selected.[**CR**] claudio1/. gold5r

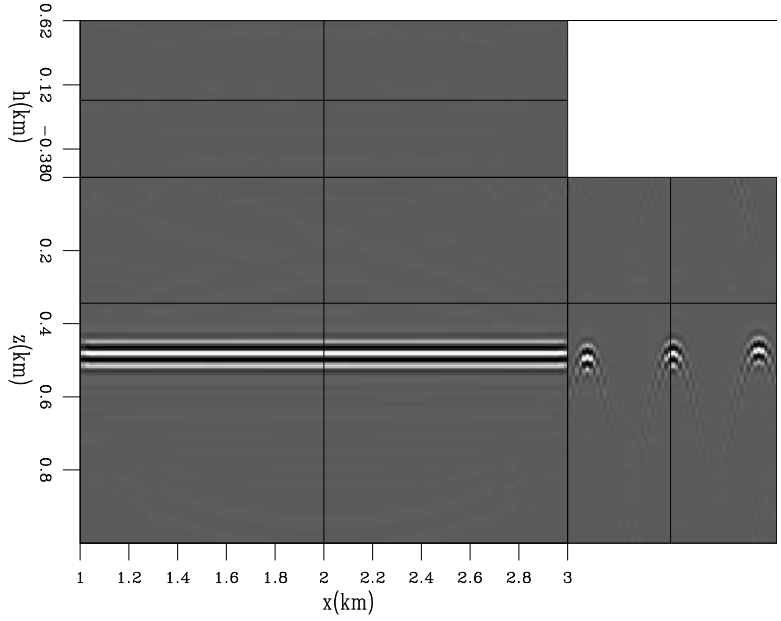


Figure 11: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are selected such that the crosstalk is shifted away from the reflector. A simple windowing should completely eliminate the crosstalk.[**CR**] claudio1/. gold1o

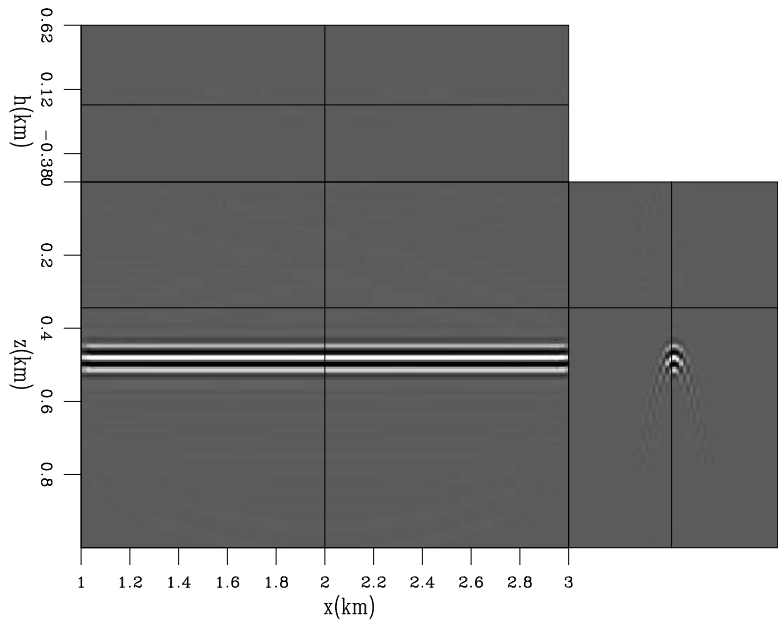


Figure 12: Crosstalk of Figure 11 can be completely eliminated by windowing the SODCIG. Compare with Figure 5.[**CR**] claudio1/. goldwd

To understand the origin of the crosstalk in Figure 11, we migrated the data without adding the results of individual migrations of the super areal data. Figure 13 shows three of these migrations. The three panels show the same spatial position. The top panel shows the migration of super areal data whose modeling includes a SODCIG coincident with that displayed. Remember that we are using a "comb" function to select every 10th SODCIG to initiate the modeling experiments comprised by a super areal data. The central panel shows the migration of the super areal data modeled from the initial image shifted five CMP positions away from that of the uppermost panel, and the panel on the bottom shows the migration of the super areal data modeled from the initial image shifted nine CMP positions away from that of the uppermost panel. The modeling experiments which generated the super areal data for Figure 13c are actually separated one CMP position from the ones for Figure 13a, given that the modeling experiments in super areal data are separated every 10th CMP. Figure 13a and Figure 13c show that much of the reflector energy comes from the migration of modeling experiments initiated at CMP positions close to the considered image point. Figure 13b shows that much of the crosstalk energy of Figure 11 is related to migration of super areal shots whose modeling experiments are initiated at SODCIGs shifted five CMP positions with respect to the considered SODCIG.

In migration velocity analysis, at every velocity iteration, data is migrated with an updated velocity. To verify how the crosstalk behaves when migrating with a different velocity, we used a 5% faster velocity to migrate the Gold encoded data modeled with the initial 5% slower velocity. We used sets of Gold codes with cross-correlation peaks at every 5th (Figure 14a), 10th (Figure 14b) and, 20th lag (Figure 14c) lags. The depth shift of the crosstalk is proportional to the increase in velocity. The behavior is similar to that of migration with a slower velocity, regarding that the interference with the region of interest decreases with the increase of the lag of the cross-correlation peaks.

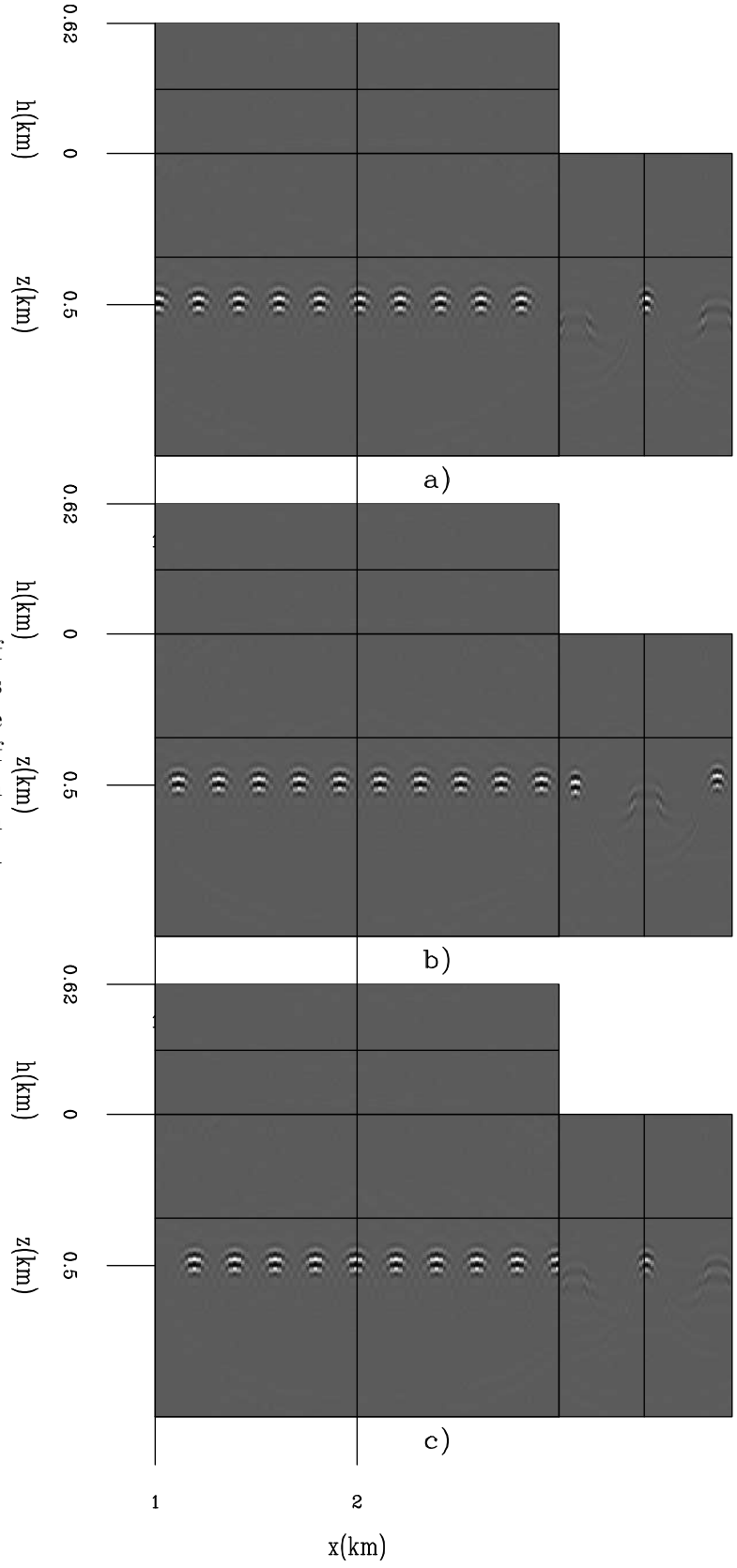


Figure 13: Areal shot migration of three super areal data. The results were not stacked together. Notice that much of the crosstalk energy of Figure 11 is related to the migration of super areal shots (b) other than the one which modeling was initiated at that CMP position (a and c).[**CR**]

claudio1/. goldsp

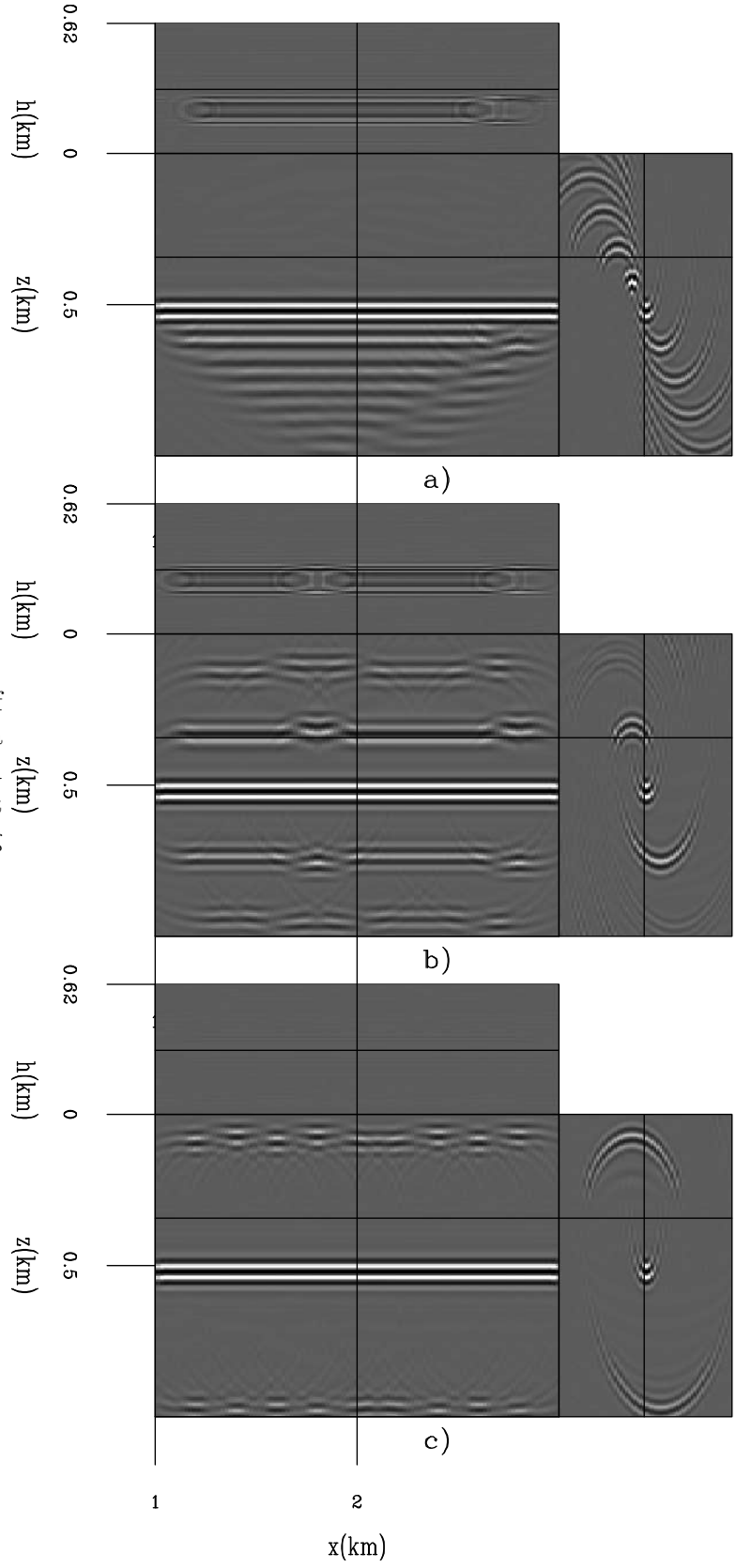


Figure 14: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes have cross-correlation peaks at every 5th (a), 10th (b) and, 20th lag (c).[**CR**] claudio1/. gold2x

The use of Gold codes can be less costly than conventional random codes, given that just one realization is necessary to achieve an almost perfect result, while even using more realizations of conventional random encoding does not produce an image with similar quality. In addition, it is suited to be used in a horizon based strategy to migration velocity analysis, where a few important reflectors are chosen to do velocity update.

However, an inadequate choice of the Gold codes is potentially more dangerous than using conventional random codes. Because the crosstalk may not be shifted out of the region of interest, coherent artifacts can coincide with reflectors and obliterate the moveout information. Conventional random codes, in turn, randomizes the crosstalk. This can lead to a noisy residual moveout scan or a noisy gradient. In a companion paper Tang et al. (2008) show that the prestack exploding-reflector modeled data, phase encoded with conventional random codes, is able to provide a reasonable direction for velocity update.

CONCLUSION

We showed that the use of Gold codes in phase encoding can virtually eliminate crosstalk if the codes are satisfactorily selected. Its advantage over conventional random codes is not only in the image quality but also cost. The elimination of the crosstalk can be achieved with only one realization of Gold codes, while for conventional random codes more realizations are needed to obtain reasonable images, but still of lower quality when compared to the ones obtained by using Gold codes. The need for more realizations increases the cost of modeling and migration.

We also showed that the phase of the cross-correlation of Gold codes is given by the lag in time where the cross-correlation peaks multiplied by frequency. This amounts to apply linear phase encoding to the encoded wavefields.

We used a simple example with just one reflector to illustrate the encoding techniques. The performance of the Gold codes in the presence of more than one reflector needs to be tested. However, as just one realization is sufficient to drastically reduce the crosstalk, they can be affordable to use in a horizon-based approach to the modeling in which each reflector is modeled separately. Additional research is needed to analyze the performance of Gold codes in more complex situations, where the depth shift of the crosstalk may not be a simple function of velocity.

ACKNOWLEDGEMENTS

We thank Jerry Harris and Youli Quan for the suggestion of using codes applied in wireless communications.

REFERENCES

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analyis: SEP-Report, **124**, 45–60.

———, 2007, Prestack modeling of image events for migration velocity analysis: SEP-131, 101–118.

Dinan, E. and B. Jabbari, 1998, Spreading codes for direct sequence CDMA and wideband CDMA cellular networks: IEEE Communications Magazine, **36**, 48–54.

Golay, M., 1961, Complementary sequences: IRE Trans. on Info.Theory, **IT-7**, 82–87.

Gold, R., 1967, Optimal binary sequences for spread spectrum multiplexing: IEEE Transactions on Information Theory, **14**, 619–621.

Gran, F., 2005, Spatio-temporal encoding in medical ultrasound imaging: PhD thesis, Technical University of Denmark.

Guerra, C. and B. Biondi, 2008, Prestack exploding reflector modeling: The crosstalk problem: SEP-134, 79–91.

Kasami, T., 1966, Weight distribution formula for some class of cyclic codes.

Levanon, N. and E. Mozeson, 2004, Radar Signals: John Wiley & Sons.

Lidl, R. and H. Niederreiter, 1994, Introduction to Finite Fields and their Applications: Cambridge Univeristy.

Moon, T. K. and W. C. Stirling, 2000, Mathematical Methods and Algorithms for Signal Processing: Prentice Hall.

Quan, Y. and J. M. Harris, 1991, Orthogonal coded signals as simultaneous source signatures: Report of Seismic Tomography Project at Stanford University, **2**, H1–H14.

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

Shi, Z. and C. Schelgel, 2003, Spreading code construction for CDMA: IEEE Communications letters, **1**, 4–6.

Tang, Y., C. Guerra, and B. Biondi, 2008, Image-space wave-equation tomography in the generalized source domain: SEP-136.

Tseng, C., 1972, Complementary sets of sequences: IEEE Trans. on Info.Theory, **IT-18**, 644–652.

APPENDIX

We show that phase encoding using Gold codes is equivalent to the linear phase encoding introduced by (Romero et al., 2000). The cross-correlation function of Gold codes is given by equation 13. It corresponds to a spike of amplitude $\pm 2^n$ at lag λ, which depends on the difference between the number of circular shifts applied to the m-sequences to compute the Gold codes, plus a DC term, ∓ 1. The phase of the cross-correlation function is given by the phase of the spike, $t_\lambda \omega$, and is equal to the phase difference of the input signals. If Gold codes have phases $\gamma_1(\omega)$ and $\gamma_2(\omega)$, the phase of their cross-correlation is

$$\gamma_1(\omega) - \gamma_2(\omega) = t_\lambda \omega. \tag{A-1}$$

According to equation A-1, the phase of the cross-correlation of Gold codes is a linear function of the frequency. Equation A-1 is equal to equation 26 of Romero's paper.

An image-focusing semblance functional for velocity analysis

Biondo Biondi

ABSTRACT

Analyzing the focusing and defocusing of migrated images provides valuable velocity information that can supplement the velocity information routinely extracted from migrated common-image gathers. However, whereas qualitative focusing analysis is readily performed on ensemble of images generated by prestack residual migration, quantitative focusing analysis remains a challenge. I use two simple synthetic-data examples to show that the maximization of a minimum-entropy norm, a commonly-used measure of image focusing, yields accurate estimates for diffracted events, but it can be misleading in the presence of continuous but curved reflectors.

I propose to measure image focusing by computing coherency across structural dips, in addition to coherency across aperture/azimuth angles. Images can be efficiently decomposed according to structural dips during residual migration. I introduce a semblance functional to measure image coherency simultaneously across the aperture/azimuth angles and the dip angles. Using 2D synthetic data examples, I show that the simultaneous evaluation of semblance across aperture-angles and dips can be effective in quantitatively measuring image focusing and also avoiding the biases induced by reflectors' curvature.

INTRODUCTION

Even a superficial analysis of depth migrated seismic images obtained with different migration velocities clearly shows that velocity information could be extracted by measuring image focusing along the spatial dimensions (i.e. horizontal axes and depth). This information is particularly abundant in areas where complex structure and discontinuous reflectors reveal lack of focusing caused by velocity errors; such as in presence of faults, point diffractors, buried channels, uncomformities or rough salt/sediment interfaces.

If we were able to extract this focusing-velocity information reliably from migrated images it could supplement the velocity information that we routinely extract by analyzing residual moveout along the offsets or aperture-angles axes, and thus enhance velocity estimation by increasing resolution and decrease uncertainties. It would be particularly useful to improve the interpretability of the final image and the accuracy of time-to-depth conversion in areas where the reflection aperture range is narrow either because of unfavorable depth/offset ratio or because of the presence of fast body in the overburden (e.g. salt bodies) that deflect the propagating waves. Today, the most common application of image focusing is to migration-velocity scans for subsalt imaging (Wang et al., 2006). However, current practical applications exploit the image-focusing information by using subjective interpretation criteria instead of quantitative measurements (Sava et al., 2005). This limitation makes almost impossible to automate the process and potentially reduces its reliability, and thus it is a serious obstacle to its extensive application.

Minimum entropy has been often proposed as a quantitative measure of image focusing, starting with Harlan et al. (1984), De Vries and Berkhout (1984), and more recently by Stinson et al. (2005) and Fomel et al. (2007). Minimizing the "spatial entropy" measured on image windows privileges images that consist of isolated spikes. If the reflectivity function consists of isolated diffractors, minimum entropy is a good indicator of image focusing. However, field data are usually a combination of diffracted events, specular reflections from planar reflectors, and reflections for high-curvature reflectors. In these cases minimum entropy may yield bias estimates unless the diffractions are successfully separated from the other events before performing the analysis (Fomel et al., 2007). In complex geology, this separation can be unreliable when performed in the data space, and even more challenging when performed in the image space because it is biased by the initial migration velocity. In the following section I show that in presence of reflector curvature (e.g. a sinusoidal reflector) measuring focusing by minimum entropy leads to under-migrated images of convex reflectors (e.g. an anticline,) and over-migrated images of concave reflectors (e.g. a syncline.)

I aim to overcome these shortcomings by generalizing the conventional concept of semblance commonly used in velocity analysis. In addition to measuring semblance along the reflection-aperture angle (or offset for Kirchhoff migration,) as is routinely done, I propose to measure semblance along the structural-dip axes. In this paper I work with 2D data, and thus I compute semblance on 2D patches (structural dip and aperture angle.) With 3D full-azimuth data, semblance would be computed on 4D patches (indexed by two structural dips, reflection aperture and reflection azimuth.)

The proposed method can be applied to locally select the best-focused image among an ensemble of images obtained with different migration velocities. I use residual prestack depth migration in the angle domain (Biondi, 2008) to generate this ensemble of images starting from prestack depth-migrated image in the angle domain. Stolt prestack depth migration could be used as well to perform residual prestack migration (Sava, 2003). With either choice of residual migration, the image decomposition according to structural dip is easily performed within the residual prestack migration process, since both migrations require the image to be transformed into the spatial Fourier domain. The final goal, not addressed by this paper, is to use the image focusing information to enhance interval-velocity estimation for depth migration. In particular, I plan to update the interval-velocity model by using the wave-equation migration velocity analysis method starting from a spatially-varying field of optimal-focusing parameters (Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Sava, 2004).

THE CHALLENGE OF QUANTIFYING IMAGE FOCUSING

In this section, I introduce two simple synthetic data sets that illustrate the opportunities and challenges of measuring image focusing for velocity analysis. I start by showing how the application of a minimum entropy functional can help to determine the correct migration velocity, but also it can mislead the estimation.

Figure 1a and Figure 2a show the reflectivity functions assumed to generate the two data sets. The first one contains a strong diffractor and two dipping planar reflectors broken by a fault. Focusing analysis of the diffractor and the reflectors' truncations provides velocity information additional to the one available by conventional analysis of the reflections from the planar interfaces. The second model consists of a continuous sinusoidal reflector. It

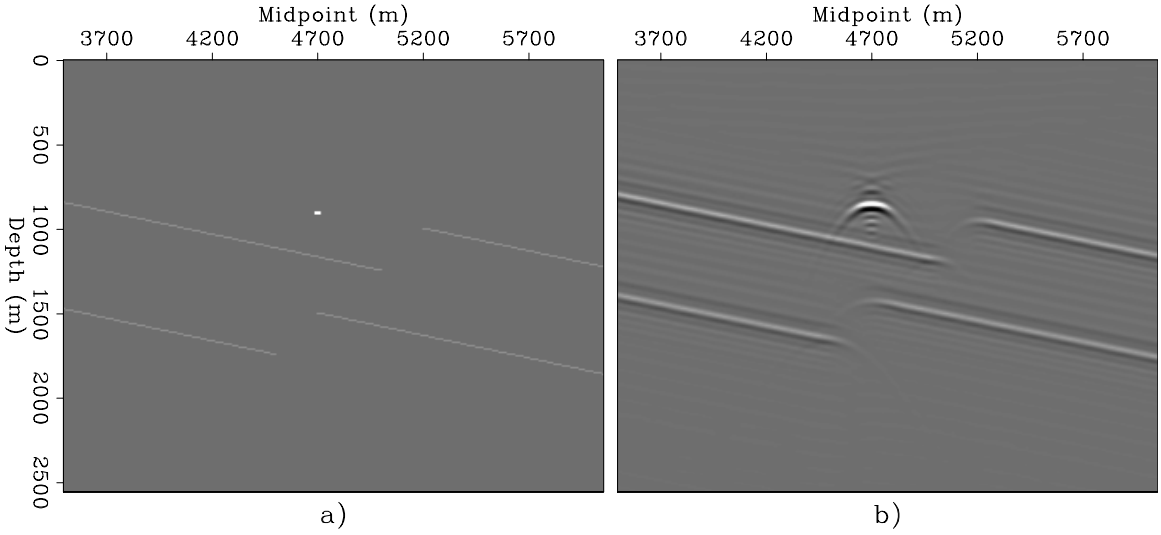


Figure 1: (a) Reflectivity function assumed to compute a synthetic prestack data set, and (b) the stacked section obtained by migrating the data set with a low velocity. [**CR**] biondo1/. Refl-Mig-fault-overn

shows the potential pitfalls of measuring image focusing in presence of curvature in the structure. Figure 1b and Figure 2b show the result of migrating with a low velocity the modeled data corresponding to the reflectivity functions shown in Figure 1a and Figure 2a, respectively. In Figure 1b both the image of the point diffractor and the image of the reflectors' truncations show the typical signs of undermigration; that is, not fully collapsed diffracted events. In Figure 2b the bottom of the syncline shows triplication that are signs of undermigration, whereas the top of the anticline does not show any clear defocusing problems.

Pitfalls of Minimum Entropy functional

Minimizing the image entropy measured on moving spatial windows is a well-known approach to measuring image focusing. The varimax norm (Wiggins, 1985) is commonly used to measure the "entropy" of an image instead of the conventional entropy functional. The varimax norm is cheaper to evaluate than the conventional entropy functional because it does not require the evaluation of a logarithmic function. A peak in the varimax corresponds to a point of minimum entropy. I computed the varimax for local windows extracted from image ensembles computed by applying residual prestack migration to an initial prestack migration performed with a low velocity.

I define $\mathbf{R}\left(\mathbf{x},\gamma,\rho\right)$ as an ensemble of prestack images obtained by residual prestack migration where the parameter ρ is the ratio between the new migration velocity and the migration velocity used for the initial migration. The aperture angle is γ and $\mathbf{x}=\{z,x\}$ is the vector of spatial coordinates, where z is depth and x is the horizontal location.

I define the image window $\bar{\mathbf{x}}$ as:

$$\bar{\mathbf{x}}:\left\{\bar{z}-\Delta z\leq z\leq\bar{z}+\Delta z,\bar{x}-\Delta x\leq x\leq\bar{x}+\Delta x\right\}, \tag{1}$$

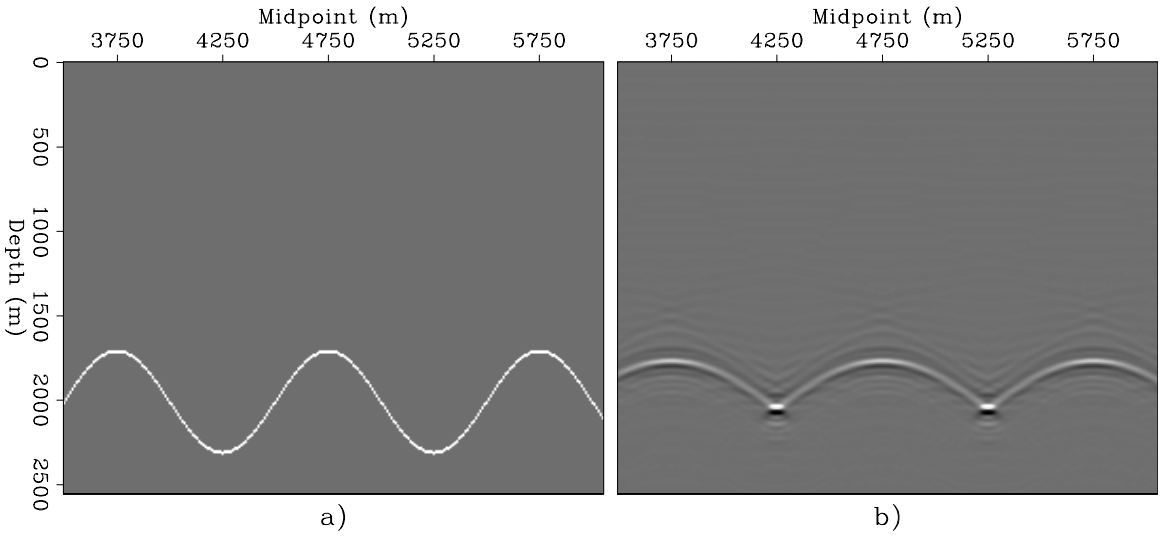


Figure 2: (a) Reflectivity function assumed to compute a synthetic prestack data set, and (b) the stacked section obtained by migrating the data set with a low velocity. [**CR**] biondo1/. Refl-Mig-sinus-overn

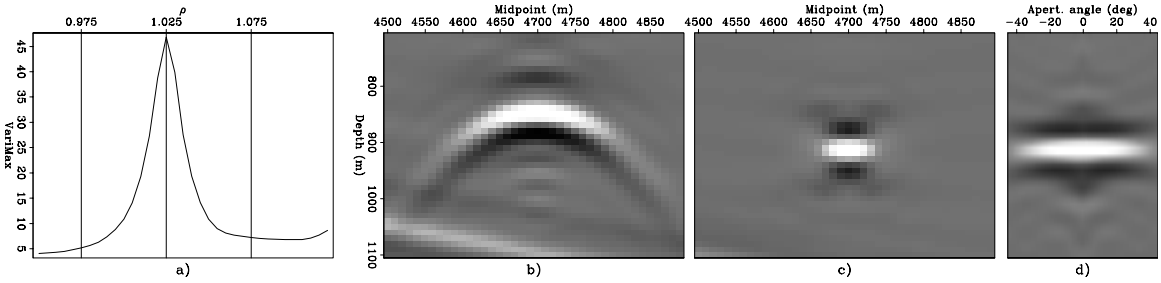


Figure 3: (a) Graph of the varimax norm as a function of ρ, (b) stacked section for $\rho = 1$, (c) stacked section for $\rho = 1.025$, and (d) angle-domain common image gather for $\rho = 1.025$ at $x = 4,700$ meters. [**CR**] biondo1/. Probl-4700-diffr-overn

where $2\Delta z$ is the height of the window and $2\Delta x$ is its width, and $\bar{z}$ and $\bar{x}$ are the coordinates of the window's center.

The varimax norm computed for $\bar{\mathbf{x}}$ is defined as:

$$E_{\mathbf{x}}(\rho) = \frac{N_{\bar{\mathbf{x}}} \sum_{\bar{\mathbf{x}}} \left[\sum_{\gamma} \mathbf{R}(\mathbf{x}, \gamma, \rho) \right]^4}{\left\{ \sum_{\bar{\mathbf{x}}} \left[\sum_{\gamma} \mathbf{R}(\mathbf{x}, \gamma, \rho) \right]^2 \right\}^2}, \tag{2}$$

where $\sum_{\mathbf{x}}$ signifies summation over all the image points in $\bar{\mathbf{x}}$ and $N_{\bar{\mathbf{x}}}$ is the number of points in $\bar{\mathbf{x}}$. Notice that the varimax in equation 2 includes stacking over the aperture angle γ.

For the first data set (Figure 1,) I computed the varimax in equation 2 as a function of ρ in two windows: the first centered on the point diffractor, the second centered on the reflector truncation. Figure 3 shows the following four plots for the point-diffractor window:

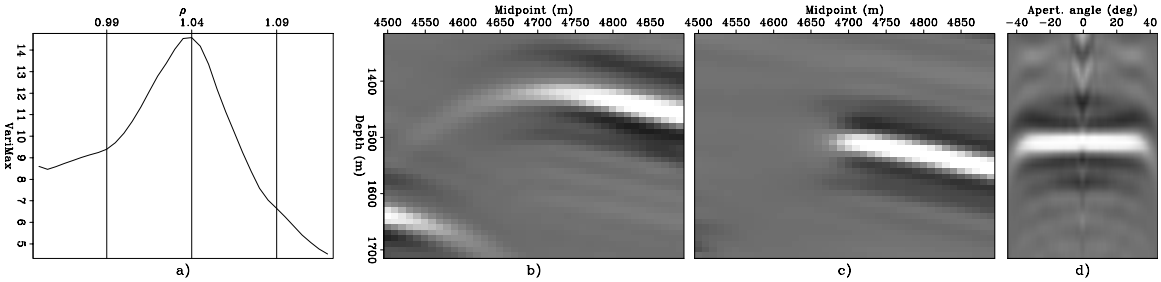


Figure 4: (a) Graph of the varimax norm as a function of ρ, (b) stacked section for $\rho = 1$, (c) stacked section for $\rho = 1.04$, and (d) angle-domain common image gather for $\rho = 1.04$ at $x = 4,700$ meters. [**CR**] biondo1/. Probl-4700-trunc-overn

a) the graph of the varimax norm as a function of ρ, b) the stacked section for $\rho = 1$; that is, the window of the initial undermigrated section in Figure 1b, c) the stacked section for $\rho = 1.025$; that is; for the peak of the curve shown in Figure 3a, and d) the angle-domain common image gather for the same value of $\rho = 1.025$ and extracted from the prestack cube at the horizontal location of the point diffractor.

Figure 4 shows analogous plots as the ones shown in the previous figure, but for the reflector-truncation window. Figure 4a shows the graph of the varimax as a function of ρ. Figure 4b shows the stacked section for $\rho = 1$. Figure 4c shows the stacked section for $\rho = 1.04$; that is, for the peak of the curve shown in Figure 4a, whereas Figure 4d shows the angle-domain common image gather for the same value of $\rho = 1.04$ and extracted from the prestack cube at the horizontal location of the reflector's truncation.

For both windows, the maximum of the varimax norm corresponds to the value of ρ that best focuses the prestack image and best flattens the angle-domain common image gathers. The semblance peak for the point diffractor is sharper than for the reflector truncation, suggesting that point diffractors provide higher-resolution information on migration velocity than reflectors' truncations.

I also computed the varimax in equation 2 as a function of ρ in two windows of the prestack migrated image corresponding to the sinusoidal reflector (Figure 2.) The first window is centered on the bottom of the syncline and the second centered on the top of the anticline. Figure 5 and Figure 6 show: a) graphs of the varimax as function of ρ, b) the stacked sections corresponding the correct values of ρ ($\rho = 1.06$ for Figure 5b and $\rho = 1.045$ for Figure 6b,) c) the stacked sections corresponding the the varimax peaks ($\rho = .995$ for Figure 5c and $\rho = 1.105$ for Figure 6c,) and d) the angle-domain common image gathers extracted at the very bottom of the syncline in Figure 5d and top of the anticline in Figure 6d.

For the first window, the peak of the varimax corresponds to a value of ρ that is too low, whereas for the second window the peak of the varimax corresponds to a value of ρ that is too high. The cause of these errors is that the image of concave reflectors can be made more spiky (i.e. lower entropy) by undermigration than by migration with the correct velocity. Similarly, the image of a convex reflector can be made more spiky by overmigration than by migration with the correct velocity. If the varimax norm were used to determine the residual-migration parameter ρ it would lead to images with wrong structure and non-

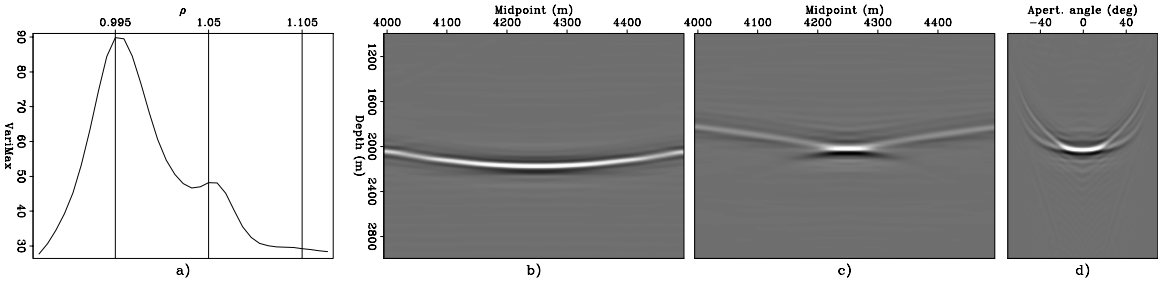


Figure 5: (a) Graph of the varimax norm as a function of ρ, (b) stacked section for $\rho = 1.06$, (c) stacked section for $\rho = .995$, and (d) angle-domain common image gather for $\rho = .995$ at $x = 4,250$ meters. [**CR**] biondo1/. Probl-4250-overn

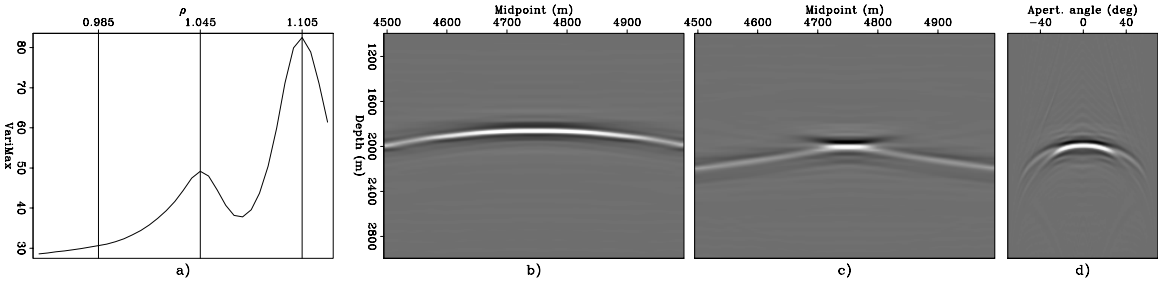


Figure 6: (a) Graph of the varimax norm as a function of ρ, (b) stacked section for $\rho = 1.045$, (c) stacked section for $\rho = 1.105$, and (d) angle-domain common image gather for $\rho = 1.105$ at $x = 4,750$ meters. [**CR**] biondo1/. Probl-4750-overn

flat common-image gathers. However, the secondary peaks of the varimax norm in both Figure 5 and Figure 6 are approximately located at the correct value of ρ. This secondary peaks indicate that there is potentially useful focusing information in the images, but to be practically useful we must devise a method that is not biased by the reflectors' curvature.

Measuring image coherency across structural dips

As an alternative to minimizing entropy, in this paper I propose to measure image focusing by maximizing coherency along both the structural-dip axes and the aperture/azimuth axes. The simultaneous use of dips and aperture angles is discussed in the next section. In this section, I show that measuring coherency along the structural dips does provide information on image focusing and I illustrate the concept by using the same two 2D synthetic data sets shown above. I will also demonstrate that maximizing coherency only along the structural dips may lead us to similar problems as the minimization of entropy.

To measure coherency along the structural dip α, I first create the dip-decomposed prestack image $\mathbf{R}(\mathbf{x}, \gamma, \alpha, \rho)$ by residual prestack migration, and then I compute the follow-

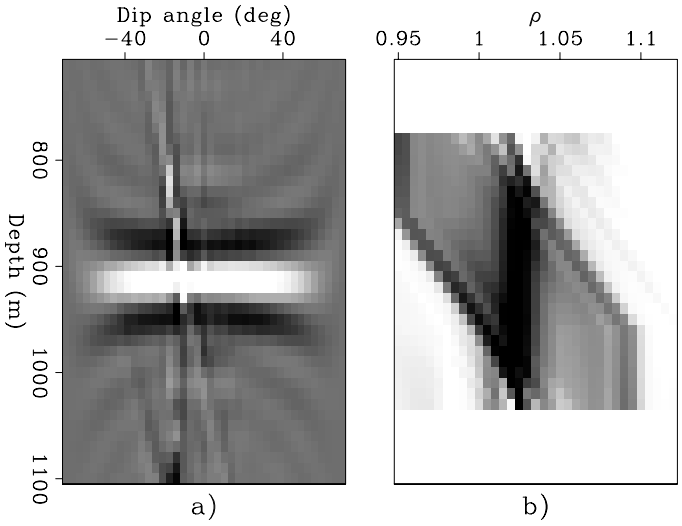


Figure 7: a) Dip-decomposed stack image of the diffractor-point window as a function of the dip angle extracted at $x = 4,700$ meters and $\rho = 1.025$), and (b) semblance ρ-spectrum computed at $x = 4,700$ meters. [**CR**] biondo1/. Dips-4700-diffr-overn

ing semblance functional:

$$S_\alpha(\mathbf{x},\rho) = \frac{\left[\sum_\alpha \sum_\gamma \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}{N_\alpha \sum_\alpha \left[\sum_\gamma \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}, \tag{3}$$

where N_α is the number of dips to be included in the computation. Notice that, as for the varimax in equation 2, semblance along structural dips is computed after stacking over the aperture angle γ.

The determination of the dip summation range at each image location and for each value of the parameter ρ is a practical problem of the proposed method. For the examples shown in this paper I determined the summation ranges for both α and γ by applying an amplitude thresholding criterion based on both local and global amplitude maxima measured from the images. To improve the smoothness of the semblance spectra, I averaged the evaluation of equation 3, and of all the other semblance functionals introduced in this paper, over spatial windows extending along both the z and x directions.

Figure 7a shows the dip-decomposed stack image of the diffractor-point window as a function of the dip angle α extracted from $\mathbf{R}(\mathbf{x},\alpha,\rho)$ at the point-diffractor's horizontal position and for $\rho = 1.025$; that is, the correct value of ρ. The image is consistent as a function of dips, with the exception of an image artifact caused by interference with the image from the planar reflectors below the point diffractor. Figure 7b shows the semblance computed by applying equation 3 at the horizontal position of the point diffractor. It has a sharp peak for $\rho = 1.025$. The dip-coherency analysis has thus the potential to provide accurate velocity information.

Figure 8a shows the dip-decomposed stack image of the reflector-truncation window as a function of the dip angle α at extracted from $\mathbf{R}(\mathbf{x},\alpha,\rho)$ at the horizontal position of the reflector's truncation for $\rho = 1.04$; that is, the correct value of ρ. The dip-decomposed image is strongly peaked at $\alpha = -15^o$; that is the dip of the reflector. The event is weak away from $\alpha = -15^o$; and much weaker than the point-diffractor event shown in Figure 7a. Furthermore, polarity of the event switches at $\alpha = -15^o$. At the transition corresponding

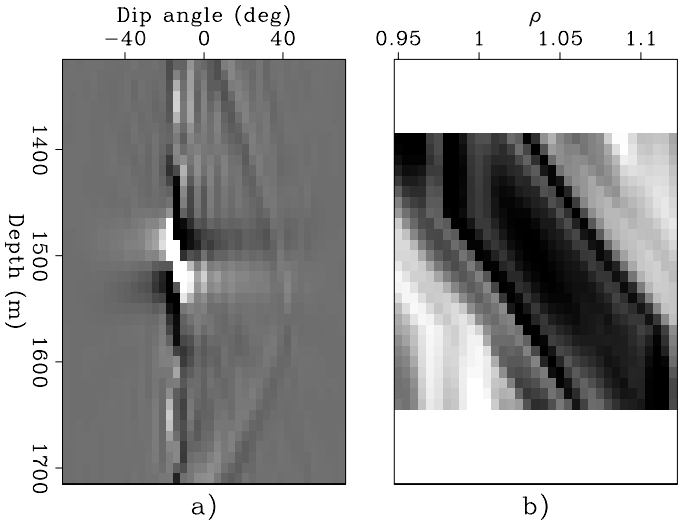


Figure 8: a) Dip-decomposed stack image of the reflector-truncation window as a function of the dip angle extracted at $x = 4,700$ meters and $\rho = 1.04$), and (b) semblance ρ-spectrum computed at $x = 4,700$ meters. [**CR**] biondo1/. Dips-4700-trunc-overn

to the reflector dip, the image is actually rotated by 45 degrees. To compute a higher-quality semblance spectrum, I zeroed the image at $\alpha = -15^o$ and split the computation of the numerator in equation 3 between dips larger than 15 degrees and dips smaller than 15 degrees; that is I computed the following modified semblance functional:

$$S_{\bar{\alpha}}(\mathbf{x},\rho) = \frac{\left[\sum_{\alpha<\bar{\alpha}}\sum_{\gamma}\mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2 + \left[\sum_{\alpha>\bar{\alpha}}\sum_{\gamma}\mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}{(N_\alpha - 1)\sum_{\alpha\neq\bar{\alpha}}\left[\sum_{\gamma}\mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}, \tag{4}$$

where $\bar{\alpha}$ is the structural dip of the truncated reflector. The need to identify a reflector truncation and to estimate the local dip of the reflector is potentially a practical problem with using dip coherency to extract velocity information from reflector's truncations.

The semblance spectrum shown in Figure 8b was computed by applying equation 4 with $\bar{\alpha} = -15^o$. The semblance peak is at the correct value of $\rho = 1.04$ but it is much broader than the peak corresponding to the point diffractor shown in Figure 7b. As noted when comparing Figure 3a with Figure 4a, the velocity information provided by focusing analysis of reflectors' truncations seems to be more difficult to use than the one provided by point diffractors.

The computation of the dip spectra for the data set with sinusoidal reflector illustrates the limitations and potential dangers of relying on dip-only spectra when continuous reflectors have a strong curvature. Figures 9a and 9b show the image decomposed according to structural dips for the bottom of the syncline window for two different values of ρ: $\rho = .995$ for Figure 9a, and $\rho = 1.06$ for Figure 9b (same values of ρ as for Figure 5c and Figure 5b, respectively.) The image is flat as a function of the dip angle for the wrong value of ρ and is frowning for the correct value of ρ. Consequently the dip spectrum shown in Figure 9c peaks at a low value of ρ and would mislead velocity estimation.

The analysis of Figure 10 leads to similar conclusions. In this case the image is flat for a higher value of ρ ($\rho = 1.105$) than the correct one ($\rho = 1.045$), for which the image is actually smiling. The semblance spectrum is also biased toward higher values of ρ.

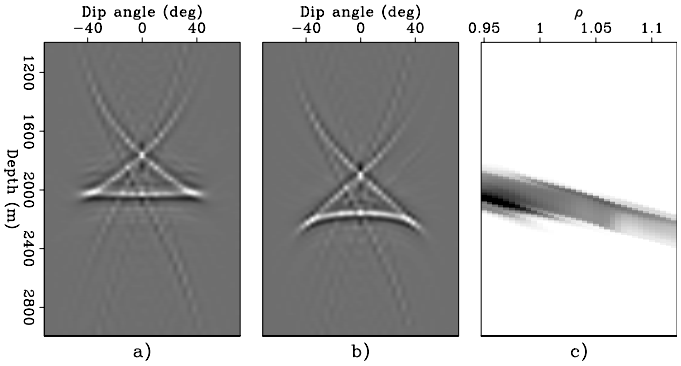


Figure 9: a) Dip-decomposed stack image of the bottom of the syncline window as a function of the dip angle extracted at $x = 4,250$ meters and $\rho = .995$), (b) dip-decomposed stack image for $\rho = 1.06$), and (c) semblance ρ-spectrum computed at $x = 4,250$ meters. [**CR**] biondo1/. Dips-4250-overn

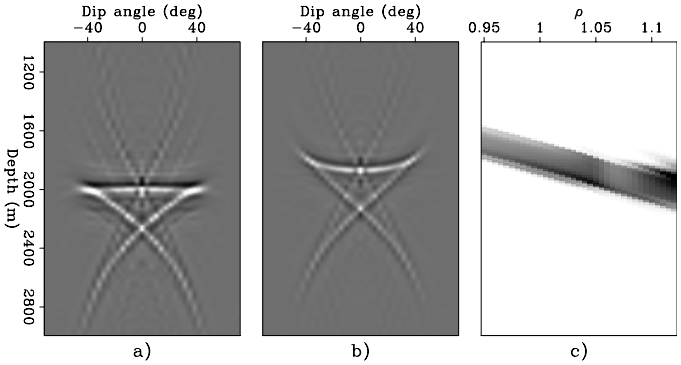


Figure 10: a) Dip-decomposed stack image of the top of the anticline window as a function of the dip angle extracted at $x = 4,750$ meters and $\rho = 1.105$), (b) dip-decomposed stack image for $\rho = 1.045$), and (c) semblance ρ-spectrum computed at $x = 4,750$ meters. [**CR**] biondo1/. Dips-4750-overn

IMAGE-FOCUSING SEMBLANCE

In the previous section, I showed that we can measure image coherency across the structural dip axes to extract focusing information from stacked images. I also showed the shortcomings of this procedure in presence of reflector's curvature. In this section, I introduce a generalization of the semblance functional that measures coherency simultaneously along the dip axes and the aperture/azimuth axes. I name this semblance functional the *Image-focusing semblance*. In 2D it is defined as:

$$S_{(\gamma,\alpha)}(\mathbf{x},\rho) = \frac{\left[\sum_\gamma \sum_\alpha \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}{N_\gamma N_\alpha \sum_\gamma \sum_\alpha \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)^2}, \tag{5}$$

where N_γ is the number of aperture angles to be included in the computation.

As discussed in the previous section, the polarity of reflectors' truncation reverses at the reflectors' dip (Figure 8.) The semblance functional introduced in equation 5 can be modified to better measure image focusing of reflectors' truncation in a way analogous to the way I modified equation 3 to become equation 4. For reflector truncations, the image-focusing semblance can thus be computed as:

$$S_{(\gamma,\bar{\alpha})}(\mathbf{x},\rho) = \frac{\left[\sum_\gamma \sum_{\alpha<\bar{\alpha}} \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2 + \left[\sum_\gamma \sum_{\alpha>\bar{\alpha}} \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)\right]^2}{N_\gamma (N_\alpha - 1) \sum_\gamma \sum_{\alpha\neq\bar{\alpha}} \mathbf{R}(\mathbf{x},\gamma,\alpha,\rho)^2}. \tag{6}$$

To better evaluate the amount of additional information provided by measuring coherency along the structural dips, I also computed a conventional semblance functional

that measured coherency only along the aperture angle from the residual prestack migration results. I computed this conventional semblance function according to the following expression:

$$S_\gamma\left(\mathbf{x},\rho\right)=\frac{\left[\sum_\gamma \mathbf{R}\left(\mathbf{x},\gamma,\rho\right)\right]^2}{N_\gamma \sum_\gamma \mathbf{R}\left(\mathbf{x},\gamma,\rho\right)^2}. \tag{7}$$

The ρ spectrum shown in Figure 11a was computed by applying equation 7. To compute the ρ spectrum shown in Figure 11b I used a combination of the semblance functional expressed in equation 5 for the two shallower events, and of the semblance functional expressed in equation 6 for the deepest event, which corresponds to the reflector's truncation. The semblance peak corresponding to the point diffractor (the top event) is much sharper in Figure 11b than in Figure 11a. This result confirms that the use of image-focusing semblance instead of conventional semblance has the potential of enhancing velocity estimation. In Figure 11b the semblance peaks corresponding to the planar dipping event (second from the top) and the reflector's truncation (first from the bottom) are substantially smaller than the one for the point diffractor, but are still located at the correct value of ρ. The relative scaling between the semblance peaks could be improved.

Figure 12 compares conventional aperture-angle ρ spectrum with the proposed image-focusing spectrum evaluated at the horizontal location of the bottom of the syncline in the model shown in Figure 2a. Both spectra peak for the correct value of ρ; that is $\rho = 1.06$. The spectrum computed using the proposed method has a small secondary peak for low ρs, but not as strong as the one for only-dip spectrum (Figure 9c) or the varimax norm (Figure 5a.) Similarly, the spectra computed at the horizontal location of the top of the anticline in the same model peak for the correct value of ρ, as shown in Figure 13.

CONCLUSIONS

Image-focusing analysis can provide useful velocity information, in particular in areas where conventional velocity analysis lacks resolution. Measuring image coherency across the structural-dip axes provides quantitative information on image focusing. However, in the presence of curved reflectors coherency across dips suffers similar shortcomings as conventional minimum entropy functional. To overcome these difficulties, I propose to perform image-focusing analysis by measuring image coherency simultaneously across structural dips and reflection aperture/azimuth angles. The application of the proposed image-focusing semblance to two synthetic data sets demonstrates its potential as a tool to extract quantitative velocity information from image-focusing analysis.

REFERENCES

Biondi, B., 2008, Automatic wave-equation migration velocity analysis: 2008, **134**, 65–77.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

De Vries, D. and A. J. Berkhout, 1984, Velocity analysis based on minimum entropy: Geophysics, **49**, 2132–2142.

Fomel, S., E. Landa, and M. T. Taner, 2007, Poststack velocity analysis by separation and imaging of seismic diffractions: Geophysics, **72**, U89–U94.

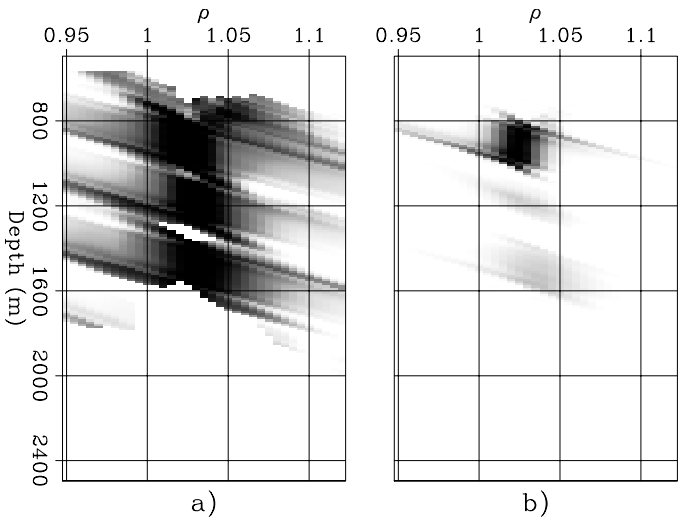


Figure 11: Semblance ρ spectra computed from the first data set (point diffractor and reflector truncations) at $x = 4,700$ meters with: (a) conventional aperture-angle semblance (equation 7,) and (b) aperture-angle and dip-angle semblance (equations 5 and 6.) [**CR**] biondo1/. Sembl-4700-both-overn

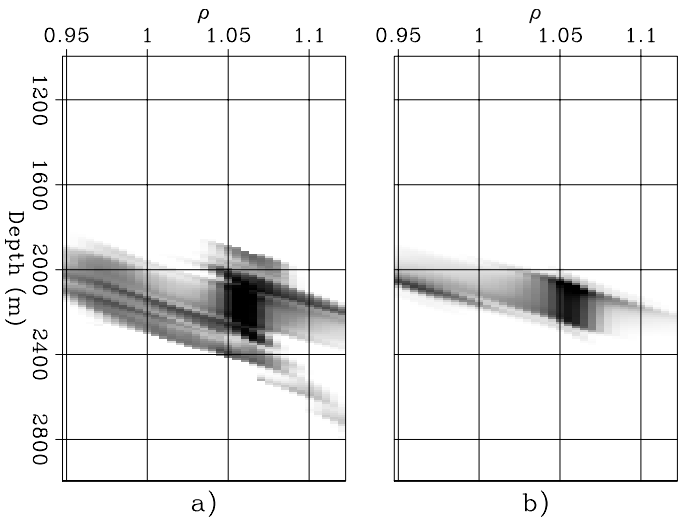


Figure 12: Semblance ρ spectra computed from the second data set (sinusoidal reflector) at $x = 4,250$ meters with: (a) conventional aperture-angle semblance (equation 7,) and (b) aperture-angle and dip-angle semblance (equation 5.) [**CR**] biondo1/. Sembl-4250-both-overn

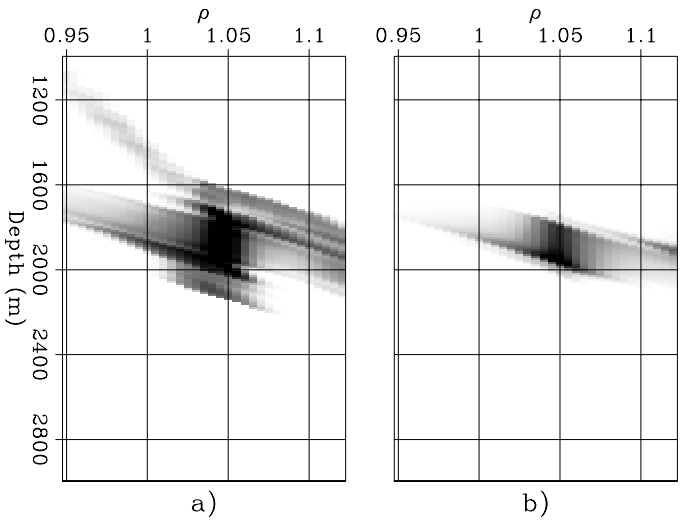


Figure 13: Semblance ρ spectra computed from the second data set (sinusoidal reflector) at $x = 4,750$ meters with: (a) conventional aperture-angle semblance (equation 7,) and (b) aperture-angle and dip-angle semblance (equation 5.) [**CR**] biondo1/. Sembl-4750-both-overn

Harlan, W. S., J. F. Claerbout, and F. Rocca, 1984, Signal/noise separation and velocity estimation: Geophysics, **49**, 1869–1880.

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. 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.

Stinson, K., E. Crase, W.-K. Chan, and S. Levy, 2005, Optimized determination of migration velocities: Recorder, **30**, 5–6.

Wang, B., V. Dirks, P. Guillaume, F. Audebert, and D. Epili, 2006, A 3d subsalt tomography based on wave-equation migration-perturbation scans: Geophysics, **71**, E1–E6.

Wiggins, R., 1985, Entropy guided deconvolution: Geophysics, **50**, 2720–2726.

Migration velocity analysis with cross-gradient constraint

Mohammad Maysami

ABSTRACT

Velocity analysis plays a fundamental to seismic imaging. A variety of techniques using pre-stack seismic data exist for migration-velocity analysis, including reflection tomographic inversion methods. However, when the wavefield propagation is complex, reflection tomography may fail to converge to a geologically reasonable velocity estimation. Non-seismic geological properties can be integrated in the reflection-seismic tomography problem to achieve better velocity estimation. Here, I propose to use cross-gradients function as a similarity measure to constrain the tomography problem and enforce a general geometrical structure for the seismic velocity estimates.

INTRODUCTION

Precise estimation of subsurface velocities is a requirement for high quality seismic imaging. Without an accurate velocity, seismic reflectors are misplaced, the image is unfocused, and seismic images can easily mislead earth scientists (Claerbout, 1999; Clapp, 2001). Defining a reliable velocity model for seismic imaging is a difficult task, especially when sharp lateral and vertical velocity variations are present. Velocity estimation becomes even more challenging when seismic data are noisy. Therefore it is harder to extract velocity information (Clapp, 2001).

In areas with complex structures, and significant lateral velocity variations, velocity analysis is a challenging task. In these areas, reflection-tomography methods are often more effective than conventional velocity-estimation methods based on measurements of stacking velocities (Biondi, 1990; Clapp, 2001). Unfortunately, the reflection-tomography problem is ill-posed and under-determined. Furthermore, it may not converge to a realistic velocity model without *a priori* information, e.g., regularization constraints and other types of geophysical properties in addition to seismic data (Clapp, 2001).

The main challenge in integrating different geophysical data sets is the absence of an analytical relationship between properties exploited by different geophysical surveys. Most often, probabilistic relations among these geophysical properties are used to address this shortcoming. A different approach would be to use gradients field as an objective measure of geometrical similarity. This is true since the variations of geophysical properties can be described by a magnitude and a direction (Gallardo and Meju, 2004, 2007).

Here, I use the cross-gradients function introduced by Gallardo and Meju (2004, 2007) to integrate the resistivity field measured by electromagnetic surveys into the reflection-seismic tomography problem. The integration of this additional piece of information may lead to velocity estimates that are geologically reasonable.

My paper is organized as follows. First, I present a short overview of reflection tomography, followed by introducing the cross-gradients function as a structural similarity mea-

sure. Next, I show how we can use this measure as an extra constraint for the reflection-tomography problem, with the goal of obtaining at a more accurate velocity estimation. Last, I discuss the future work based on these ideas.

REFLECTION SEISMIC-TOMOGRAPHY

By definition, tomography is an inverse problem, where a field is reconstructed from its known linear path integrals, i.e., projections (Clayton, 1984; Iyer and Hirahara, 1993). We can think of tomography as a matrix operator $\mathbf{T}$, which integrates slowness along the raypath. The tomography problem can then be stated as

$$\mathbf{t} = \mathbf{T}\,\mathbf{s}, \tag{1}$$

where $\mathbf{t}$ and $\mathbf{s}$ are travel time and slowness vector, respectively (Clapp, 2001).

The raypaths are dependent on the velocity field. Consequently, the tomography operator is a function of the model parameters. This dependency causes the tomography problem to be nonlinear, which makes it difficult to solve. A common technique to overcome this non-linearity is to iteratively linearize the operator around *a prior* estimation of the slowness field $\mathbf{s}_0$ (Biondi, 1990; Etgen, 1990; Clapp, 2001). The linearization of the tomography problem by using a Taylor expansion is then given by

$$\mathbf{t} \approx \mathbf{T}\mathbf{s}_0 + \frac{\partial \mathbf{T}}{\partial \mathbf{s}}\Big|_{\mathbf{s}=\mathbf{s}_0}\Delta\mathbf{s}. \tag{2}$$

Here, $\Delta\mathbf{s} = \mathbf{s} - \mathbf{s}_0$ represents the update in the slowness field with respect to the *a priori* slowness estimation, $\mathbf{s}_0$. Equation 2 can be simplified as

$$\Delta\mathbf{t} = \mathbf{t} - \mathbf{T}\mathbf{s}_0 \approx \mathbf{T}_L\Delta\mathbf{s}, \tag{3}$$

where $\mathbf{T}_L = \frac{\partial \mathbf{T}}{\partial \mathbf{s}}\big|_{\mathbf{s}=\mathbf{s}_0}$ is a linear approximation of $\mathbf{T}$. A second, but not least, difficulty arises because the location of reflection points are unknown and a function of the velocity field (van Trier, 1990; Stork, 1992).

Clapp (2001) attempts to resolve some of the non-linearity issues with the introduction of a new tomography operator in the tau domain and use of steering filters. In addition to geologic models other types of geophysical data can also be extremely important. In the following section, I show how the cross-gradients function can be used to add constraints to the seismic tomography problem.

THE CROSS-GRADIENTS FUNCTION AS A CONSTRAINT FOR THE TOMOGRAPHY PROBLEM

As mentioned in the previous section, integrating different types of geophysical data can lead to improvements in reflection-seismic tomography results due to reduction of model uncertainty. For this purpose, I propose the cross-gradients function, which can also be considered as a metric to measure the structural similarity between two fields. Following Gallardo and Meju (2004), we can define the cross-gradients function for the tomography problem as

$$\mathbf{g} = \nabla\mathbf{r} \times \nabla\mathbf{s}, \tag{4}$$

where $\mathbf{r}$ and $\mathbf{s}$ can represent any two model parameters, e.g., resistivity and slowness in our case, respectively. Zero values of the cross-gradients function correspond to points where spatial changes in both geophysical properties, i.e., $\nabla\mathbf{r}$ and $\nabla\mathbf{s}$, align. However, the function is also zero where the magnitude of spatial variations of either field is negligible, e.g., where either property is smooth. Note that the cross-gradients function is a non-linear function of $\mathbf{s}$ and $\mathbf{r}$ if both are unknowns.

Figure 1 shows a synthetic 2-D resistivity profile with two anomalies in a constant resistivity background. Ideally, we expect different types of geophysical measurements to produce a geometrically similar image of the subsurface. The cross-property relations between pairs of geophysical properties, e.g., seismic velocity and electrical resistivity of rocks (for more details refer to Hacikoylu et al., 2006; Carcione et al., 2007) also support this similarity. Figure 2 shows the corresponding 2-D velocity profile of the modeled subsurface region in Figure 1, which includes both fast and slow anomalies in comparison to the background velocity. The velocity profile is computed using the *Archie/time-average* cross-property relation (Carcione et al., 2007) with arbitrary parameter values. Note that, the structural similarity of Figures 1 and 2 suggest that the cross-gradients function should vanish almost everywhere.

In a 2-D problem, $\mathbf{g}$ simplifies to a scalar function at each point, given by

$$\mathbf{g} = \frac{\partial \mathbf{s}}{\partial x}\,\frac{\partial \mathbf{r}}{\partial z} - \frac{\partial \mathbf{s}}{\partial z}\,\frac{\partial \mathbf{r}}{\partial x}\ , \tag{5}$$

where the model parameters are given in $x-z$ plane. In order to compute the cross-gradients function, we can further simplify it by using first-order forward differences approximation of the first derivative operators. Figure 3 shows the estimated cross-gradients function. Note that it is approximately zero everywhere as expected. Negligible non-zero values are caused by errors in forward-difference estimation. This implies that geometrical changes, e.g., layer boundaries and other subsurface structures, should be sensed by measurement of both geophysical properties, i.e., seismic slowness and electrical resistivity. Therefore, the cross-gradients function can be used as a constraint for joint data inversion problems or to integrate *a priori* information from other fields into the seismic tomography problem.

If an accurate estimate of the electrical resistivity profile is provided, we can use the cross-gradients function as a constraint for the reflection-seismic tomography problem to improve the accuracy of the velocity estimations. In this case, we can write the cross-gradients function given in equation 5 as a linear operator $\mathbf{G}$ on the slowness field, $\mathbf{s}_0 + \Delta\mathbf{s}$. We can then extend the linearized tomography problem by employing $\mathbf{G}$ as an additional constraint. The objective function, $\mathcal{P}(\Delta\mathbf{s})$, of this extended problem becomes

$$\mathcal{P}(\Delta\mathbf{s}) = ||\Delta\mathbf{t} - \mathbf{T}_L\Delta\mathbf{s}||^2 + \epsilon_1^2\ ||\mathbf{A}\Delta\mathbf{s}||^2 + \epsilon_2^2\ ||\mathbf{G}(\mathbf{s}_0 + \Delta\mathbf{s})||^2, \tag{6}$$

where ϵ_1 and ϵ_2 are problem-specific weights, and $\mathbf{A}$ represents any regularization operator other than cross-gradients function such as smoothing operator.

The important advantage of using the cross-gradients function over using steering filters may not be very clear in this synthetic example. Steering filters are most effective for continuous anomalies with smooth boundaries. However, in the case of sharp boundaries, e.g., Gaussian anomalies or salt boundaries, the cross-gradients function is better able to handle the seismic tomography problem. As mentioned in previous section, we can also

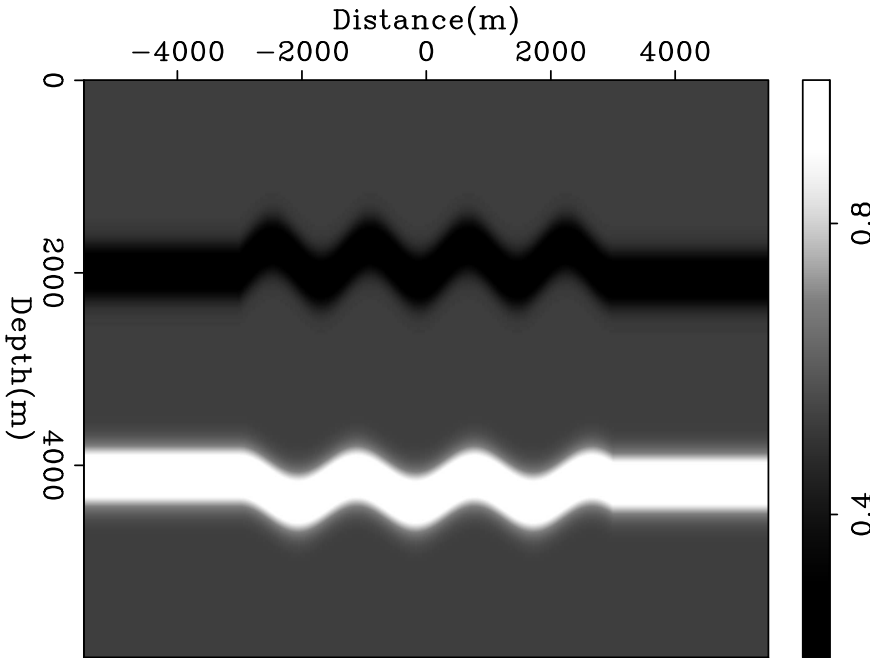


Figure 1: Synthetic resistivity ($Ohm.m$). **[ER]** mohammad1/. fltp-res

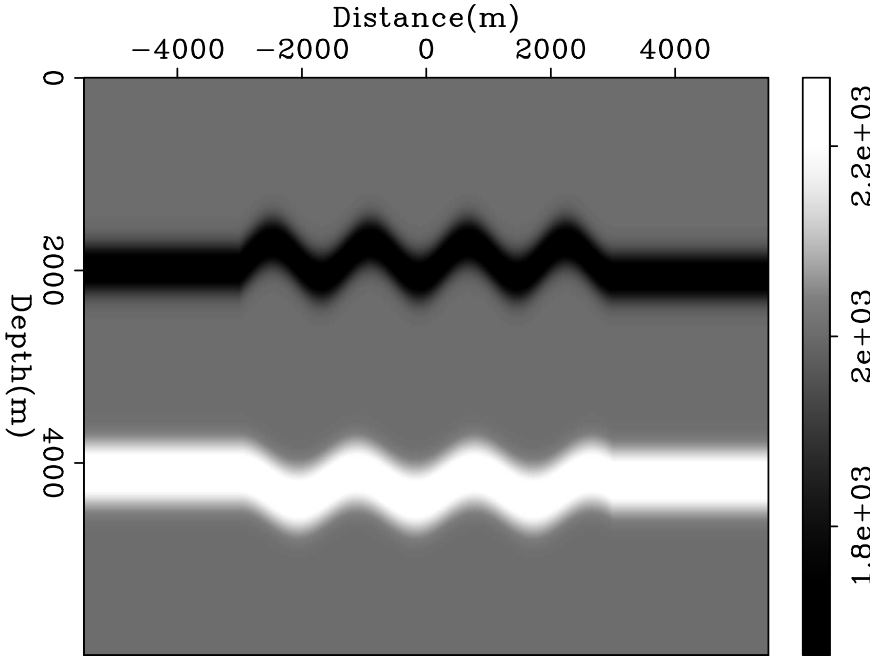


Figure 2: Synthetic velocity (m/s) associated with Figure 1. **[ER]** mohammad1/. fltp-vel

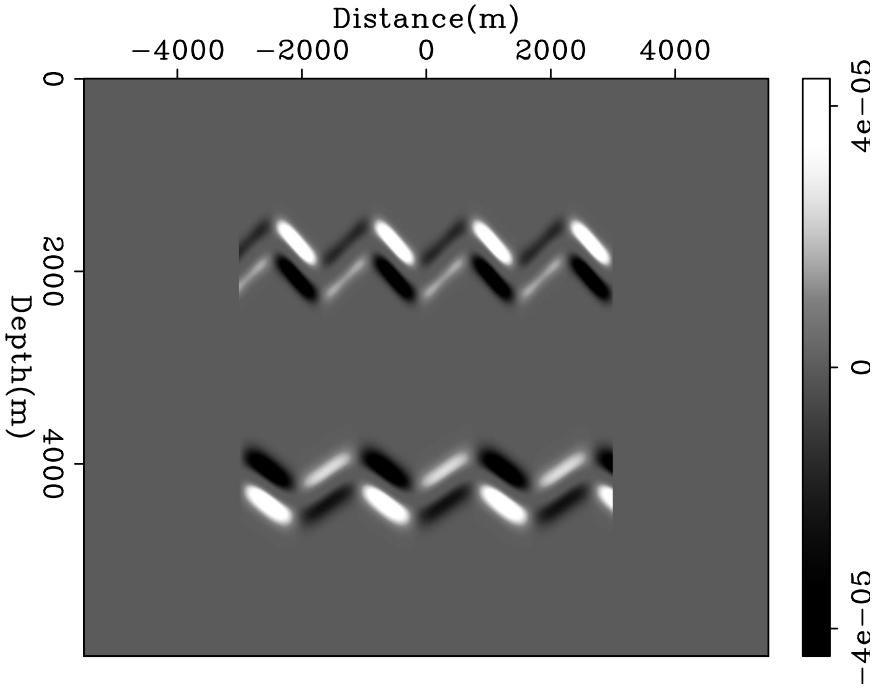


Figure 3: Cross-gradients function of the velocity and resistivity shown in Figures 1 and 2. [**ER**] mohammad1/. fltp-xgrad

use the cross-gradients function as a constraint for joint inversion, where steering filters are not effective. This is true because steering filters assume *a priori* knowledge of the model parameters while the cross-gradients function use the collocated data field to build this information.

FUTURE WORK

The tomography problem stated in Equation 6 is based on the assumption that we have a reasonably accurate estimate of the collocated resistivity field. Given this assumption, I expect the similarity constraint to improve the estimation of slowness profile. I will first incorporate the similarity constraint into the reflection-seismic tomography problem for the synthetic model shown above. Then, I will extend the application of the idea to the tomography problem for 2-D sections of a field data set.

REFERENCES

Biondi, B., 1990, Velocity analysis using beam stacks: PhD thesis, Stanford University.

Carcione, J. M., B. Ursin, and J. I. Nordskag, 2007, Cross-property relations between electrical conductivity and the seismic velocity of rocks: Geophysics, **72**, E193–E204.

Claerbout, J., 1999, Everything depends on v(x,y,z): SEP-Report, **100**, 1–10.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.

Clayton, R. W., 1984, Seismic tomography: Eos, Trans., Am. Geophys Union., **65**.

Etgen, J., 1990, Residual prestack migration and interval velocity estimation: PhD thesis, Stanford University.

Gallardo, L. A. and M. A. Meju, 2004, Joint two-dimensional dc resistivity and seismic travel time inversion with cross-gradients constraints: J. Geophys. Res., **109**.

———, 2007, Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltime data for structural and lithological classication: Geoph. J. Int., **169**, 1261–1272.

Hacikoylu, P., J. Dvorkin, and G. Mavko, 2006, Resistivity-velocity transforms revisited: The Leading Edge, **25**, 1006 – 1009.

Iyer, H. and K. Hirahara, eds., 1993, Seismic tomography: Theory and practice, 1st. ed.: Chapman & Hall.

Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, 680–692.

van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: PhD thesis, Stanford University.

Transmission effects of localized variations of Earth's visco-acoustic parameters

Abdullah Al Theyab and Biondo Biondi

ABSTRACT

In an effort to understand the transmission effects of localized heterogenities in the subsurface, we present the travel-time and amplitude distortions caused by localized variations in velocity and absorption. To examine the relative impact of velocity and absorption heterogeneities on seismic events, we conducted numerical experiments using visco-acoustic finite-difference modeling of the linearized wave-equation for Newtonian fluids. We analyzed the distortions in the midpoint-offset domain. We find that the distortion caused by an anomaly that is both slow and absorptive is different from that an anomaly that is either slow or absorptive, but not both. Our results also indicate that amplitude distortion of highly absorptive anomalies ($Q < 50$) can be comparable to that of small velocity variation (less than 4%), and therefore absorption must be considered in seismic amplitude inversion and AVO analysis.

INTRODUCTION

Localized heterogeneities in the subsurface cause amplitude and travel-time distortions of seismic reflections from underlying reflectors. These distortions are problematic to imaging and AVO analysis. The distortions come in almost regular patterns and usually are recognizable by V-shaped trajectories in the midpoint-offset domain (X-shapes for split-spread acquisition geometry). These distortions can be used to find the locations of the anomalies, which can reveal valuable information for interpreters such as fault locations (Hatchell, 2000). Moreover, they can be used to invert for velocity and absorption anomalies. The analyses of several authors Vlad (2005), Hatchell (2000) and Harlan (1994) have considered mostly velocity anomalies, which cause focusing and defocusing effects.

In this report, we stress that absorption must be considered in the analysis of these distortions. A seismic amplitude inversion that disregards absorption is likly to be biased, especially if velocity perturbations of interest are less than 4%. We examine the relative impact of localized velocity anomalies versus absorption anomalies on seismic amplitude.

BACKGROUND

For a constant background velocity with non-dipping reflector, the distortion trajectory (i.e. the location in prestack data space) caused by a single anomaly can be described by

$$h = \frac{t}{t - t_a} \mid m - m_a \mid \quad , \tag{1}$$

where h is the half offset, t is travel-time, m is the midpoint, and m_a and t_a are the midpoint and traveltime location of the anomaly (Vlad, 2005). The trajectory of the distortion is controlled by the background velocity and the geometry of the reflectors (Vlad, 2005). Besides the trajectory, the distortion has a time-signature (i.e. changes in travel time caused by the presence of the anomaly) and an amplitude-signature (i.e. changes in amplitude). Hatchell (2000) showed real data examples of different amplitude signatures caused by faulting. He also showed that the asymmetry of a velocity anomaly causes different focusing effects depending on whether it is encountered in the receiver leg or source leg, which means the signature can be asymmetric. The signature of an anomaly depends on it's size, shape, type (i.e. velocity, absorption, or both), and the strength (departure from the background velocity and/or absorption).

Considering only type of the anomaly for this study, we use a constant-background velocity and constant-background Q-factor with non dipping-reflectors. Although simple, this geologic model of non-dipping reflectors exists in many geologic provinces, which justifies using it here.

VISCO-ACOUSTIC MODELING

For this study, several seismic experiments were simulated using many different geologic visco-acoustic models. The models are parametrized by three fields for velocity, density, and Q-factor (i.e. the Q-factor for the peak frequency in the source wavelet used). Direct wave-equation modeling was conducted using time-domain finite differencing of the linearized wave equation for Newtonian fluids (Mavko et al., 2003). Finite differencing was explicit in time, and the spatial derivatives were computed in the Fourier domain to attain better accuracy.

A problem with using the linearized wave equation for Newtonian fluids is the acausality, which was observed as negative time shifts in some of our test cases (see Figure 1). Those time shifts, however, are extremely small and can be observed only after dense resampling, which is unreasonable considering the accuracy of our numerical modeling. For this study, only the time shifts of the maximum absolute amplitude are considered.

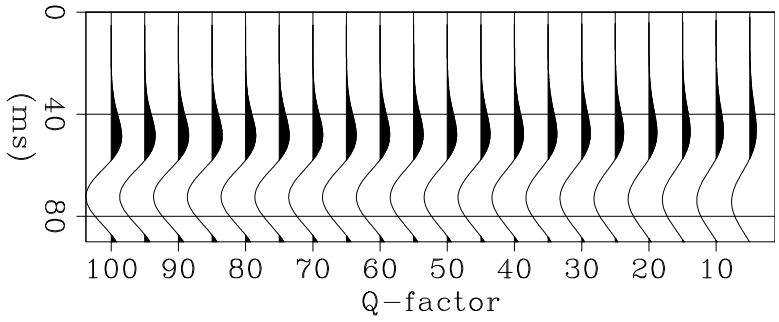


Figure 1: Distortions of a wavelet after passing an absorption anomaly. The amplitude drop as a function of Q-factor is accompanied by a wavelet stretch and a shift of the zero crossing. **[CR]** altheyab1/. avpwiggle

EFFECT ON REFLECTED WAVES

The models used to analyze transmission effects on reflection data have four reflections caused by density contrasts in a constant-background velocity and constant background absorption medium. Four different possible models are considered, two of which are shown Figure 2. In the first model (left), there are three slowness anomalies with a velocity that is 2% less than the background velocity (3000 m/s). In the second model (right), the three anomalies are replaced by absorption anomalies with same shape and size. These absorption anomalies have a Q-factor 50% less than the background Q-factor ($Q = 100$). The third model has the both the velocity and absorption anomalies. In the fourth model, a similar model, but without the anomalies, was used as a reference model for the subsequent analysis.

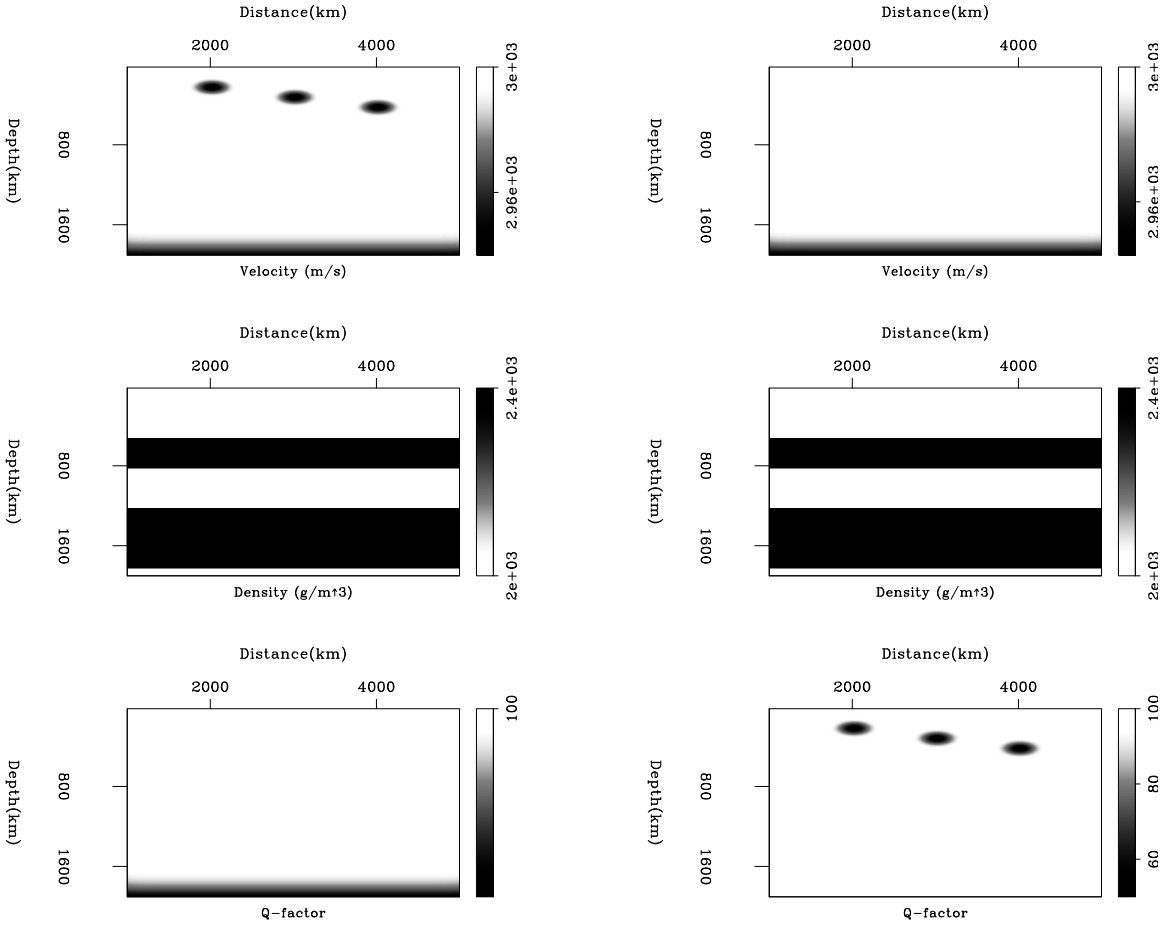


Figure 2: Visco-acoustic models with a constant background velocity (top) and Q-factor (bottom), with four density (middle) reflectors, and three anomalies at depths of 100 m, 200 m and 400 m. Left: the anomalies are velocity variations of 2% from the background velocity. Right: the anomalies are absorption variations with Q-factor 50% less than the background.[**ER**] altheyab1/. rflmodels

To measure the effect of the anomalies, we applied NMO correction to the seismic events coming from the reference model. Then, the maximum amplitude A_{ref} of each reflection event and its arrival time t_{ref} were picked (trace by trace). The same procedure was applied

to the resulting data from the models with anomalies to obtain A_{max} and t_{max}. Time delays are then computed by taking the difference of the arrival times,

$$\delta t = t_{max} - t_{ref}. \tag{2}$$

Figure 3 shows the travel-time delays δt of the maximum amplitude (caused by the presence of the anomalies) sorted into the midpoint-offset domain. The maximum amplitude differences normalized with the amplitude of reference reflections (i.e. reflections if the model had no anomalies) were computed using

$$\delta A = \frac{A_{max} - A_{ref}}{A_{ref}}, \tag{3}$$

and are shown in Figure 4.

Each row of the Figures 3 and 4 corresponds to one of the four reflectors; the top rows are for the shallowest reflector and the bottom ones are for the deepest. The left columns of the two figures are for the data resulting from the left model in Figure 2, and the right columns are for the right model in Figure 2.

Each anomaly in the model causes an X-shaped signature centered at the midpoint location of the anomaly. The arms of the shape generally spread further apart and each becomes broader with increasing offsets. The trajectory angle (i.e. the angle between the arms of the X-shape) is generally narrower for the deepest anomaly, especially for the first reflector, with which the anomaly coincides. This distinction, however, is gradually lost with increasing depth of reflectors, as shown in the fourth reflector, where the trajectory angles are almost the same for all three anomalies.

As expected, the slow anomalies cause time delays (positive shifts) as shown in the left side of Figure 3. The magnitudes of the time shifts are smaller for deeper reflectors, and span a larger range of offsets, which results in fatter patterns. Absorption anomalies cause almost no time shifts. The width of the signature is less dependent on increasing offset. Instead, it depends on the depth of the anomaly.

Amplitude distortions in Figure 4 show trajectories similar to time shifts in the midpoint-offset domain. The magnitude of the distortion generally decreases with depth, and becomes less focused with increasing offsets. The arms of signatures narrow with depth. Because of tilting of the upcoming waves, the energy is confined closer to the source and stretches with increasing offset. This causes the asymmetry of the signatures about the axes of the arms. Velocity anomalies cause focusing. Therefore, we have higher amplitudes paired with two shadow zones (drops in amplitude), as shown in the left side of Figure 4. The absorption signature, on the other hand, shows only a drop in amplitude. The width of the absorption signature is generally smaller than that of velocity because of the absence of focusing.

From Figures 3 and 4, we can observe that the time delay of the velocity signature is strictly positive, and the amplitude signature has a doublet of positive and negative amplitude changes. The absorption amplitude signature is strictly negative, with no time shifts. It should be noticed, however, that that the magnitude of the absorption amplitude distortions matches those of the velocity distortions. This is shown in the two cases presented. Figure 6 shows amplitude changes to the zero-offset reflection that passes twice through an anomaly (left). The changes in amplitude are shown as functions of percentile change of velocity (middle) relative to the background velocity ($v = 3000\ m/s$), and Q-factor

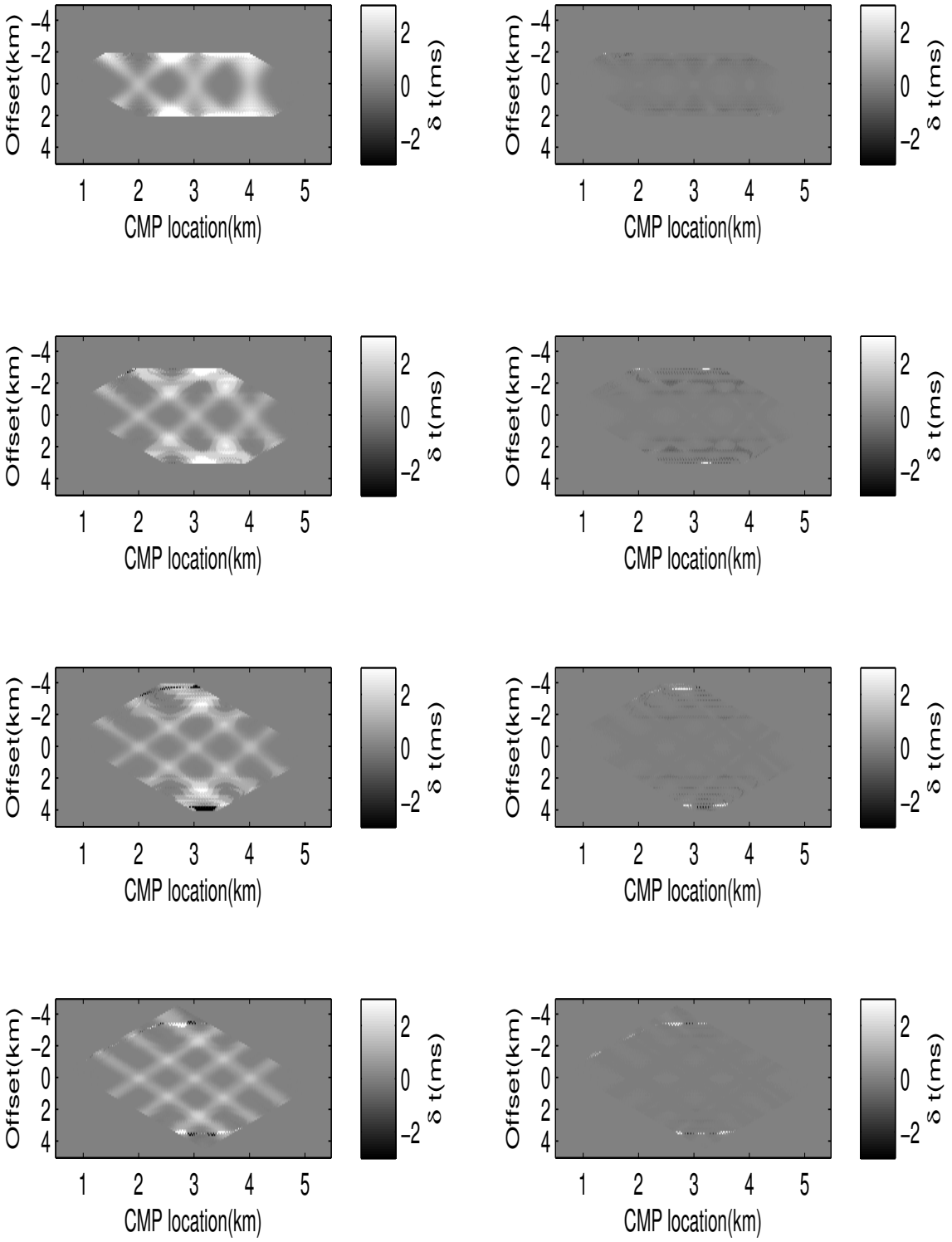


Figure 3: Time delays caused by the velocity anomalies (left) and absorption anomalies (right) on on the four primary reflections, the shallowest (top) to the deepest (bottom) in the midpoint-offset domain. **[CR]** altheyab1/. rfltimes

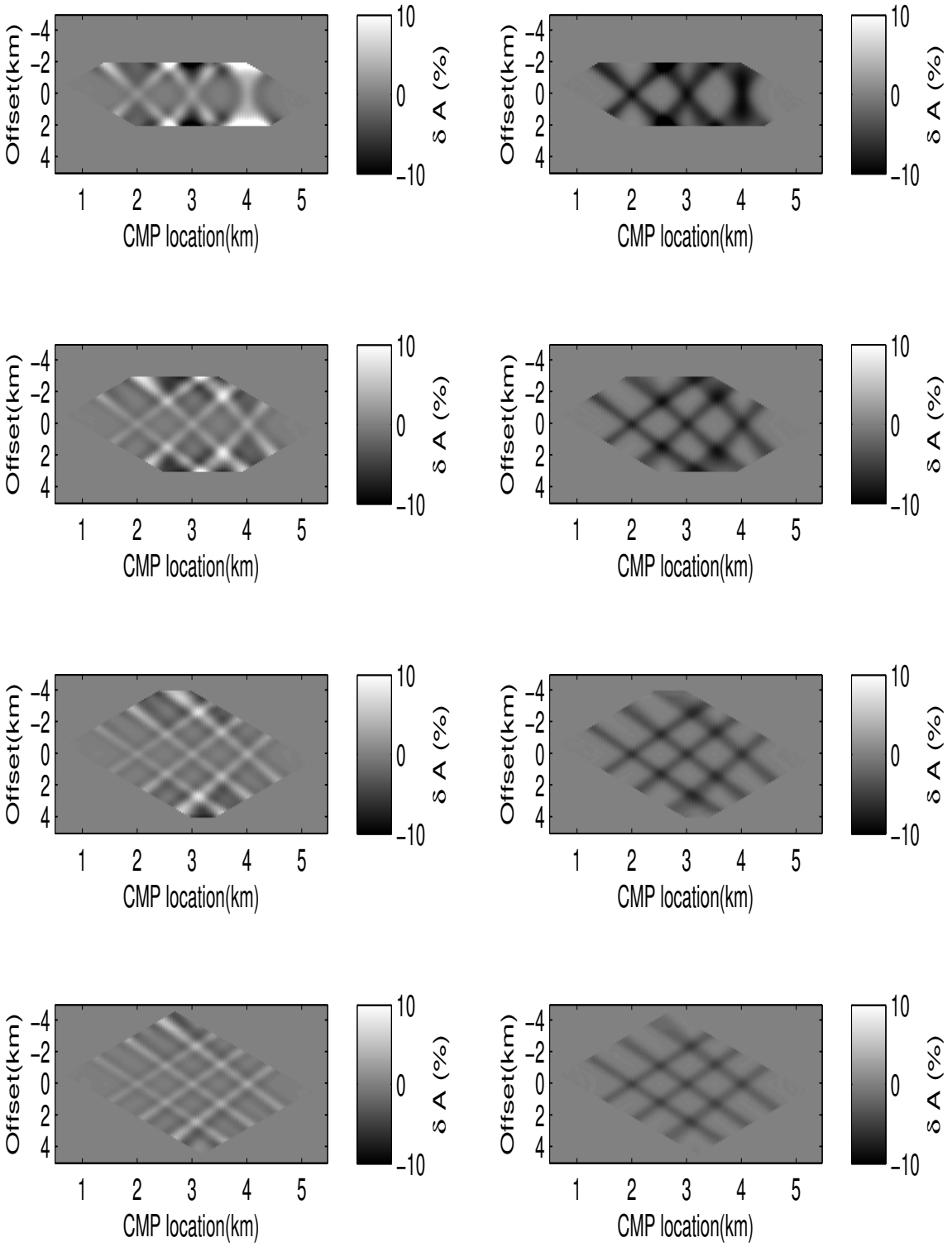


Figure 4: Amplitude changes caused by the velocity anomalies (left) and absorption anomalies (right) on the four primary reflections, the shallowest (top) to the deepest (bottom) in the midpoint-offset domain. **[CR]** altheyab1/. rflamps

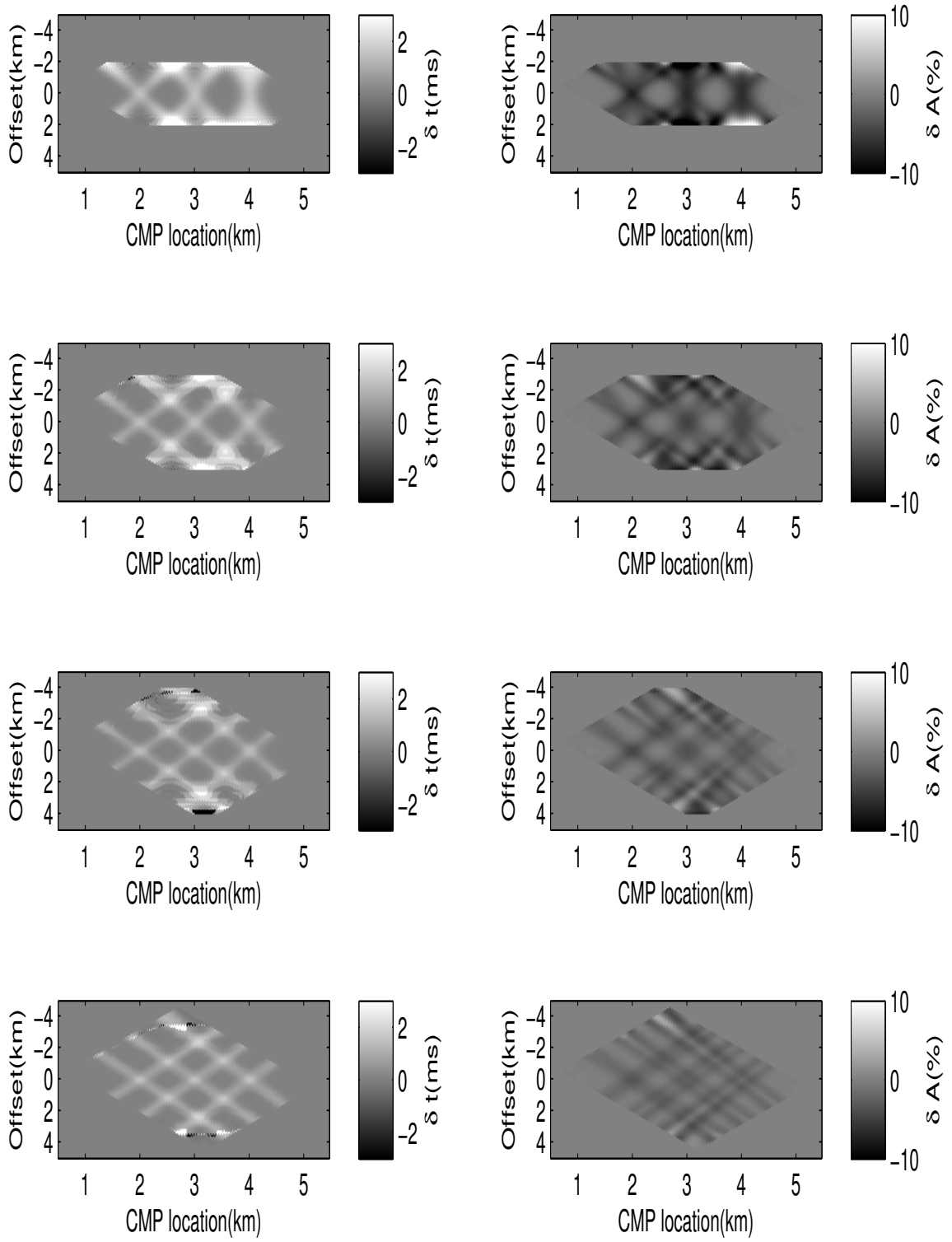


Figure 5: The time (left) and amplitude (right) changes to reflections from the shallowest (top) to deepest (bottom) caused by anomalies that are both slow and absorptive. **[CR]** altheyab1/. rflvq

(right) relative to the background Q-factor ($Q = 100$). The range of amplitude drop due to absorption is generally similar to that caused by the velocity changes of interest ($< 5\%$).

Figure 5 shows the time signature (left) and the amplitude signature (right) for the third model, i.e. the model with both velocity and absorption anomalies coinciding. The time signature looks similar to that of slowness-only anomalies. The amplitude signature is more complex than the two velocity-only and absorption-only cases. For the near offsets, the focusing effect and the absorption effect cancel each other leaving only two parallel shadow zones. The focusing effect dominates the amplitude signature at the far offsets and we see the dim-bright-dim signature again. From this, we can see that the near offsets play a significant role in determining the presence of absoroption. Missing or noisy near offsets can potentially cause the velocity-absorption effect to be mistaken for velocity-only effect.

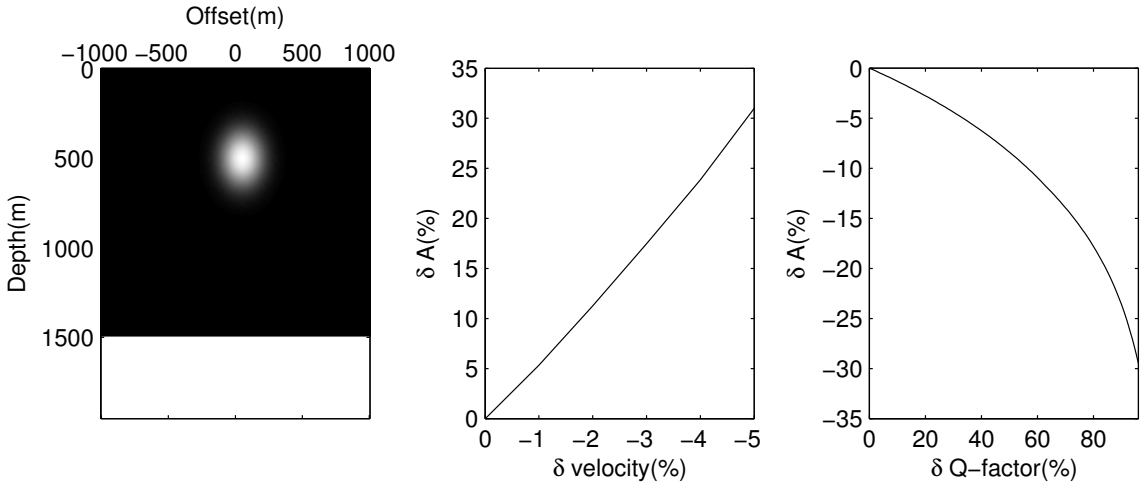


Figure 6: Left: a geologic model of a single reflector and an anomaly that will perturb the primary reflection. The changes of the zero offset reflection are normalized with respect to unperturbed reflection. The amplitude changes as a function of percent change in velocity (middle) or Q-factor (right). **[CR]** altheyab1/. avpcurve

CONCLUSION

In this report, we showed that localized variations in velocities and absorption have smooth effect on seismic amplitude with increasing offset and can hinder AVO analysis. We also showed that the effect of highly absorptive anomalies on seismic amplitude can be of the same magnitude as the effect of small velocity perturbation. Consequently, a possible cancellation of concurrent effects of velocity and absorption can cause different amplitude signatures, which could be confused with the signature of smaller absorption or velocity anomalies.

REFERENCES

Harlan, W. S., 1994, Tomographic correction of transmission distortions in reflected seismic amplitudes: SEG Technical Program Expanded Abstracts, **13**, 968–971.

Hatchell, P. J., 2000, Fault whispers: Transmission distortions on prestack seismic reflection data: Geophysics, **65**, 377–389.

Mavko, G., T. Mukerji, and J. Dvorkin, 2003, The rock physics handbook: Cambridge University Press.

Vlad, I., 2005, Focusing-effect AVO/AVA: overview of results and assessment of problems: Technical report, SEP-120.

Joint wave-equation inversion of time-lapse seismic data

Gboyega Ayeni and Biondo Biondi

ABSTRACT

We discuss two regularized least-squares inversion formulations for time-lapse seismic imaging. Differences in acquisition geometries of baseline and monitor datasets or the presence of a complex overburden can degrade the quality of the time-lapse seismic signature. In such a scenario, the time-lapse amplitude information are poor indicators of the true reservoir property changes. Although the migration operator accurately images the seismic data, it does not remove these amplitude distortions. We pose time-lapse imaging as joint linear inverse problems that utilize concatenations of a target-oriented approximation to the least squares imaging Hessian. In one of the two formulations considered, outputs are inverted time-lapse images, while in the other, outputs are evolving images of the study area. Using a 2D-synthetic sub-salt model, we demonstrate that either joint-inversion formulation can attenuate overburden and geometry artifacts in time-lapse images and that joint wave-equation inversion yields more accurate results than migration or separate inversion.

INTRODUCTION

Hydrocarbon exploration and production has gradually shifted from simple to complex geological environments. Relatively simple imaging and monitoring objectives (e.g. anticlinal-type traps) have been replaced by more complex ones (e.g., sub-salt reservoirs and stratigraphic traps). Since most of the current time-lapse seismic imaging technologies are inadequate in many emerging frontiers, new imaging and monitoring methods are required. In addition, differences almost always exist between acquisition geometries of different seismic datasets. Such geometry differences may be due to new (or more efficient) acquisition systems and design, production facilities (absent at the time of the baseline survey) or nature (e.g., ocean currents).

Our goal is to attenuate artifacts from two major sources:

1. poor and uneven sub-surface illumination in reservoirs under complex overburden, and
2. disparities in acquisition geometries of the baseline and monitor surveys.

We achieve these objectives by simultaneously inverting migrated images from different vintages with a target-oriented approximation (Valenciano, 2008) to the linear least-squares wave-equation Hessian. The Hessian operator in this problem can be regarded as a set of non-stationary deconvolution filters in a single survey, or a concatenation of sub-matrices built from such filters in multiple surveys. We discuss two joint-inversion formulations of the seismic monitoring problem:

1. regularized joint-inversion for image differences (RJID):
    - input: staggered sums of migrated images, and
    - output: inverted baseline image and image differences between successive surveys;
2. regularized joint-inversion of multiple images (RJMI)
    - input: migrated images for all surveys,
    - output: inverted images for all input surveys.

Solving a single joint-inversion problem enables the incorporation of prior knowledge of the reservoir location, extent and geometry, temporal constraints or information from other sources (e.g., production history-matching).

As previously noted, inputs into the RJID formulation are staggered sums of migrated images from multiple surveys and the outputs are inverted baseline and time-lapse images between successive surveys. Since the imaging and monitoring objectives are decoupled, different regularization schemes can be defined for the baseline and time-lapse images. Inputs and outputs to RJMI are migrated images and corresponding inverted images respectively. RJMI differs from separate inversion, because a coupling operator introduces desirable temporal constraints during inversion.

In order to arrive at both formulations, we have assumed that the background baseline velocity model is known and that it changes slowly between surveys. We also assume that such small velocity changes have negligible impact on wave propagation through the earth, at least to the top of the reservoir. Where there are noticeable displacements between images — as a result of significant velocity changes or geomechanical effects around the reservoir — an event alignment step (Hale, 2007) can be applied prior to inversion.

In this paper, we briefly summarize the seismic monitoring problem, and then we discuss the basic theory of linear least-squares inversion and its extension to the RJID and RJMI formulations for an arbitarary number of surveys. Finally, using six datasets from a 2D-synthetic sub-salt model, we show that both joint-inversion formulations yield noticeably improved results over migration or separate inversion.

BACKGROUND

There is a wide range of published work on the most important considerations for time-lapse seismic monitoring. For example, Batzle and Wang (1992) outline important rock and fluid relationships; Lumley (1995), Rickett and Lumley (2001), Calvert (2005), and Johnston (2005) discuss important processing and practical applications; and Lefeuvre et al. (2003), Whitcombe et al. (2004), and Zou et al. (2006) showed successful case studies. Ayeni and Biondi (2008) discuss additional considerations and previous work related to seismic monitoring of hydrocarbon reservoirs.

Nemeth et al. (1999), Kuhl and Sacchi (2001), Clapp (2005), and Valenciano (2008) have shown that linear least-squares wave-equation migration of seismic data improves structural and amplitude information. We demonstrate that an extension of least-squares migration to

the time-lapse imaging can improve time-lapse amplitude information, especially if all available data are jointly inverted. Previous authors have discussed joint-inversion applications, including impedance inversion (Sarkar et al., 2003), ray-tomography (Ajo-Franklin et al., 2005) and wave-equation velocity analysis (Albertin et al., 2006). Lumley et al. (2003) show that improvements can be made to time-lapse processing through *simultaneous processing*. Dynamic imaging strategies that utilize aspects of spatio-temporal regularization have also been discussed in other scientific disciplines (Schmitt and Louis, 2002; Schmitt et al., 2002; Zhang et al., 2005; Kindermann and Leitao, 2007).

A joint wave-equation inversion formulation of the time-lapse imaging problem has the advantage that the attenuation of image differences is based on the physics of wave propagation, making it less susceptible to removal of true time-lapse changes than conventional methods. The method proposed by Ajo-Franklin et al. (2005) for tomographic inversion can be directly extended to wave-equation inversion, and it actually forms a first step in the RJMI formulation. Such direct extension to wave-equation migration is too expensive, requiring at least one set of migration and modeling per survey per iteration. In most practical inversion problems, parameter selection requires that the inversion procedure be carried out more than once. By pre-computing the Hessian operators, we are able to test different regularization schemes and parameters for the inversion at several orders of magnitude cheaper than directly solving the least-squares migration problem. In addition, we avoid the use of matching filters which can have unpredictable effects on time-lapse changes within the reservoir (Lumley et al., 2003).

THEORY

Linear inversion

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

$$\mathbf{L}\mathbf{m} = \mathbf{d}, \tag{1}$$

where $\mathbf{m}$ is the reflectivity model. The modeling operator, $\mathbf{L}$, in this study, represents 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:

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

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 defining the acquisition process for the two surveys (baseline and monitor).

The quadratic cost functions for equation 2 are given by

$$\begin{aligned} S(\mathbf{m_0}) &= \|\mathbf{L}_0\mathbf{m}_0 - \mathbf{d}_0\|^2, \\ S(\mathbf{m_1}) &= \|\mathbf{L}_1\mathbf{m}_1 - \mathbf{d}_1\|^2, \end{aligned} \tag{3}$$

and the least-squares solutions are

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

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, $\mathbf{L'}_0$ and $\mathbf{L'}_1$ are the migration operators (adjoints to the modeling operators $\mathbf{L}_0$ and $\mathbf{L}_1$ respectively), and $\mathbf{H}_0 \equiv \mathbf{L'}_0\mathbf{L}_0$ and $\mathbf{H}_1 \equiv \mathbf{L'}_1\mathbf{L}_1$, are the Hessian matrices. Here, and in other parts of this paper, the symbol $'$ denotes transposed complex conjugate. These formulations are based on (but not limited to) one-way wave-equation extrapolation methods.

The Hessian matrices are the second derivatives of the cost functions (equation 3) with respect to all model points in the image. Because the Hessian matrices are generally not invertible for almost any practical scenario, equation 4 is solved iteratively as follows:

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

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

$$\Delta\hat{\mathbf{m}} = \hat{\mathbf{m}}_1 - \hat{\mathbf{m}}_0. \tag{6}$$

We will refer to the method of computing the time-lapse image using equation 6 as *separate inversion* throughout the rest of this paper.

Joint-inversion

Two joint-inversion formulations are discussed in the following sections.

Joint-inversion for image differences (JID)

First, we re-formulate the data modeling operations for the two surveys in equation 2 as follows:

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

where $\mathbf{m}_0 + \Delta\mathbf{m} = \mathbf{m}_1$. In matrix form, these expressions can be combined to give

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

In principle, using an iterative solver, a least-squares solution to equation 8 can be obtained by minimizing the cost function

$$S(\mathbf{m_0}, \Delta\mathbf{m}) = \left|\left| \begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \Delta\mathbf{m} \end{bmatrix} - \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} \right|\right|^2. \tag{9}$$

The computational cost of this approach is proportional to the number of iterations times at least twice the cost of one set of migrations — since each iteration requires at least 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

will be high. An important advantage of the JID (or JMI) formulation 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

$$\begin{bmatrix} \mathbf{L'}_0\mathbf{L}_0+\mathbf{L'}_1\mathbf{L}_1 & \mathbf{L'}_1\mathbf{L}_1 \\ \mathbf{L'}_1\mathbf{L}_1 & \mathbf{L'}_1\mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \mathbf{L'}_0 & \mathbf{L'}_1 \\ \mathbf{0} & \mathbf{L'}_1 \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0+\tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{10}$$

or

$$\begin{bmatrix} \mathbf{H}_0+\mathbf{H}_1 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0+\tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{11}$$

which can be recast as

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

Thus, the inverted baseline and time-lapse images ($\hat{\mathbf{m}}_0$ and $\Delta\hat{\mathbf{m}}$ respectively) can be obtained from equation 12. However, since the Hessian matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ (and hence the joint Hessian operator) are not invertible, equation 11 is solved iteratively. We have extended equation 11 to multiple surveys (Appendix A). When multiple surveys are available, the outputs of the JID formulation are the inverted baseline image and image differences between successive surveys.

Joint-inversion of multiple images (JMI)

The data modeling operations for two surveys can be written as follows

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}. \tag{13}$$

In principle, it is possible to solve for a least-squares solution to equation 13 by minimizing the cost function

$$S(\mathbf{m_0},\mathbf{m_1}) = \left\| \begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} \right\|^2. \tag{14}$$

As discussed in the JID formulation, this would be too expensive to be practical since the cost of one iteration is at least the cost of four migrations. Ajo-Franklin et al. (2005) have shown a tomographic example of this formulation, but since each migration is orders of magnitudes more expensive than ray-based tomography, this approach would be too expensive for wave-equation inversion. Therefore, we reformulate equation 14 as

$$\begin{bmatrix} \mathbf{L'}_0\mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L'}_1\mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{L'}_0 & \mathbf{0}_1 \\ \mathbf{0} & \mathbf{L'}_1 \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{15}$$

or

$$\begin{bmatrix} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{16}$$

which can be written as

$$\begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \tag{17}$$

Thus, the inverted baseline and monitor images ($\hat{\mathbf{m}}_0$ and $\hat{\mathbf{m}}_1$ respectively) can be obtained from equation 17 and the time-lapse image as a difference between the two images as done in equation 5. Also, note that without coupling, as done in the next section, equation 16 is equivalent to equation 5. Since the Hessian matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ (and hence the joint Hessian operator) are not invertible, equation 16 is solved iteratively. An extension of equation 16 to multiple surveys is given in Appendix A.

Joint-inversion with Regularization

In most seismic monitoring 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. As shown in the Appendix, the regularized joint-inversion for image difference (RJID) for two surveys is given by

$$\left(\begin{bmatrix} \mathbf{H_0 + H_1} & \mathbf{H_1} \\ \mathbf{H_1} & \mathbf{H_1} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{0} \\ \mathbf{-\Lambda_{10}} & \mathbf{\Lambda_{11}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{18}$$

where

$$\begin{aligned} \mathbf{R_{ij}} &= \epsilon_\mathbf{i} \mathbf{R'}_\mathbf{i} \epsilon_\mathbf{j} \mathbf{R_j}, \\ \mathbf{\Lambda_{ij}} &= \zeta_\mathbf{i} \mathbf{\Lambda'}_\mathbf{i} \zeta_\mathbf{j} \mathbf{\Lambda_j}, \end{aligned} \tag{19}$$

while $\mathbf{R_0}$ and $\mathbf{R_1}$ are the spatial/imaging constraints for the baseline and time-lapse images respectively, and $\mathbf{\Lambda_0}$ and $\mathbf{\Lambda_1}$ the temporal regularization (or coupling) between the surveys. In the implementation of equation 19, the regularization terms, $\mathbf{R_{ij}}$ and $\mathbf{\Lambda_{ij}}$ are not explicitly computed, but instead, the appropriate operators $\mathbf{R_i}$ and $\mathbf{\Lambda_i}$ (and their adjoints, $\mathbf{R'_i}$ and $\mathbf{\Lambda'_i}$ respectively) are applied at each step of the inversion. The parameters $\epsilon_\mathbf{0}$ and $\epsilon_\mathbf{1}$ determine strength of the spatial regularization on the baseline and time-lapse images respectively, while $\zeta_\mathbf{0}$ and $\zeta_\mathbf{1}$ determine the coupling between surveys. The regularized joint-inversion of multiple images (RJMI) formulation for two surveys is given as

$$\left(\begin{bmatrix} \mathbf{H_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{-\Lambda_{01}} \\ \mathbf{-\Lambda_{10}} & \mathbf{\Lambda_{11}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \tag{20}$$

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), while the temporal regularization ensures that the reservoir changes evolve according to a reasonable scheme (e.g., smooth variation over time). The temporal regularization operator in the RJMI formulation is similar to that used in spatio-temporal tomographic inversion (Ajo-Franklin et al., 2005).

As shown in Appendix A, the general regularized joint-inversion problem can be written in compact notation as

$$[\mathbf{\Xi} + \Re + \mathbf{\Gamma}] \left[\hat{\mathbf{M}}\right] = \left[\tilde{\mathbf{M}}\right], \tag{21}$$

where $\mathbf{\Xi}$ is the Hessian operator, $\Re$ is the spatial/imaging regularization operator, $\mathbf{\Gamma}$ is the temporal regularization operator, $\hat{\mathbf{M}}$ is the model vector and $\tilde{\mathbf{M}}$ the data vector. Each of the components of the RJID and RJMI formulations are fully described in Appendix A.

Note that in the RJID formulation, the imaging (baseline inversion) and monitoring (time-lapse inversion) goals are decoupled, thus allowing for application of different regularization schemes. Since the baseline and time-lapse images are expected to have different desirable properties, the baseline ($\mathbf{R_0}$ and $\mathbf{\Lambda_0}$) and monitor ($\mathbf{R_1}$ to $\mathbf{R_N}$ and $\mathbf{\Lambda_0}$ to $\mathbf{\Lambda_N}$) regularization operators are different. The RJMI formulation is cheaper to solve, since the joint Hessian operator is less dense and with appropriate regularization, the results from the two formulations should be comparable.

Target-oriented Hessian

The computational cost of the full Hessian matrix for one survey (needless to say for multiple surveys) is prohibitive and not practical for any reasonably sized survey. Several authors have discussed possible approximations to the wave-equation Hessian (Shin et al., 2001; Rickett, 2003; Guitton, 2004; Valenciano, 2008; Symes, 2008; Tang, 2008b,a) . The wave-equation Hessian for synthetic seismic data, $\mathbf{d}\left(\mathbf{s},\mathbf{r};\omega\right)$ at a given frequency, ω, recorded by receiver $\mathbf{r}\left(x_r,y_r,z_r\right)$, from a shot $\mathbf{s}\left(x_s,y_s,z_s\right)$ and scattering point $\mathbf{x}(x,y,z)$, is given by

$$\mathbf{H}\left(\mathbf{x},\mathbf{y}\right)=\sum_w \omega^4 \sum_{\mathbf{s}} \left|f'\left(s\right)\right|^2 \mathbf{G}'\left(\mathbf{x},\mathbf{s};\omega\right)\mathbf{G}\left(\mathbf{y},\mathbf{s};\omega\right)\sum_{\mathbf{r}}\mathbf{G}'\left(\mathbf{x},\mathbf{r};\omega\right)\mathbf{G}\left(\mathbf{y},\mathbf{r};\omega\right), \tag{22}$$

where $\mathbf{y}(x,y,z)$ corresponds to all model points. A detailed derivation of the explicit wave-equation Hessian is given by Mulder and Plessix (2004).

Because reservoirs are typically limited in extent, the region of interest is usually smaller than the full image space. Thus, the required Hessian matrices are constructed for a region around the target zone and not for the full survey area. In this paper, we follow the target-oriented approach of Valenciano (2008) in the Hessian computation. Phase-encoding approximations to the target-oriented Hessian (Tang, 2008a) offer improved efficiency in the Hessian computation and are currently being explored as alternatives to the explicit method used in this paper.

The target oriented Hessian (Valenciano, 2008) is given by:

$$\begin{array}{rcll}\mathbf{H}(\mathbf{x_T},\mathbf{x_T}+\mathbf{a_x}) = & \sum_w \omega^4 & \sum_{\mathbf{s}} \left|f'\left(s\right)\right|^2 \mathbf{G}'(\mathbf{x_T},\mathbf{s};\omega)\mathbf{G}(\mathbf{x_T}+\mathbf{a_x},\mathbf{s};\omega) \\ & & \sum_{\mathbf{r}} \mathbf{G}'(\mathbf{x_T},\mathbf{r};\omega)\mathbf{G}(\mathbf{x_T}+\mathbf{a_x},\mathbf{r};\omega),\end{array} \tag{23}$$

where $\mathbf{a_x}$ is the offset from the target image-point $\mathbf{x_T}$ defining the 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 migration. Examples of the target-oriented Hessian operator for the model in Figures 1 and three surveys are shown in Figures 2 and 3 for both RJID and RJMI.

In a single survey, each row of the Hessian is a point-spread function that describes the effects of the limited-bandwidth seismic waveform, geometry and illumination on a

Figure 1: Full impedance model. The box indicates the target area for which Figures 2 and 3 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 velocities ranging from 2200m/s and 2700m/s. The densities range from 2.5g/cc to 3.0g/cc. [**ER**] gayeni1/. surv5-salt-imp

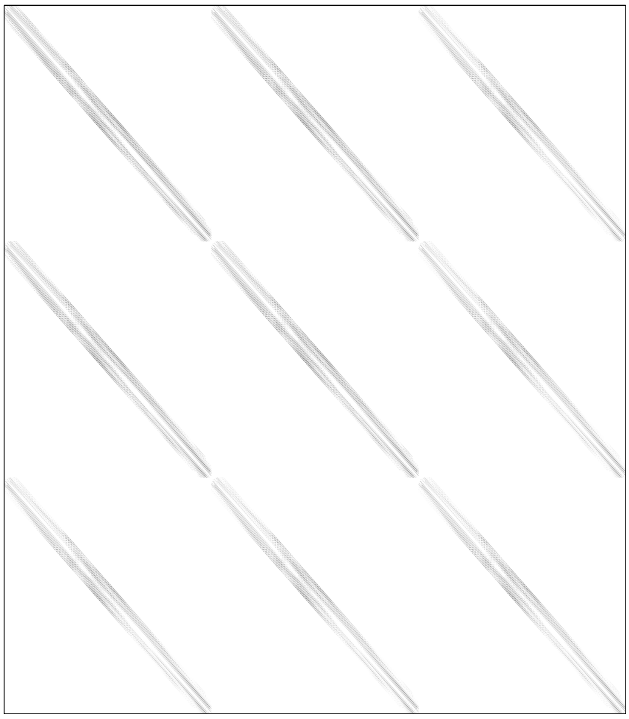


Figure 2: JID: Joint target-oriented Hessian operator for one baseline and two monitor surveys for the reservoir models in Figure 1. The dimension of the square matrix here and in Figure 3 is equal to the number of surveys times the size of the model space. This figure corresponds to the $\boldsymbol{\Xi}$ operator in equation A-24. Note however that the zeros (light regions in the matrices) were neither computed nor stored. [**NR**] gayeni1/. hesssalt3

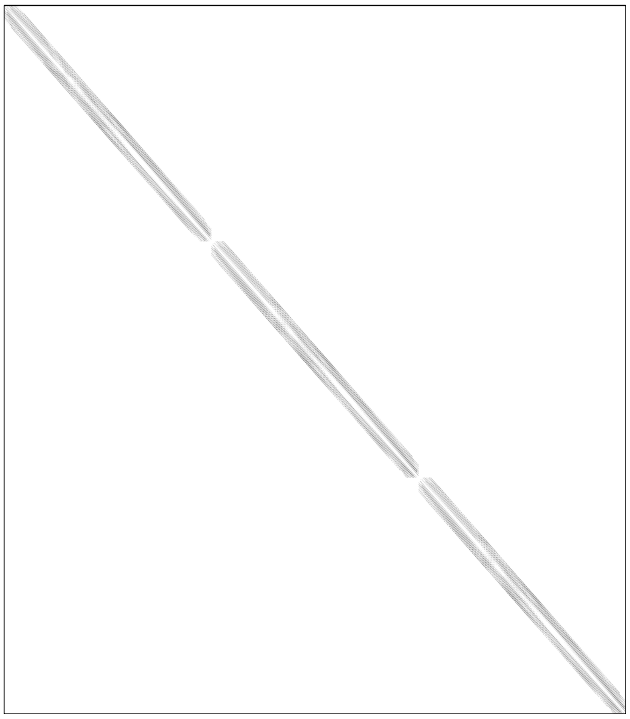


Figure 3: JMI: Joint target-oriented Hessian for one baseline and two monitor surveys for the reservoir models in Figure 1. See caption in Figure 2 for further description. [**NR**] gayeni1/. hesssaltjimi3

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 equations A-25 and A-12. In addition, note that the empty bins in Figures 2 and 3 are neither computed nor stored and that because of the matrix symmetry, only one-half of its elements needs to be computed. The structure of the problem gives a large leeway for parallelization over several domains in both 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), some Green's functions can be reused in the Hessian computation for different surveys.

NUMERICAL EXAMPLE

The inversion formulations were tested on synthetic datasets modeled for the 2D-synthetic sub-salt model in Figure 1. We simulated six datasets (representing different stages of production) using a variable-density acoustic finite-difference algorithm. Reservoir changes were modeled as an expanding Gaussian anomaly centered at x = 0m and z = 3000m. In order to simulate non-repeated acquisition geometries, we modeled all the datasets with spatially different geometries as summarized in Table 1. We modeled 76 shots spaced at 80m and 301 receivers spaced at 20m and for each survey, the receiver spread was kept constant while the shots move along. We 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. The spatial regularization operator is a gradient along reflector dips, while a temporal gradient was used to ensure temporal smoothness.

Table 1: Modeling parameters for synthetic datasets

	Shot/receiver depth	Shot/receiver spread
Geometry 1	0m	−3000 to 3000m
Geometry 2	100m	−3500 to 2500m
Geometry 3	200m	−2600 to 3400m
Geometry 4	40m	−2900 to 3100m
Geometry 5	240m	−3200 to 2800m
Geometry 6	300m	−2500 to 3500m

Figure 4 shows the migrated images for the six surveys, while corresponding illumination maps (diagonal of the Hessian) are shown in Figure 5. The irregular illumination patterns explain the uneven amplitudes of reflectors below the salt in Figure 4. Figure 6 shows the illumination-ratio (normalized rms-difference in illumination) between baseline and the monitor surveys, which measures of the variability of illumination between surveys. We represent the variation in illumination at any image point as the illumination-ratio between the point-spread functions for the different surveys. For example, Figure 7 shows the point-spread functions at image point $[x = -200m, z = 2800m]$, while Figure 8 shows the coresponging illumination-ratio. The time-evolution of the true reflectivity model is shown in Figure 9 and the inversion goal is to reconstruct these. Figure 10 is the reflectivity change obtained from migration, while Figures 11 to 13 were obtained from separate inversion, RJID and RJMI respectively. Both the RJID and RJMI results contain less noise relative to migration (Figures 10) and separate inversion (Figures 11). No pre-processing was done to remove multiples from the data and hence these are expected to adversely affect the inversion.

DISCUSSION

Uneven illumination of the reservoir region as captured by the Hessian diagonal (Figure 5) and ratio (Figure 6) explain the high-amplitude artifacts observed in the migrated time-lapse images in Figure 10. The noticeable shadow zones in parts of the reservoir below the salt (Figure 4) result from the high impedance contrast at the salt-sediment boundary and the complex wave propagation. As shown by Ayeni and Biondi (2008), even where the survey geometries are perfectly repeated, uneven illumination below complex overburden can strongly distort time-lapse seismic amplitudes.

Since different geometries were used for all surveys in the numerical test, deterioration of the time-lapse amplitudes is due to a combined effect of geometry and complex overburden. The illumination-ratio maps (Figure 5) show the variability of illumination between surveys. Disparities in point-spread functions (Figure 8) suggest that the a diagonal approximation to the Hessian is insufficient to remove the unwanted artifacts. Although separately inverted time-lapse images in Figure 11 show some improvement in resolution over migration results (Figure 10), the images are dominated by the large amplitude residual artifacts. Time-lapse images obtained from joint-inversion using the RJID (Figure 12), and RJMI (Figure 13) formulations are less noisy than those from the migration (Figure 10) and separate inversion (Figure 11) and are comparable to the true reflectivity change in Figure 9. Since one-way operators were used in this study, the Hessian contains no information regarding secondary

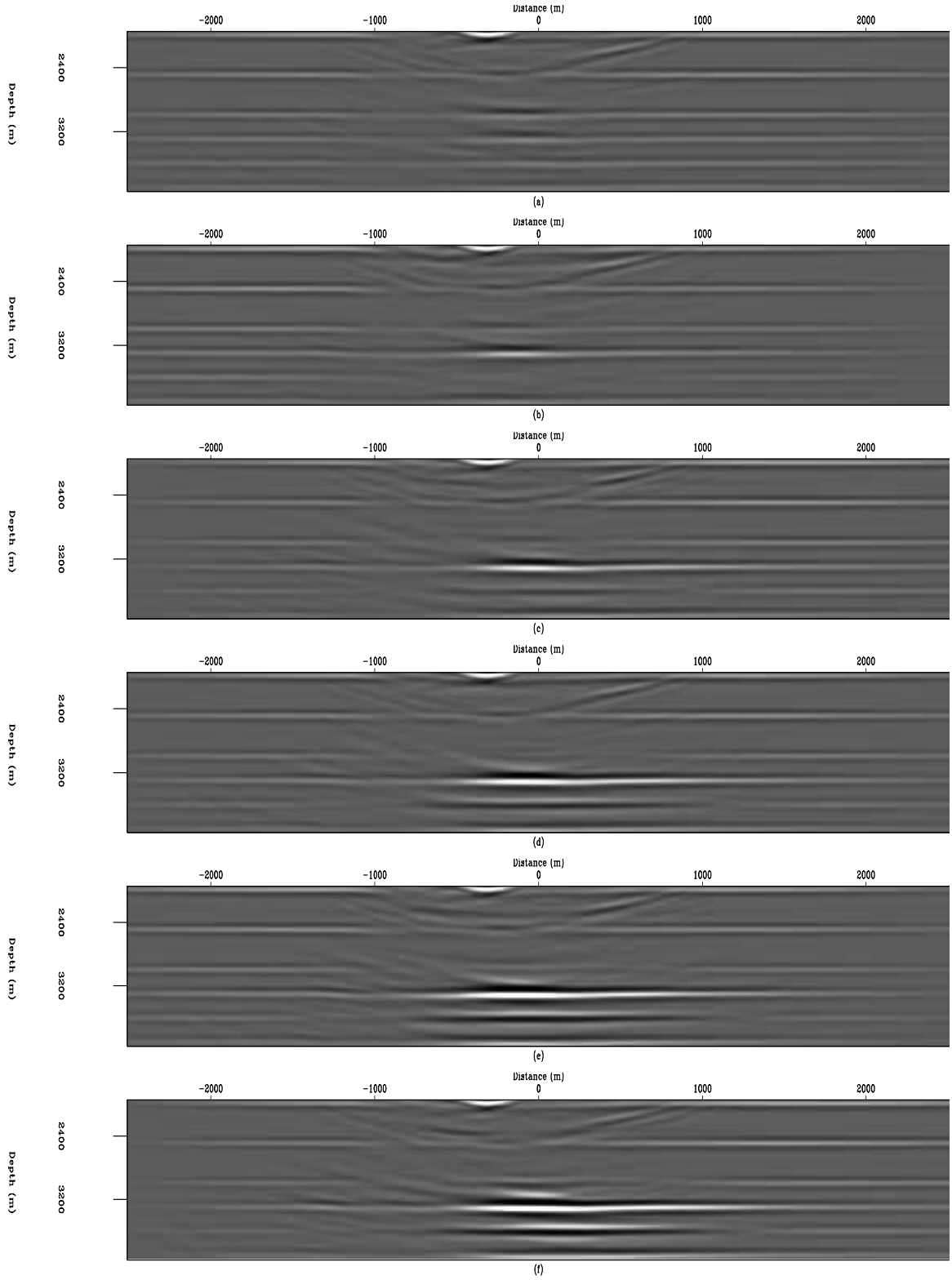


Figure 4: Migrated images obtained for six surveys (see modeling parameters in Table 1) for the target area in shown Figure 1. Figure (a) is the baseline image, while Figures (b)-(f) are images of the monitor images. Note the irregular amplitude patterns cause by the presence of the overlying salt structure.[**CR**] gayeni1/. migs

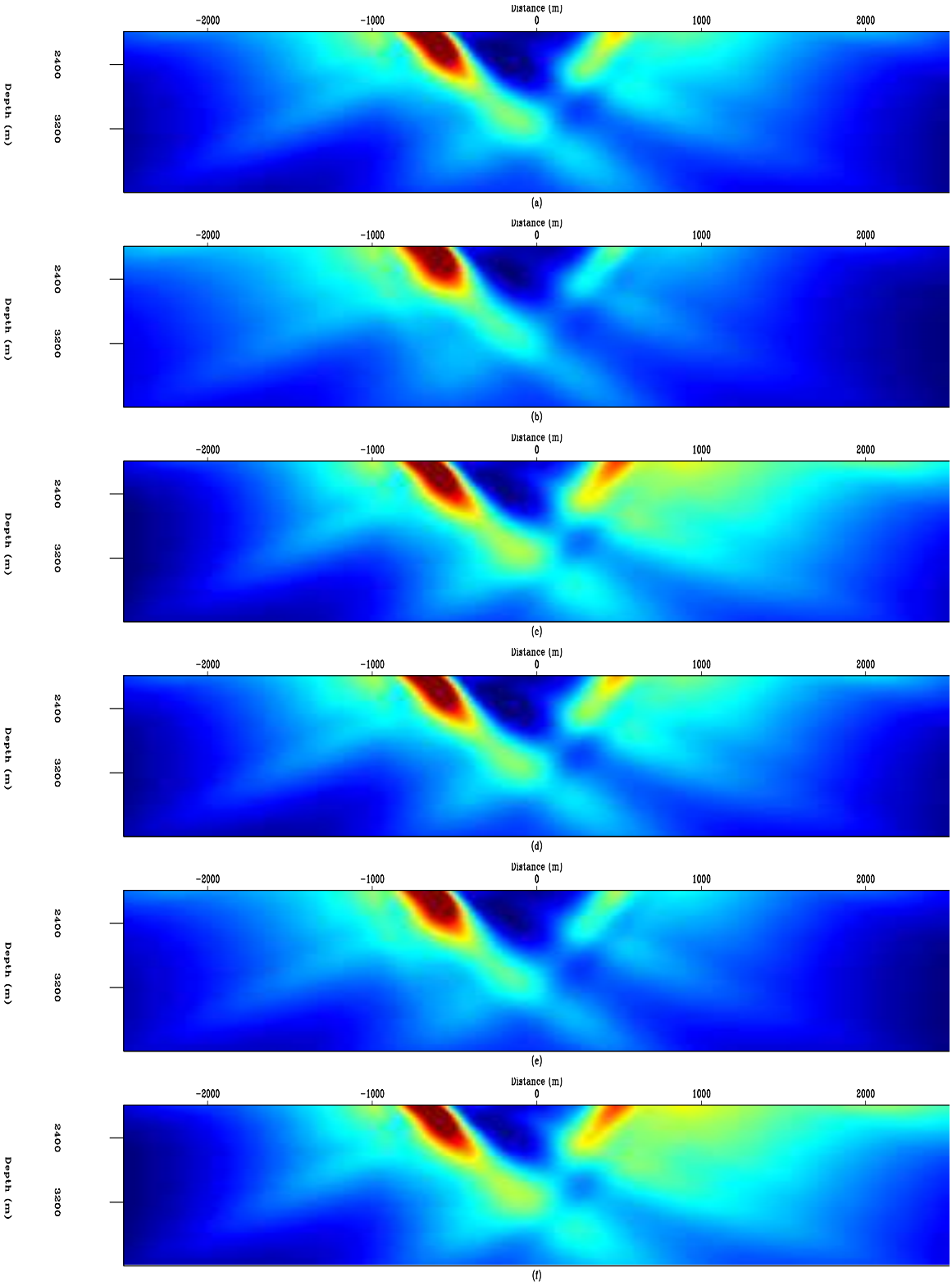


Figure 5: Illumination maps for the six surveys described in Table 1. Each section corresponds to the migrated sections in Figure 4 and explain the observed irregular seismic amplitudes. Light color represent high illumination and dark represents low illumination.[**CR**]
gayeni1/. illums-salt

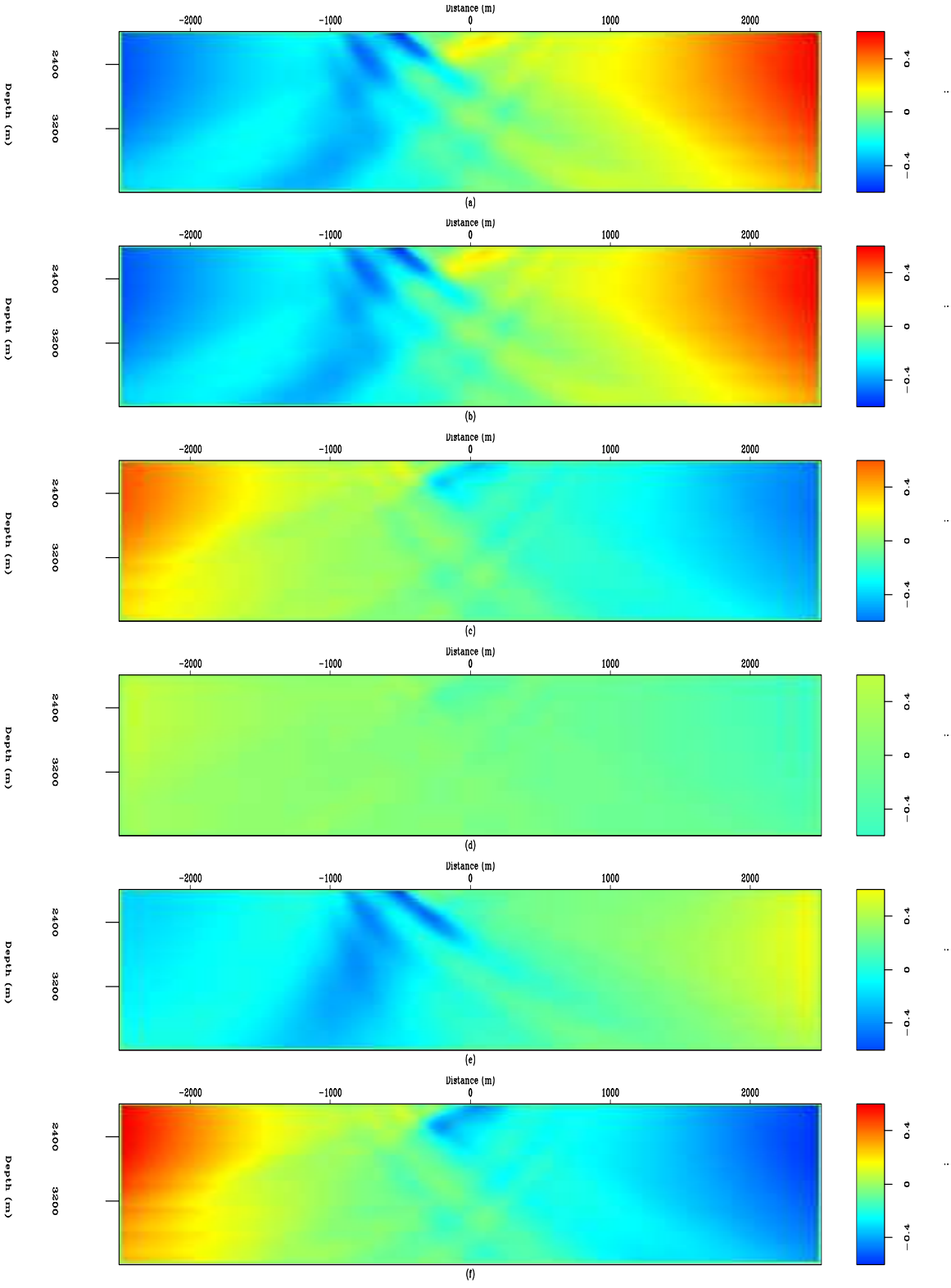


Figure 6: Base-Monitor illumination-ratio for the six surveys described in Table 1. Each section corresponds to a normalized rms-ratio (over 3x3 patches) between the Hessian diagonal of the monitor surveys (Figure 5b-f) to that of the baseline (Figure 5a).[**CR**] gayeni1/. illums-salt-r

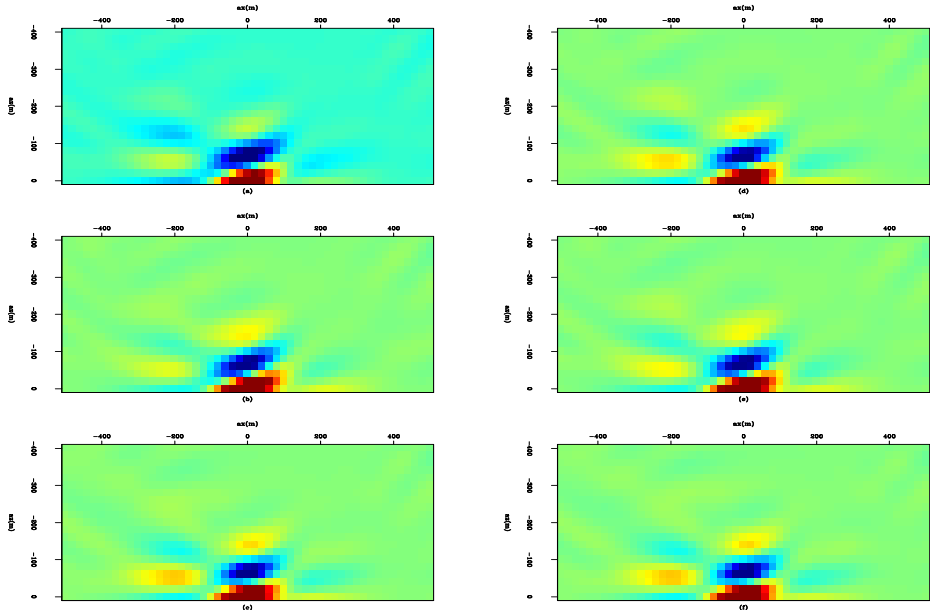


Figure 7: Point spread functions at the image point $[x = -200m, z = 2800m]$ for the six surveys in Figure 5. Each section corresponds to a row of the Hessian for the images in Figure 4a-f). Note that only one half of each filter is computed, since the Hessian matrix is symmetric.[**CR**] gayeni1/. hess-salt-off

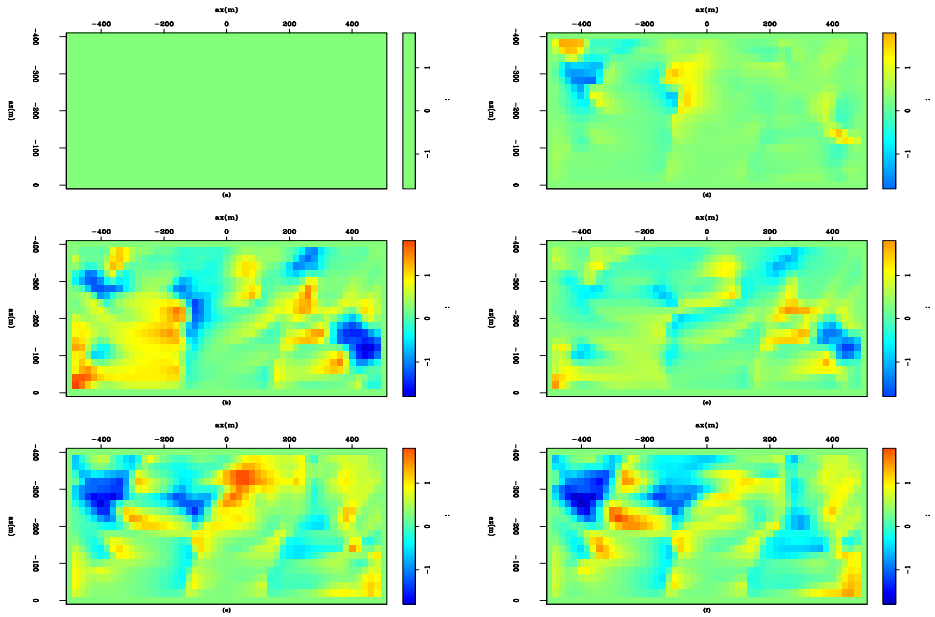


Figure 8: Base-Monitor illumination-ratio at image point $[x = -200m, z = 2800m]$ for the six surveys described in Table 1. section corresponds to a normalized rms-ratio (over 3x3 patches) between the point-spread functions (a row of the Hessian) of the monitor surveys (Figure 7b-f) to that of the baseline (Figure 7a).[**CR**] gayeni1/. illums-salt-off

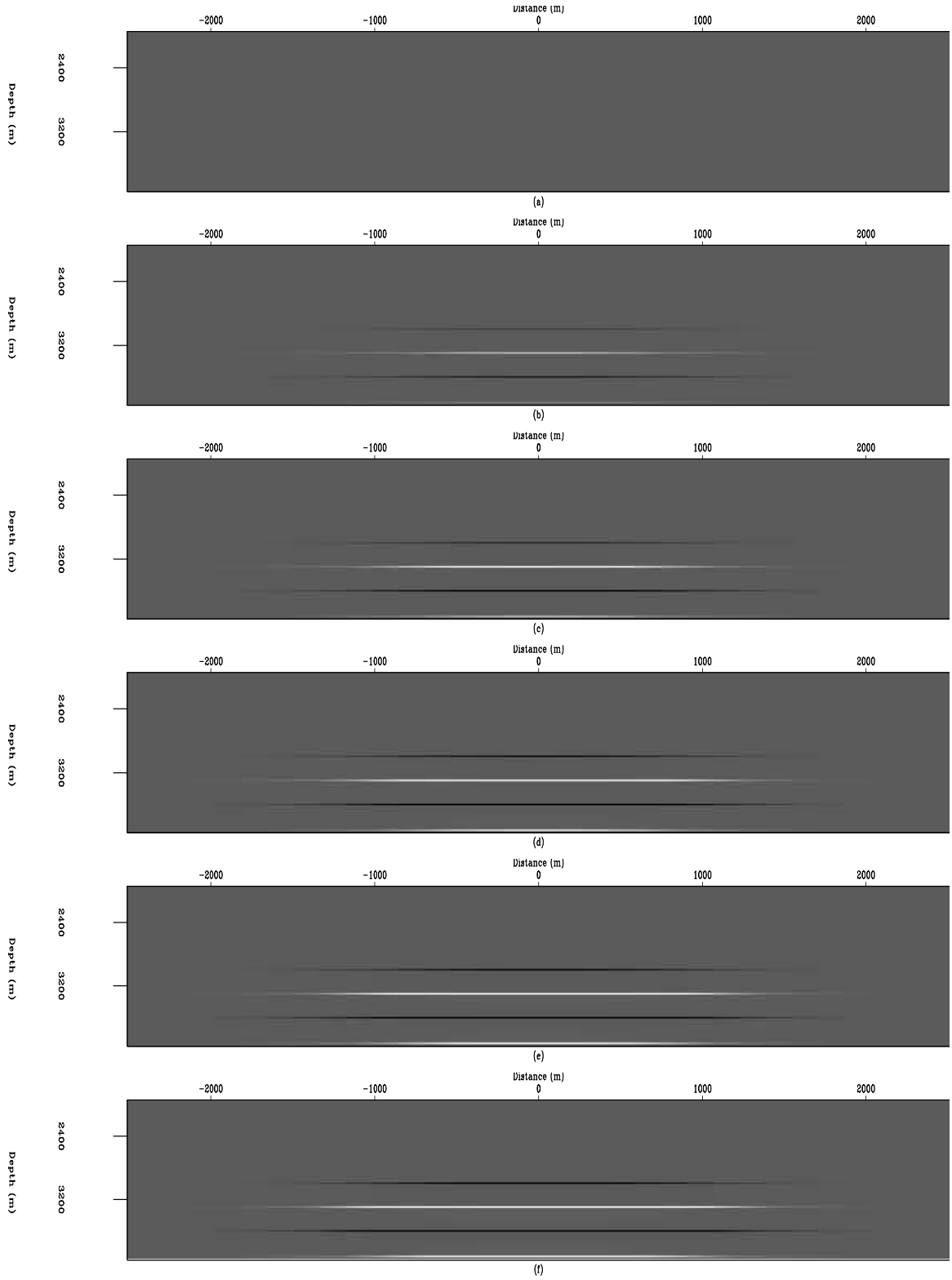


Figure 9: True cumulative time-lapse reflectivity images at the times for which the six surveys (Table 1) were modeled. These should be compared with the results in Figures 10 to 13.[**CR**] gayeni1/. refs4d

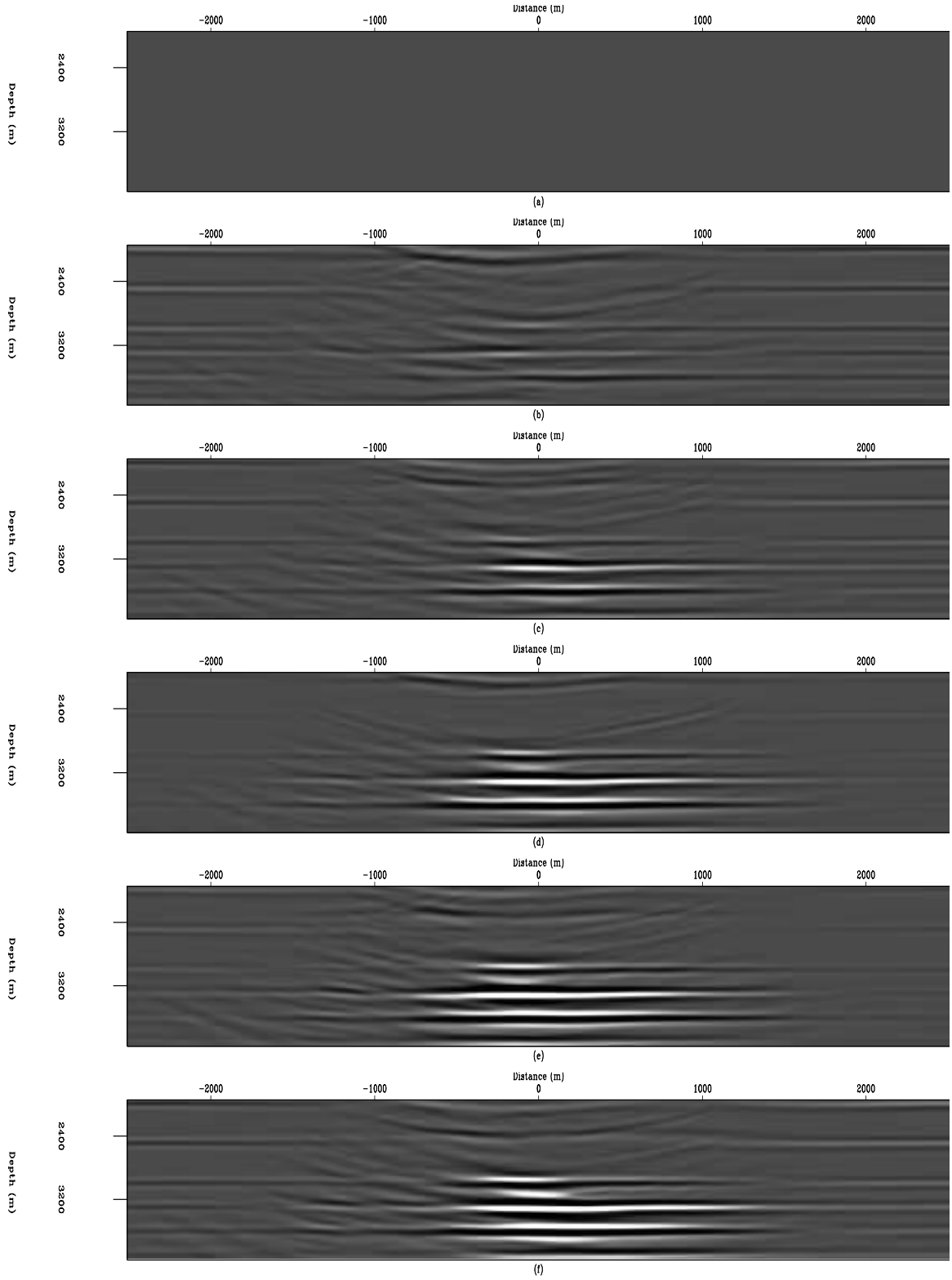


Figure 10: Migrated time-lapse at six different times corresponding to Figure 4. Each section shows the amplitude change between time 1 (baseline) and the time of the monitor survey. Note that the inversion has resulted in an increase in the noise amplitudes relative to migrated time-lapse images shown in Figure 10.[**CR**] gayeni1/. migs4d

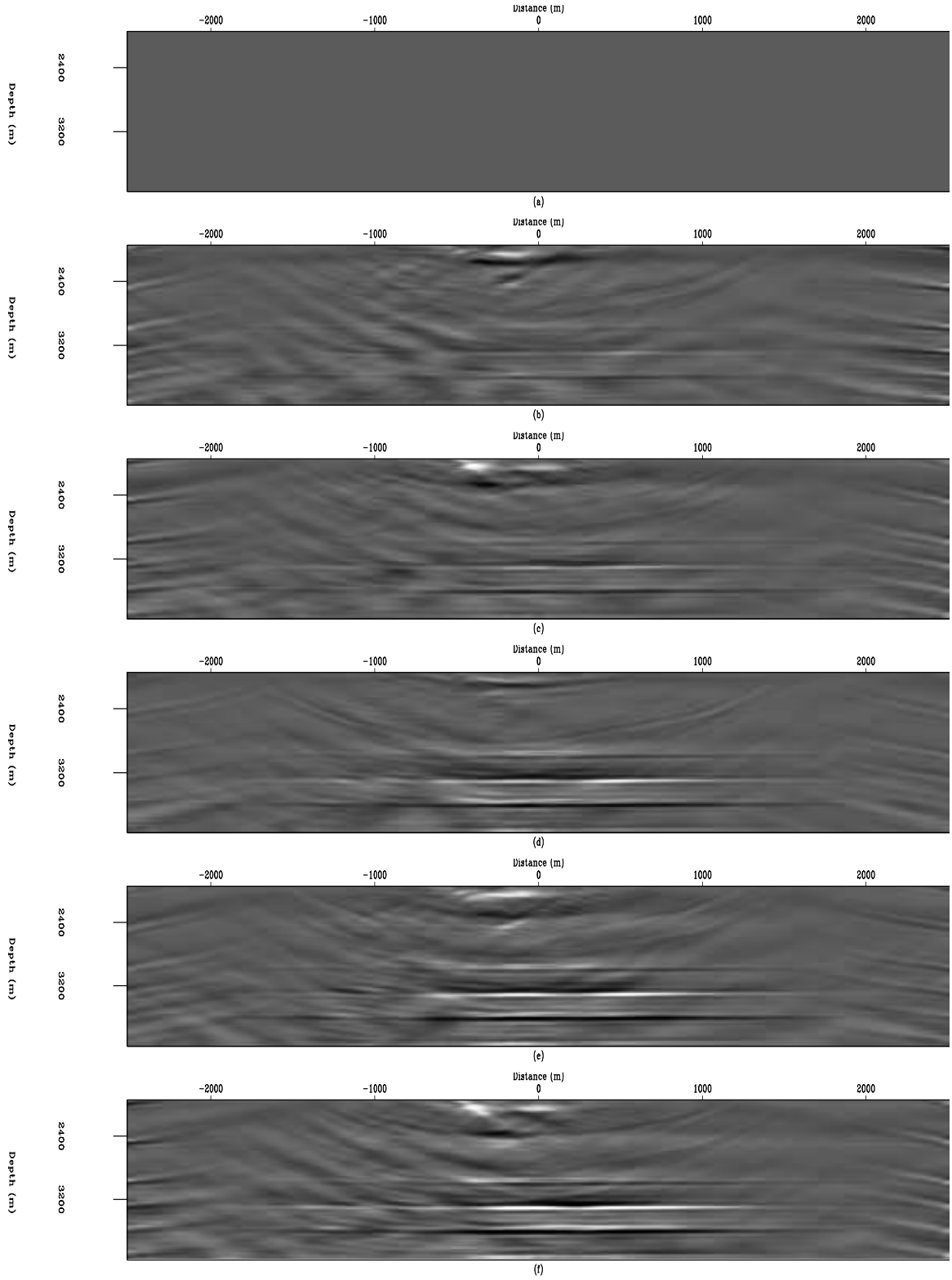


Figure 11: Separately inverted time-lapse images at six different times. Each section shows the amplitude change between time 1 (baseline) and the time of the monitor survey. Note that the inversion has resulted in an increase in the noise amplitudes relative to migrated time-lapse images shown in Figure 10.[**CR**] gayeni1/. invs4d

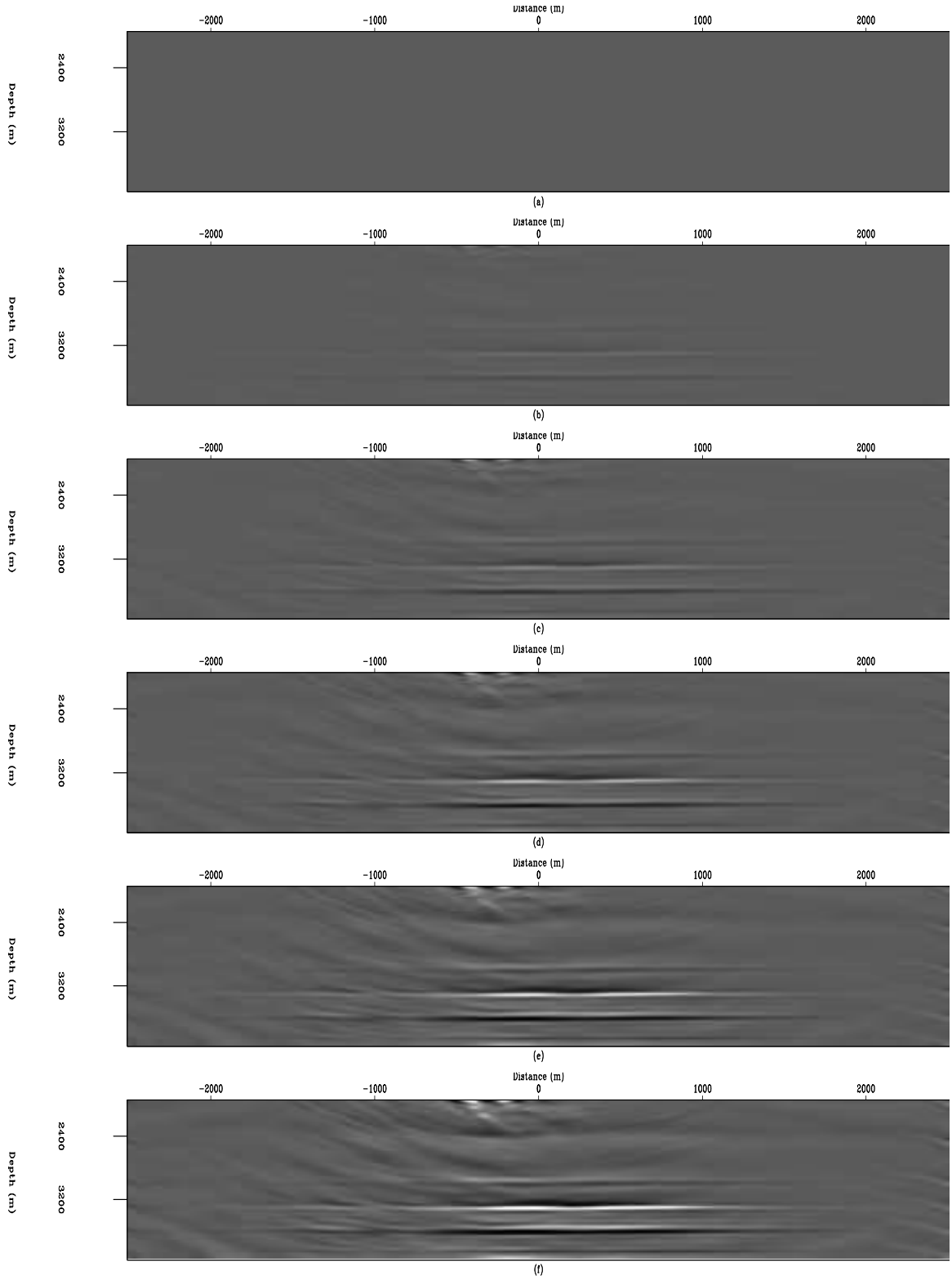


Figure 12: RJID: Jointly inverted time-lapse images at six different times. Each section shows the amplitude change between time 1 (baseline) and the time of the monitor survey. Compare these results to Figures 9, 10, 11 and 13.[**CR**] gayeni1/. inv14d

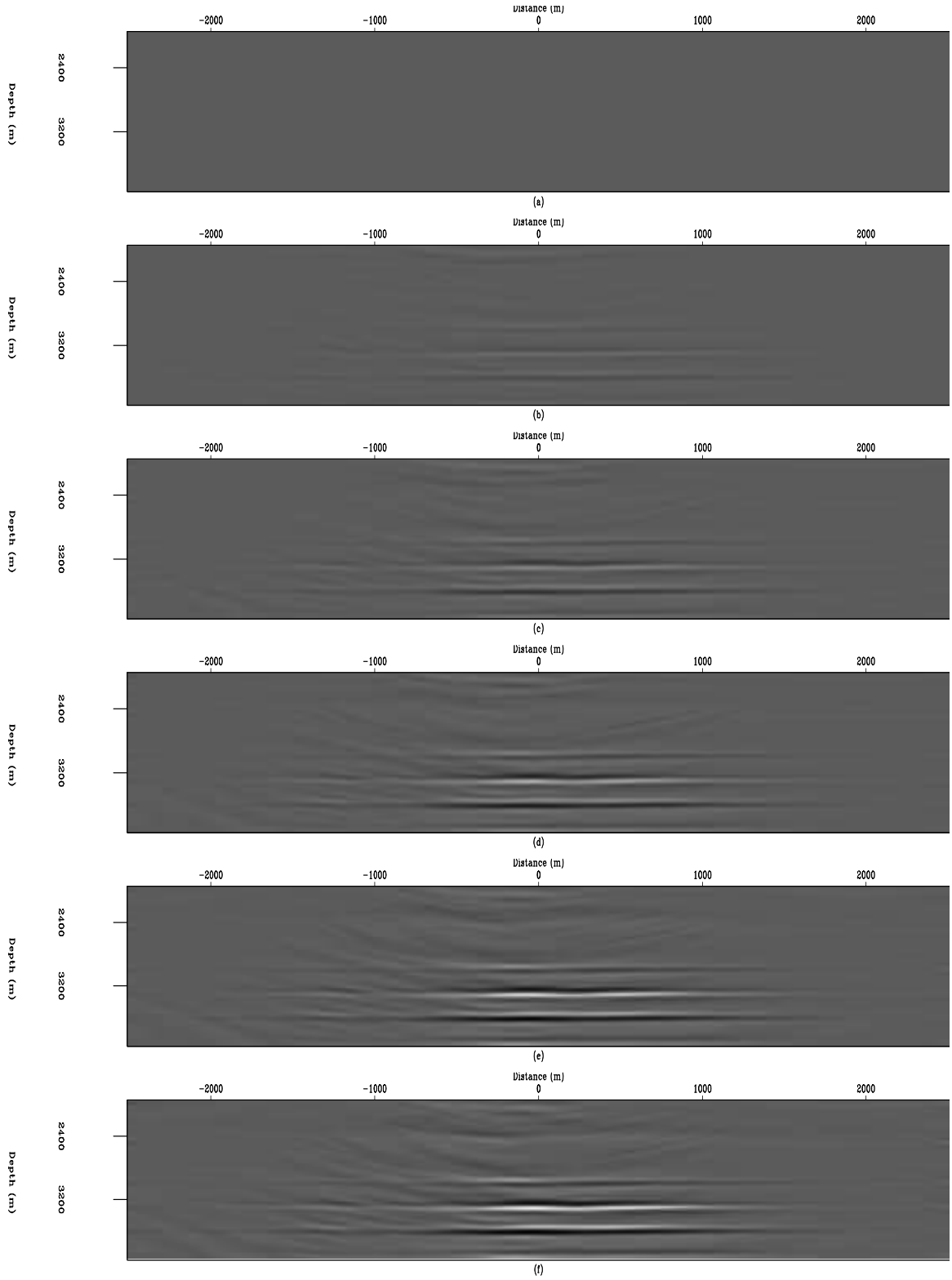


Figure 13: RJMI: Jointly inverted time-lapse images at six different times. Each section shows the amplitude change between time 1 (baseline) and the time of the monitor survey. Compare these results to Figures 9, 10, 11 and 12.[**CR**] gayeni1/. inv24d

events such as multiples. Residual effects of internal multiples due to the salt persist in the inversion results but are significantly suppressed in the jointly inverted images. In many cases of interest, multiple energy will be sufficiently attenuated or may not fall within the target region, and thus have little impact on the inversion.

CONCLUSIONS

We have proposed two regularized least-squares inversion formulations for time-lapse seismic imaging. -hese formulations arise from the linearized least-square wave-equation inversion. By using a target-oriented approximation to the least-squares Hessian with appropriate spatial and temporal regularization, we have shown that image difference due to geometry dissimilarity and complex overburden can be attenuated. We show that we can directly invert for image differences (RJID) or multiple images (RJMI) using a concatenation of target-oriented Hessian operators and combinations of migrated images. From numerical tests using a synthetic 2D-subsalt model, we conclude that both the RJID and RJMI joint-inversion formulations give more accurate time-lapse images than either migration or separate inversion.

REFERENCES

Ajo-Franklin, J. B., J. Urban, and J. M. Harris, 2005, Temporal integration of seismic traveltime tomography: SEG Technical Program Expanded Abstracts, **24**, 2468–2471.

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3345–3349.

Ayeni, G. and B. Biondi, 2008, Time-lapse inversion: SEP 134.

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).

Clapp, M. L., 2005, Imaging under salt: illumination compensation by regularized inversion: PhD thesis, Stanford University.

Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024.

Hale, D., 2007, A method for estimating apparent displacement vectors from time-lapse seismic images: CWP Report-566.

Johnston, D., 2005, Time-lapse 4D technology: Reservoir surveillance: AAPG Search and Discovery.

Kindermann, S. and A. Leitao, 2007, Regularization by dynamic programming: Journal of Inverse and Ill-posed Problems, **15**, 295–310.

Kuhl, H. and M. Sacchi, 2001, Generalized least-squares dsr migration using a common angle imaging condtion: 71st Annual International Meeting, Expanded Abstracts, 1025–1028, SEG.

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., D. C. Adams, M. Meadows, S. Cole, and R. Wright, 2003, 4d seismic data processing issues and examples: SEG Technical Program Expanded Abstracts, **22**, 1394–1397.

Lumley, D. E., 1995, Seismic time-lapse monitoring of subsurface fluid flow: PhD thesis, Stanford University, http://sepwww.stanford.edu/public/docs/sep91/.

Mulder, W. and R.-E. Plessix, 2004, Frequency-domain finite-frequency amplitude-preserving migration: Geophysical Journal International, **157**, 975–985.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

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.

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.

Schmitt, U. and A. K. Louis, 2002, Efficient algorithms for the regularization of dynamic inverse problems: I. theory: Inverse Problems, **18**, 645–658.

Schmitt, U., A. K. Louis, C. Wolters, and M. Vauhkonen, 2002, Efficient algorithms for the regularization of dynamic inverse problems: Ii. applications: Inverse Problems, **18**, 659–676.

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., 2008a, Modeling, migration and inversion in the generalized source and receiver domain: SEP 136.

———, 2008b, Wave-equation hessian by phase encoding: SEP 134.

Valenciano, A., 2008, Imaging by Wave-Equation Inversion: PhD thesis, Stanford University.

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, Y., A. Ghodrati, and D. H. Brooks, 2005, An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework: Inverse Problems, **21**, 357–382.

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 FORMULATIONS FOR MULTIPLE SURVEYS

Here, we show a brief derivation of the joint-inversion formulation for two surveys and its generalization to multiple surveys.

Regularized joint-inversion

The process of acquiring two seismic datasets over an evolving earth model can be represented as

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}, \tag{A-1}$$

where $\mathbf{d}_0$ and $\mathbf{d}_1$ are the baseline and monitor datasets, and $\mathbf{m}_0$ and $\mathbf{m}_1$ are the baseline and monitor reflectivity models respectively. The linear operators ($\mathbf{L}_0$ and $\mathbf{L}_1$) define the modeling/acquisition experiments for datasets $\mathbf{d}_0$ and $\mathbf{d}_1$ respectively. We rewrite equation A-1 to include spatial and regularization operators ($\mathbf{R}_0$ and $\mathbf{\Lambda}_0$ respectively), and we seek to minimize the objective function

$$S(\mathbf{m_0}, \mathbf{m_1}) = \left\| \left[\begin{array}{cc} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \\ \hline \mathbf{R}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_1 \\ \hline -\mathbf{\Lambda}_0 & \mathbf{\Lambda}_1 \end{array} \right] \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right\|^2. \tag{A-2}$$

This cost function can be expanded as follows:

$$\left(\begin{bmatrix} \mathbf{L'_0 L_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L'_1 L_1} \end{bmatrix} + \begin{bmatrix} \mathbf{R'_0 R_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R'_1 R_1} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda'_0 \Lambda_0} & \mathbf{-\Lambda'_0 \Lambda_1} \\ \mathbf{-\Lambda'_1 \Lambda_0} & \mathbf{\Lambda'_1 \Lambda_1} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{A-3}$$

which can be written as

$$\left(\begin{bmatrix} \mathbf{H_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{-\Lambda_{01}} \\ \mathbf{-\Lambda_{10}} & \mathbf{\Lambda_{11}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}, \tag{A-4}$$

where

$$\begin{aligned} \mathbf{R_{ij}} &= \mathbf{\epsilon_i R'_i \epsilon_j R_j}, \\ \mathbf{\Lambda_{ij}} &= \mathbf{\zeta_i \Lambda'_i \zeta_j \Lambda_j}, \end{aligned} \tag{A-5}$$

while $\mathbf{R_0}$ and $\mathbf{R_1}$ are the spatial/imaging constraints for the baseline and monitor images respectively, and $\mathbf{\Lambda_0}$ and $\mathbf{\Lambda_1}$ the temporal constraints between the surveys. The parameters $\mathbf{\epsilon_0}$ and $\mathbf{\epsilon_1}$ determine the strength of the spatial regularization on the baseline and monitor images respectively, while $\mathbf{\zeta_0}$ and $\mathbf{\zeta_1}$ determine the coupling between surveys. Equation A-4 is the RJMI formulation. Using a similar procedure, the RJID formulation for two seismic datasets can be shown to be

$$\left(\begin{bmatrix} \mathbf{H_0 + H_1} & \mathbf{H_1} \\ \mathbf{H_1} & \mathbf{H_1} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{0} \\ \mathbf{-\Lambda_{10}} & \mathbf{\Lambda_{11}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \tag{A-6}$$

In the next sections, we derive the RJID and RJMI formulations for multiple surveys.

Regularized joint-inversion of multiple images: RJMI

The data modeling process for three seismic datasets (a baseline and two monitors) over an evolving earth model can be written as

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_2 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \tag{A-7}$$

where $\mathbf{d}_0$, $\mathbf{d}_1$ and $\mathbf{d}_2$ are respectively datasets for the baseline, first and second monitor, $\mathbf{m}_0$, $\mathbf{m}_1$, and $\mathbf{m}_2$ are the baseline and monitor reflectivity models. The linear operators ($\mathbf{L}_0$, $\mathbf{L}_1$ and $\mathbf{L}_2$) define the modeling/acquisition experiments for datasets $\mathbf{d}_0$, $\mathbf{d}_1$ and $\mathbf{d}_2$ respectively. The least-squares solution to equation A-7 is given as

$$\begin{bmatrix} \mathbf{L'_0 L_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L'_1 L_1} & \mathbf{0} \\ \mathbf{0} & 0 & \mathbf{L'_2 L_2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \\ \hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{L'}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L'}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{L'}_2 \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \tag{A-8}$$

where the symbol $'$ denotes transposed complex conjugate.

We rewrite equation A-8 as

$$\begin{bmatrix} \mathbf{H_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H_2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \\ \hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m_0}} \\ \tilde{\mathbf{m_1}} \\ \tilde{\mathbf{m_2}} \end{bmatrix}, \tag{A-9}$$

where $\tilde{\mathbf{m}}_\mathbf{i}$ is the migrated image from the $\mathbf{i}^{th}$ survey, and $\mathbf{H_i}$ is the corresponding Hessian matrix. Introducing spatial and temporal constraints into equation A-9 we obtain

$$\left(\begin{bmatrix} \mathbf{H_0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H_2} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_{22}} \end{bmatrix} + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{-\Lambda_{01}} & \mathbf{0} \\ \mathbf{-\Lambda_{10}} & \mathbf{2\Lambda_{11}} & \mathbf{-\Lambda_{12}} \\ \mathbf{0} & \mathbf{-\Lambda_{21}} & \mathbf{\Lambda_{22}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \\ \hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m_0}} \\ \tilde{\mathbf{m_1}} \\ \tilde{\mathbf{m_2}} \end{bmatrix}. \tag{A-10}$$

Equation A-10 can be generalized to an arbitrary number of surveys as follows

$$[\mathbf{\Xi} + \Re + \mathbf{\Gamma}] \left[\hat{\mathbf{M}}\right] = \left[\tilde{\mathbf{M}}\right], \tag{A-11}$$

where, $\mathbf{\Xi}$ is the Hessian operator, defined as

$$\mathbf{\Xi} = \begin{bmatrix} \mathbf{H_0} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H_2} & ... & ... & \mathbf{0} \\ : & : & : & : & : & : \\ : & : & : & : & \mathbf{H_{N-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{H_N} \end{bmatrix}. \tag{A-12}$$

The spatial and temporal regularization operators, $\Re$ and $\mathbf{\Gamma}$ are defined as

$$\begin{aligned} \Re &= \mathbf{R'R}, \\ \mathbf{\Gamma} &= \mathbf{\Lambda'\Lambda}, \end{aligned} \tag{A-13}$$

where,

$$\mathbf{R} = \begin{bmatrix} \mathbf{R_0} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_1} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_2} & ... & \mathbf{0} & \mathbf{0} \\ : & : & : & : & : & : \\ : & : & : & \mathbf{0} & \mathbf{R_{N-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R_N} \end{bmatrix}, \tag{A-14}$$

and,

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda_0} & \mathbf{\Lambda_1} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{-\Lambda_1} & \mathbf{\Lambda_2} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{-\Lambda_2} & \mathbf{-\Lambda_3} & ... & \mathbf{0} \\ : & : & : & : & : & : \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{-\Lambda_{N-1}} & \mathbf{\Lambda_N} \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} & \mathbf{\Lambda_N} \end{bmatrix}. \tag{A-15}$$

The input vector into the RJMI formulation, $\tilde{\mathbf{M}}$ is given as

$$\tilde{\mathbf{M}} = \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_2 \\ : \\ \tilde{\mathbf{m}}_{N-1} \\ \tilde{\mathbf{m}}_N \end{bmatrix}, \tag{A-16}$$

while the inversion targets are in the vector:

$$\hat{\mathbf{M}} = \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \\ \hat{\mathbf{m}}_2 \\ : \\ \hat{\mathbf{m}}_{N-1} \\ \hat{\mathbf{m}}_N \end{bmatrix}. \tag{A-17}$$

Regularized joint-inversion for image differences: RJID

The data modeling process for three seismic datasets over an evolving earth model can be written as

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 & \mathbf{0} \\ \mathbf{L}_2 & \mathbf{L}_2 & \mathbf{L}_2 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \Delta\mathbf{m}_1 \\ \Delta\mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \tag{A-18}$$

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

$$\begin{aligned} \Delta\mathbf{m}_1 &= \mathbf{m}_1 - \mathbf{m}_0, \\ \Delta\mathbf{m}_2 &= \mathbf{m}_2 - \mathbf{m}_1, \end{aligned} \tag{A-19}$$

where $\mathbf{m}_1$ and $\mathbf{m}_2$ are respectively the monitor reflectivities at the times data $\mathbf{d}_1$ and $\mathbf{d}_2$ were acquired (with survey geometries defined by the linear $\mathbf{L}_1$ and $\mathbf{L}_2$).

The least-squares solution to equation A-18 is given as

$$\begin{bmatrix} \mathbf{L'_0L_0 + L'_1L_1 + L'_2L_2} & \mathbf{L'_1L_1 + L'_2L_2} & \mathbf{L'_2L_2} \\ \mathbf{L'_1L_1 + L'_2L_2} & \mathbf{L'_1L_1 + L'_2L_2} & \mathbf{L'_2L_2} \\ \mathbf{L'_2L_2} & \mathbf{L'_2L_2} & \mathbf{L'_2L_2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{L'}_0 & \mathbf{L'}_1 & \mathbf{L'}_2 \\ & \mathbf{L'}_1 & \mathbf{L'}_2 \\ & & \mathbf{L'}_2 \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \tag{A-20}$$

where the symbol $'$ denotes transpose complex conjugate. We rewrite equation A-20 as

$$\begin{bmatrix} \mathbf{H_0+H_1+H_2} & \mathbf{H_1+H_2} & \mathbf{H_2} \\ \mathbf{H_1+H_2} & \mathbf{H_1+H_2} & \mathbf{H_2} \\ \mathbf{H_2} & \mathbf{H_2} & \mathbf{H_2} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m_0}}+\tilde{\mathbf{m_1}}+\tilde{\mathbf{m_2}} \\ \tilde{\mathbf{m_1}}+\tilde{\mathbf{m_2}} \\ \tilde{\mathbf{m_2}} \end{bmatrix}, \tag{A-21}$$

where $\tilde{\mathbf{m_i}}$ is the migrated image from the $\mathbf{i}^{th}$ survey, and $\mathbf{H_i}$ is the corresponding Hessian matrix. Introducing spatial and temporal regularization goals that incorporates prior knowledge of the reservoir geometry and location as well as constraints on the inverted time-lapse images into equation A-21 we obtain

$$\begin{aligned} \left(\begin{bmatrix} \mathbf{H_0+H_1+H_2} & \mathbf{H_1+H_2} & \mathbf{H_2} \\ \mathbf{H_1+H_2} & \mathbf{H_1+H_2} & \mathbf{H_2} \\ \mathbf{H_2} & \mathbf{H_2} & \mathbf{H_2} \end{bmatrix} + \begin{bmatrix} \mathbf{R_{00}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_{22}} \end{bmatrix} \right. \\ \left. + \begin{bmatrix} \mathbf{\Lambda_{00}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{2\Lambda_{11}} & \mathbf{-\Lambda_{12}} \\ \mathbf{0} & \mathbf{-\Lambda_{21}} & \mathbf{\Lambda_{22}} \end{bmatrix} \right) \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m_0}}+\tilde{\mathbf{m_1}}+\tilde{\mathbf{m_2}} \\ \tilde{\mathbf{m_1}}+\tilde{\mathbf{m_2}} \\ \tilde{\mathbf{m_2}} \end{bmatrix}, \end{aligned} \tag{A-22}$$

where,

$$\begin{aligned} \mathbf{R_{ij}} &= \epsilon_\mathbf{i}\mathbf{R'}_\mathbf{i}\epsilon_\mathbf{j}\mathbf{R_j} \\ \mathbf{\Lambda_{ij}} &= \zeta_\mathbf{i}\mathbf{\Lambda'}_\mathbf{i}\zeta_\mathbf{j}\mathbf{\Lambda_j} \end{aligned}, \tag{A-23}$$

with $\mathbf{R_i}$ being the spatial regularization terms for the baseline and time-lapse images respectively while $\mathbf{\Lambda_i}$ is the temporal regularization between the surveys. Note that $\mathbf{R_{ij}}$ and $\mathbf{\Lambda_{ij}}$ are not explicitly computed, but instead, the regularization operators $\mathbf{R_i}$ and $\mathbf{\Lambda_i}$ (and their adjoints) are applied at each step of the inversion. Parameters ϵ_i and ζ_i determine the relative strengths of the spatial and temporal regularization respectively. Equation A-22 can be generalized to an arbitrary number of surveys as follows

$$[\mathbf{\Xi}+\Re+\mathbf{\Gamma}]\left[\hat{\mathbf{M}}\right] = \left[\tilde{\mathbf{M}}\right], \tag{A-24}$$

where, $\mathbf{\Xi}$ is the Hessian operator, defined as

$$\mathbf{\Xi} = \begin{bmatrix} \mathbf{H_0+..+H_N} & \mathbf{H_1+..+H_N} & \mathbf{H_2+..+H_N} & ... & \mathbf{H_{N-1}+H_N} & \mathbf{H_N} \\ \mathbf{H_1+..+H_N} & \mathbf{H_1+..+H_N} & \mathbf{H_2+..+H_N} & ... & \mathbf{H_{N-1}+H_N} & \mathbf{H_N} \\ \mathbf{H_2+..+H_N} & \mathbf{H_2+..+H_N} & \mathbf{H_2+..+H_N} & ... & ... & \mathbf{H_N} \\ : & : & : & : & : & : \\ : & : & : & : & \mathbf{H_{N-1}+H_N} & \mathbf{H_N} \\ \mathbf{H_N} & \mathbf{H_N} & \mathbf{H_N} & ... & \mathbf{N_N} & \mathbf{H_N} \end{bmatrix}. \tag{A-25}$$

The regularization operators $\Re$ and $\mathbf{\Gamma}$ are defined as

$$\begin{aligned} \Re &= \mathbf{R'R}, \\ \mathbf{\Gamma} &= \mathbf{\Lambda'\Lambda}, \end{aligned} \tag{A-26}$$

where,

$$\mathbf{R} = \begin{bmatrix} \mathbf{R_0} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_1} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_2} & ... & \mathbf{0} & \mathbf{0} \\ : & : & : & : & : & : \\ : & : & : & \mathbf{0} & \mathbf{R_{N-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R_N} \end{bmatrix}, \tag{A-27}$$

and,

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda_0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{-\Lambda_1} & \mathbf{\Lambda_2} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{-\Lambda_2} & \mathbf{-\Lambda_3} & ... & \mathbf{0} \\ : & : & : & : & : & : \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{-\Lambda_{N-1}} & \mathbf{\Lambda_N} \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} & \mathbf{\Lambda_N} \end{bmatrix}. \tag{A-28}$$

The input vector into the RJID formulation, $\tilde{\mathbf{M}}$ is given as

$$\tilde{\mathbf{M}} = \begin{bmatrix} \tilde{\mathbf{m_0}} + ... + \tilde{\mathbf{m}}_N \\ \tilde{\mathbf{m_1}} + ... + \tilde{\mathbf{m}}_N \\ \tilde{\mathbf{m_2}} + ... + \tilde{\mathbf{m}}_N \\ : \\ \tilde{\mathbf{m}}_{N-1} + \tilde{\mathbf{m}}_N \\ \tilde{\mathbf{m}}_N \end{bmatrix}, \tag{A-29}$$

while the inversion targets are

$$\hat{\mathbf{M}} = \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \Delta\hat{\mathbf{m}}_1 \\ \Delta\hat{\mathbf{m}}_2 \\ : \\ : \\ \Delta\hat{\mathbf{m}}_N \end{bmatrix}. \tag{A-30}$$

The temporal constraint on the baseline image, $\mathbf{\Lambda_0}$ may be set to zero, since it is assumed that the original geological structure is unchanged over time or that geomechanical changes are accounted for before/during inversion.

Modeling, migration, and inversion in the generalized source and receiver domain

Yaxun Tang

ABSTRACT

I extend the theory of Born modeling, migration and inversion to the generalized source and receiver domain, a transformed domain that is obtained by linear combination of the encoded sources and receivers. I show how to construct the target-oriented imaging Hessian with encoded sources, with encoded receivers and with simultaneously encoded sources and receivers. I also demonstrate the connection of the imaging Hessian in the generalized source and receiver domain to the phase-encoded Hessian that I developed in *SEP-134*. As an application of the theory, I introduce a mixed phase-encoding scheme to compute the Hessian operator. The new scheme combines the advantages of both random phase encoding and plane-wave phase encoding. My preliminary tests of this new method on a simple model show promising results.

INTRODUCTION

Shot-profile migration is an accurate imaging technique. The computation is performed in each shot gather; thus it closely mimics the actual physical experiment. However, migrating shot by shot is expensive, since the number of shot gathers is usually very big for a typical seismic survey. To reduce the cost, Whitmore (1995), Zhang et al. (2005) and Duquet and Lailly (2006) develop plane-wave source or delayed-shot migration, which migrates a small number of synthesized shots made by linear combination of the original shot gathers after linear time delays. In fact, plane-wave source or delayed-shot migration is only a special case of a more general class of migration technique, phase-encoding migration (Romero et al., 2000; Liu et al., 2006), where the source encoding functions can be any type of phase functions, such as linear phase functions, random phase functions, etc.. Though the methods mentioned above involve encoding only the original sources, there is no reason that the receivers could not be encoded. For example, besides assuming tilted line sources (plane-wave source or delayed-shot migration), we can also assume the data are recorded by tilted line receivers, or assume both line sources and line receivers at the same time. Such ideas have been explored by Stoffa et al. (2006), who develop the ray-based asymptotic theory for plane-wave source migration, plane-wave receiver migration and plane-wave source and receiver migration. In this paper, I extend those ideas from the plane-wave domain to more general cases and unify them under the generalized source and receiver domain.

Another important aspect of imaging is how to preserve the amplitude information of the reflectors. It is widely known that because of the non-unitary nature of the Born forward modeling operator, its adjoint, the migration operator, can only preserve the kinematic information of the reflectors (Lailly, 1983). To better preserve the amplitude, the inversion should be used. Bleistein (2007) derives closed-form asymptotic inversion formulas based

on the synthesized shot gathers (e.g. plane-wave sources) under the assumption that the acquisition geometry is infinite. However, we never have infinite acquisition geometry in practice. In fact, the limited acquisition geometry is an important factor distorting the amplitude of the reflectors, especially in complex geologies; hence it should not be neglected.

In this paper, I also extend the target-oriented inversion theory (Valenciano, 2008) to the generalized source and receiver domain for limited acquisition geometry. The effect of limited acquisition geometry is then taken into account in the least-squares sense and corrected by the pseudo-inverse of the target-oriented Hessian, the second derivative of the least-squares misfit functional with respect to the model parameters (Plessix and Mulder, 2004; Valenciano, 2008; Tang, 2008). I demonstrate that in the generalized source and receiver domain, the target-oriented Hessian can be more efficiently computed without storing the Green's functions, which is a major obstacle for the Hessian computation in the original shot-profile domain. I also show that the Hessian obtained in the generalized source and receiver domain is essentially the same as the phase-encoded Hessian (Tang, 2008), the physics of which, however, was not carefully discussed in Tang (2008). Therefore, from this perspective, this paper completes the discussion of the phase-encoded Hessian from the physical point of view. The modeling, migration and target-oriented Hessian formulas are all derived in terms of Green's functions, so that any type of Green's functions can be used under this framework, such as ray-based asymptotic Green's functions, Green's functions obtained by solving one-way wave equations, and Green's functions obtained by solving two-way wave equations. Anisotropy can also be taken into account, provided that the Green's functions are properly modeled.

This paper is organized as follows: I first briefly review the theory of Born modeling, migration and the target-oriented Hessian in the original shot-profile domain. Then I extend the theory to the encoded source domain, the encoded receiver domain, and the encoded source and receiver domain. Finally, I introduce a new phase-encoding scheme, which mixes both random and plane-wave phase encoding, to compute the Hessian operator. The new scheme combines advantages of both random phase encoding and plane-wave phase encoding. Finally, I apply the mixed phase-encoding scheme to a simple synthetic model.

BORN MODELING AND INVERSION IN THE SHOT-PROFILE DOMAIN

By using the Born approximation to the two-way wave equation, the primaries can be modeled by a linear operator as follows:

$$d(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_s, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) m(\mathbf{x}), \tag{1}$$

where $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is the modeled data for a single frequency ω with source and receiver located at $\mathbf{x}_s = (x_s, y_s, 0)$ and $\mathbf{x}_r = (x_r, y_r, 0)$ on the surface; $G(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions connecting the source and receiver, respectively, to the image point $\mathbf{x} = (x, y, z)$ in the subsurface; and $m(\mathbf{x})$ denotes the reflectivity at image point $\mathbf{x}$. In Equation 1, we assume $\mathbf{x}_s$ and $\mathbf{x}_r$ are infinite in extent and independent of each other. For a particular survey, however, we do not have infinitely long cable and infinitely many sources; thus we have to introduce an acquisition mask matrix to limit the size of the modeling. We

define

$$w(\mathbf{x}_r, \mathbf{x}_s) = \begin{cases} 1 & \text{if } \mathbf{x}_r \text{ is within the recording range of a shot at } \mathbf{x}_s; \\ 0 & \text{otherwise .} \end{cases} \tag{2}$$

For the marine acquisition geometry, $w(\mathbf{x}_r, \mathbf{x}_s)$ is similar to a band-limited diagonal matrix; for Ocean Bottom Cable (OBC) or land acquisition geometry, where all shots share the same receiver array, $w(\mathbf{x}_r, \mathbf{x}_s)$ is a rectangular matrix. Figure 1 illustrates the acquisition mask matrices for these two typical geometries in 2-D cases.

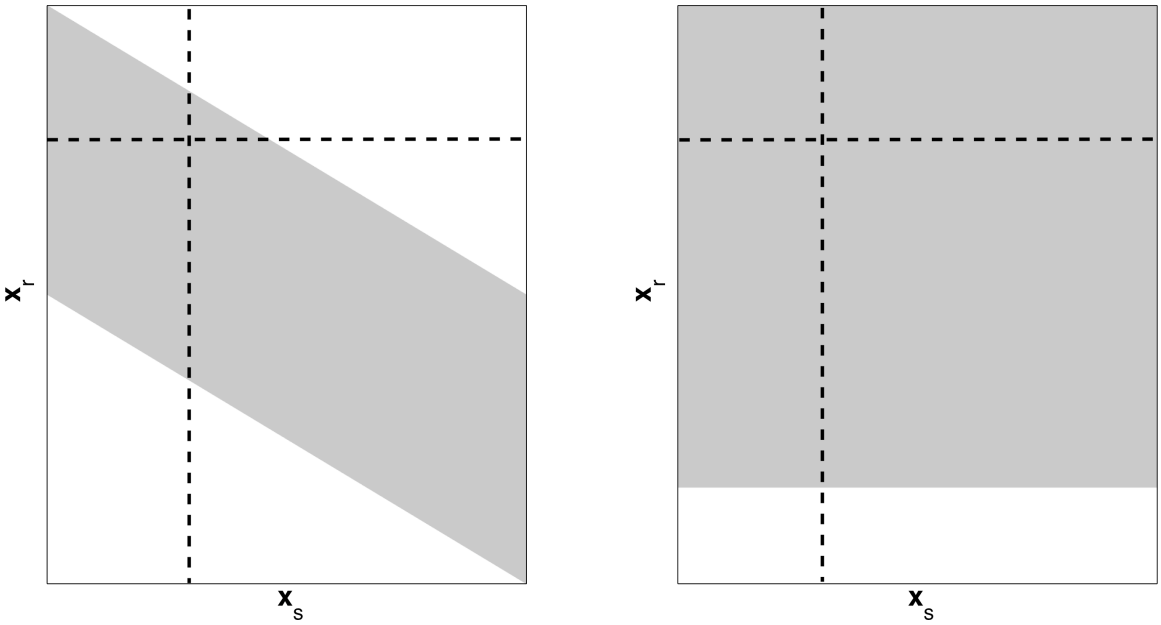


Figure 1: Acquisition mask matrices for different geometries in 2-D cases. Greys denote ones while whites denote zeros. The left panel shows the acquisition mask matrix for a typical marine acquisition geometry; the right panel shows the acquisition mask matrix for a typical OBC or land acquisition geometry. [**NR**] yaxun2/. acquisition-mask

To find a model that best fits the observed data, we can minimize the following data-misfit function in the least-squares sense:

$$J(m(\mathbf{x})) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} |w(\mathbf{x}_r, \mathbf{x}_s)(d(\mathbf{x}_r, \mathbf{x}_s, \omega) - d_{\text{obs}}(\mathbf{x}_r, \mathbf{x}_s, \omega))|^2. \tag{3}$$

The gradient of the above objective function gives the conventional shot-profile migration algorithm:

$$\nabla J(\mathbf{x}) = \Re \left(\sum_{\omega} \sum_{\mathbf{x}_s} G'(\mathbf{x}, \mathbf{x}_s, \omega) \sum_{\mathbf{x}_r} G'(\mathbf{x}, \mathbf{x}_r, \omega) w'(\mathbf{x}_r, \mathbf{x}_s) r(\mathbf{x}_r, \mathbf{x}_s, \omega) \right), \tag{4}$$

where $\Re$ denotes the real part of a complex number and $'$ means the complex conjugate; $r(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is the weighted residual defined as follows:

$$r(\mathbf{x}_r, \mathbf{x}_s, \omega) = w(\mathbf{x}_r, \mathbf{x}_s)(d(\mathbf{x}_r, \mathbf{x}_s, \omega) - d_{\text{obs}}(\mathbf{x}_r, \mathbf{x}_s, \omega)). \tag{5}$$

The gradient or migration is only a rough estimate of the model $m(\mathbf{x})$; to get a better recovery of the model space, the inverse of the Hessian, the second derivatives of the objective function, should be applied to the gradient:

$$\mathbf{m} \approx \mathbf{H}^{-1}\nabla\mathbf{J}. \tag{6}$$

The Hessian can be explicitly constructed by taking the second-order derivatives of the objective function with respect to the model parameters as follows (Plessix and Mulder, 2004; Valenciano, 2008; Tang, 2008):

$$\begin{aligned} H(\mathbf{x},\mathbf{y}) &= \Re\left(\sum_{\omega}\sum_{\mathbf{x}_s} G(\mathbf{x},\mathbf{x}_s,\omega)G'(\mathbf{y},\mathbf{x}_s,\omega)\ \times \right. \\ &\left. \sum_{\mathbf{x}_r} w^2(\mathbf{x}_r,\mathbf{x}_s)G(\mathbf{x},\mathbf{x}_r,\omega)G'(\mathbf{y},\mathbf{x}_r,\omega)\right), \end{aligned} \tag{7}$$

where $\mathbf{y}$ is a neighbor point around the image point $\mathbf{x}$ in the subsurface.

Valenciano (2008) demonstrates that the Hessian can be directly computed using the above formula; however it requires storing a large number of Green's functions, which is inconvenient for dealing with large 3-D data set. Tang (2008) shows that with some minor alteration of Equation 7, an approximate Hessian can be efficiently computed using the phase-encoding method. However, Tang (2008) focuses more on the algorithm development, and the physics behind the Hessian by phase-encoding has not been carefully discussed. In this companion paper, I complete the discussion of the actual physics behind using phase-encoding methods, such as plane-wave phase encoding and random phase encoding, to obtain the Hessian. In the subsequent sections, I start with the modeling equation in the encoded source, encoded receiver and simultaneously encoded source and receiver domains. I show that the corresponding imaging Hessian in the generalized source and receiver domain is the same as those phase-encoded Hessians discussed in Tang (2008).

ENCODED SOURCES

Let us define the encoding transform along the $\mathbf{x}_s$ coordinate of the surface data as follows:

$$d(\mathbf{x}_r,\mathbf{p}_s,\omega) = \sum_{\mathbf{x}_s} w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega)\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega), \tag{8}$$

where $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)$ is the source phase-encoding function. Equation 8 integrates along the horizontal dashed lines shown in Figure 1 for each receiver location $\mathbf{x}_r$ and transforms the surface data from $(\mathbf{x}_s,\mathbf{x}_r)$ domain into the $(\mathbf{p}_s,\mathbf{x}_r)$ domain. For the plane-wave phase encoding, the encoding function is:

$$\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega) = e^{i\omega\mathbf{p}_s\mathbf{x}_s}, \tag{9}$$

where $\mathbf{p}_s$ is defined to be the ray parameter of the source plane waves. For random phase encoding, the encoding function is

$$\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega) = e^{i\gamma(\mathbf{x}_s,\mathbf{p}_s,\omega)}, \tag{10}$$

where $\gamma(\mathbf{x}_s, \mathbf{p}_s, \omega)$ is a random sequence in $\mathbf{x}_s$ and ω, and $\mathbf{p}_s$ defines the index of different realizations of the random sequence.

Substituting Equation 1 into 8, rearranging the order of summation, we get the forward modeling equation in the encoded source domain:

$$d(\mathbf{x}_r, \mathbf{p}_s, \omega) = \sum_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_r, \omega) G(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r) m(\mathbf{x}), \tag{11}$$

where the encoded source Green's function $G(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r)$ is defined as follows:

$$G(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r) = \sum_{\mathbf{x}_s} w(\mathbf{x}_r, \mathbf{x}_s) G(\mathbf{x}, \mathbf{x}_s, \omega) \alpha(\mathbf{x}_s, \mathbf{p}_s, \omega). \tag{12}$$

Note that $G(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r)$ depends on $\mathbf{x}_r$ because of the acquisition mask $w(\mathbf{x}_r, \mathbf{x}_s)$ inside the summation. It should only integrate the grey segment for each horizontal dashed line shown in Figure 1.

As derived in Appendix A, the objective function in the encoded source domain can be written as follows:

$$J(m(\mathbf{x})) = \frac{1}{2} \sum_{\omega} |c|^2 \sum_{\mathbf{x}_r} \sum_{\mathbf{p}_s} |d(\mathbf{x}_r, \mathbf{p}_s, \omega) - d_{\text{obs}}(\mathbf{x}_r, \mathbf{p}_s, \omega)|^2, \tag{13}$$

where $c = \omega$ for plane-wave phase encoding, and $c = 1$ for random phase encoding.

The gradient of the objective function in Equation 13 gives the following migration formula in the encoded source domain:

$$\nabla J(\mathbf{x}) = \Re \left(\sum_{\omega} |c|^2 \sum_{\mathbf{p}_s} G'(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r) \sum_{\mathbf{x}_r} G'(\mathbf{x}, \mathbf{x}_r, \omega) r(\mathbf{x}_r, \mathbf{p}_s, \omega) \right), \tag{14}$$

where the residual $r(\mathbf{x}_r, \mathbf{p}_s, \omega)$ is defined as follows:

$$r(\mathbf{x}_r, \mathbf{p}_s, \omega) = d(\mathbf{x}_r, \mathbf{p}_s, \omega) - d_{\text{obs}}(\mathbf{x}_r, \mathbf{p}_s, \omega). \tag{15}$$

It is easy to see that Equation 14 defines the phase-encoding migration (Liu et al., 2006; Romero et al., 2000). By taking the second-order derivatives of the objective function defined in Equation 13 with respect to the model paramters, we obtain the Hessian in the encoded source domain:

$$\begin{aligned} H(\mathbf{x}, \mathbf{y}) = \Re \Bigg(& \sum_{\omega} |c|^2 \sum_{\mathbf{x}_r} G(\mathbf{x}, \mathbf{x}_r, \omega) G'(\mathbf{y}, \mathbf{x}_r, \omega) \times \\ & \sum_{\mathbf{p}_s} G(\mathbf{x}, \mathbf{p}_s, \omega; \mathbf{x}_r) G'(\mathbf{y}, \mathbf{p}_s, \omega; \mathbf{x}_r) \Bigg). \end{aligned} \tag{16}$$

ENCODED RECEIVERS

Let us define the encoding transform along the $\mathbf{x}_r$ coordinate of the surface data as follows:

$$d(\mathbf{p}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}_r} w(\mathbf{x}_r, \mathbf{x}_s) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \beta(\mathbf{x}_r, \mathbf{p}_r, \omega), \tag{17}$$

where $\beta(\mathbf{x}_r, \mathbf{p}_r, \omega)$ is the receiver phase-encoding function. Equation 17 integrates along the vertical dashed lines shown in Figure 1 for each source location $\mathbf{x}_s$ and transforms the surface data from $(\mathbf{x}_s, \mathbf{x}_r)$ domain into the $(\mathbf{x}_s, \mathbf{p}_r)$ domain. The receiver phase-encoding function is defined similar to the source phase-encoding function discussed in the previous section. For receiver plane-wave phase encoding,

$$\beta(\mathbf{x}_r, \mathbf{p}_r, \omega) = e^{i\omega \mathbf{p}_r \mathbf{x}_r}, \tag{18}$$

where $\mathbf{p}_r$ is the ray parameter for the receiver plane-waves. For receiver random phase encoding,

$$\beta(\mathbf{x}_r, \mathbf{p}_r, \omega) = e^{i\gamma(\mathbf{x}_r, \mathbf{p}_r, \omega)}, \tag{19}$$

where $\gamma(\mathbf{x}_r, \mathbf{p}_r, \omega)$ is the $\mathbf{p}_r$th random realization. Substituting Equation 1 into 17, we get the forward modeling equation in the receiver plane-wave domain:

$$d(\mathbf{p}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_s, \omega) G(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s) m(\mathbf{x}), \tag{20}$$

where $G(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s)$ is the encoded receiver Green's function defined as follows:

$$G(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s) = \sum_{\mathbf{x}_r} w(\mathbf{x}_r, \mathbf{x}_s) G(\mathbf{x}, \mathbf{x}_r, \omega) \beta(\mathbf{x}_r, \mathbf{p}_r, \omega). \tag{21}$$

Also note that $G(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s)$ depends on $\mathbf{x}_s$ because of the acquisition mask matrix inside the summation. It should only integrate the grey segment for each vertical dashed line shown in Figure 1.

We minimize the following objective function in the encoded receiver domain (see Appendix B for derivation):

$$J(m(\mathbf{x})) = \sum_{\omega} |c|^2 \sum_{\mathbf{p}_r} \sum_{\mathbf{x}_s} |d(\mathbf{p}_r, \mathbf{x}_s, \omega) - d_{\text{obs}}(\mathbf{p}_r, \mathbf{x}_s, \omega)|^2, \tag{22}$$

The gradient of the objective function in Equation 22 gives the following migration formula in the encoded receiver domain:

$$\nabla J(\mathbf{x}) = \Re \left(\sum_{\omega} |c|^2 \sum_{\mathbf{x}_s} G'(\mathbf{x}, \mathbf{x}_s, \omega) \sum_{\mathbf{p}_r} G'(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s) r(\mathbf{p}_r, \mathbf{x}_s, \omega) \right), \tag{23}$$

where the residual $r(\mathbf{p}_r, \mathbf{x}_s, \omega)$ is defined as follows:

$$r(\mathbf{p}_r, \mathbf{x}_s, \omega) = d(\mathbf{p}_r, \mathbf{x}_s, \omega) - d_{\text{obs}}(\mathbf{p}_r, \mathbf{x}_s, \omega). \tag{24}$$

The Hessian in the encoded receiver domain is then obtained by taking the second derivative of the objective function:

$$\begin{aligned} H(\mathbf{x}, \mathbf{y}) \quad = \quad & \Re \left(\sum_{\omega} |c|^2 \sum_{\mathbf{x}_s} G(\mathbf{x}, \mathbf{x}_s, \omega) G'(\mathbf{y}, \mathbf{x}_s, \omega) \times \right. \\ & \left. \sum_{\mathbf{p}_r} G(\mathbf{x}, \mathbf{p}_r, \omega; \mathbf{x}_s) G'(\mathbf{y}, \mathbf{p}_r, \omega; \mathbf{x}_s) \right). \end{aligned} \tag{25}$$

We can rewrite Equation 25 as follows

$$H(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{p}_r} H(\mathbf{x},\mathbf{y},\mathbf{p}_r), \tag{26}$$

where

$$\begin{aligned} H(\mathbf{x},\mathbf{y},\mathbf{p}_r) &= \Re\left(\sum_{\omega}|c|^2\sum_{\mathbf{x}_s} G(\mathbf{x},\mathbf{x}_s,\omega)G'(\mathbf{y},\mathbf{x}_s,\omega)\ \times \right. \\ &\quad \left. G(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)G'(\mathbf{y},\mathbf{p}_r,\omega;\mathbf{x}_s)\ \right) \\ &= \Re\left(\sum_{\omega}|c|^2\sum_{\mathbf{x}_s} G(\mathbf{x},\mathbf{x}_s,\omega)G'(\mathbf{y},\mathbf{x}_s,\omega)\ \times \right. \\ &\quad \left(\sum_{\mathbf{x}_r} w(\mathbf{x}_r,\mathbf{x}_s)G(\mathbf{x},\mathbf{x}_r,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\right)\times \\ &\quad \left.\left(\sum_{\mathbf{x}'_r} w(\mathbf{x}_r,\mathbf{x}_s)G(\mathbf{y},\mathbf{x}'_r,\omega)\beta(\mathbf{x}'_r,\mathbf{p}_r,\omega)\right)'\right). \end{aligned} \tag{27}$$

Equation 27 is equivalent to Equations 9 and B-1 in Tang (2008), which are called the receiver-side encoded Hessian. As I show here, the receiver-side encoded Hessian is the same as the Hessian in the encoded receiver domain; both of them are derived from the same forward modeling equation defined in Equation 20.

ENCODED SOURCES AND RECEIVERS

We can simultaneously encode the sources and the receivers as follows:

$$d(\mathbf{p}_r,\mathbf{p}_s,\omega) = \sum_{\mathbf{x}_r}\sum_{\mathbf{x}_s} w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega)\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega), \tag{28}$$

where $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)$ and $\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)$ are defined by Equations 9 and 18, respectively, for plane-wave phase encoding, and by Equations 10 and 19, respectively, for random phase encoding.

Substituting Equations 1 into 28, notice that $w^2(\mathbf{x}_r,\mathbf{x}_s) = w(\mathbf{x}_r,\mathbf{x}_s)$. With the definition of $G(\mathbf{x},\mathbf{p}_s,\omega;\mathbf{x}_r)$ and $G(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)$ by Equations 12 and 21, we obtain the forward modeling equation in the encoded source and receiver domain:

$$d(\mathbf{p}_r,\mathbf{p}_s,\omega) = \sum_{\mathbf{x}} G(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)G(\mathbf{x},\mathbf{p}_s,\omega;\mathbf{x}_r)m(\mathbf{x}). \tag{29}$$

Now we minimize the following objective function in the encoded source and receiver domain (see Appendix C for derivation):

$$J(m(\mathbf{x})) = \sum_{\omega}|c|^4\sum_{\mathbf{p}_r}\sum_{\mathbf{p}_s}|d(\mathbf{p}_r,\mathbf{p}_s,\omega) - d_{\text{obs}}(\mathbf{p}_r,\mathbf{p}_s,\omega)|^2. \tag{30}$$

The gradient of the objective function in Equation 30 gives the following migration formula in the encoded source and receiver domain:

$$\nabla J(\mathbf{x}) = \Re\left(\sum_{\omega}|c|^4\sum_{\mathbf{p}_r}G'(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)\sum_{\mathbf{p}_s}G'(\mathbf{x},\mathbf{p}_s,\omega;\mathbf{x}_r)r(\mathbf{p}_r,\mathbf{p}_s,\omega)\right), \tag{31}$$

where $r(\mathbf{p}_r,\mathbf{p}_s,\omega)$ is the residual in the encoded source and receiver domain:

$$r(\mathbf{p}_r,\mathbf{p}_s,\omega) = d(\mathbf{p}_r,\mathbf{p}_s,\omega) - d_{\text{obs}}(\mathbf{p}_r,\mathbf{p}_s,\omega). \tag{32}$$

The Hessian is obtained as follows:

$$\begin{aligned} H(\mathbf{x},\mathbf{y}) &= \Re\left(\sum_{\omega}|c|^4\sum_{\mathbf{p}_r}G(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)G'(\mathbf{y},\mathbf{p}_r,\omega;\mathbf{x}_s)\times\right. \\ &\left.\sum_{\mathbf{p}_s}G(\mathbf{x},\mathbf{p}_s,\omega;\mathbf{x}_r)G'(\mathbf{y},\mathbf{p}_s,\omega;\mathbf{x}_r)\right). \end{aligned} \tag{33}$$

We can also rewrite Equation 33 as follows:

$$H(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{p}_s}\sum_{\mathbf{p}_r}H(\mathbf{x},\mathbf{y},\mathbf{p}_s,\mathbf{p}_r), \tag{34}$$

where

$$\begin{aligned} &H(\mathbf{x},\mathbf{y},\mathbf{p}_s,\mathbf{p}_r) \\ &= \Re\left(\sum_{\omega}|c|^4G(\mathbf{x},\mathbf{p}_s,\omega;\mathbf{x}_r)G'(\mathbf{y},\mathbf{p}_s,\omega;\mathbf{x}_r)G(\mathbf{x},\mathbf{p}_r,\omega;\mathbf{x}_s)G'(\mathbf{y},\mathbf{p}_r,\omega;\mathbf{x}_s)\right) \\ &= \Re\left(\sum_{\omega}|c|^4\times\right. \\ &\left(\sum_{\mathbf{x}_s}w(\mathbf{x}_r,\mathbf{x}_s)G(\mathbf{x},\mathbf{x}_s,\omega)\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\right)\left(\sum_{\mathbf{x}'_s}w(\mathbf{x}_r,\mathbf{x}'_s)G(\mathbf{y},\mathbf{x}'_s,\omega)\alpha(\mathbf{x}'_s,\mathbf{p}_s,\omega)\right)'\times \\ &\left.\left(\sum_{\mathbf{x}_r}w(\mathbf{x}_r,\mathbf{x}_s)G(\mathbf{x},\mathbf{x}_r,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\right)\left(\sum_{\mathbf{x}'_r}w(\mathbf{x}'_r,\mathbf{x}_s)G(\mathbf{y},\mathbf{x}'_r,\omega)\beta(\mathbf{x}'_r,\mathbf{p}_r,\omega)\right)'\right). \end{aligned} \tag{35}$$

Equation 35 is equivalent to Equations 17 and C-1 in Tang (2008), which are called the simultaneously encoded Hessian. However, Equation 35 is more general, because it is not limited to OBC or land acquisition geometry. Up to this point, we have proved that both the simultaneously encoded Hessian and the Hessian in the encoded source and receiver domain are derived from the same forward modeling equation defined in 29.

A MIXED PHASE-ENCODING SCHEME

Tang (2008) compares the computational cost for the Hessian obtained using different phase-encoding and shows that the most efficient way to compute the Hessian is to use simultaneous random phase encoding. The cost is just two downward continuations (using one-way

wave equation to model the Green's functions) of the encoded wavefields plus the cross-correlation (for a single realization of the random phases). However, the random phase encoding may not be very effective in attenuating the cross-talk when many Green's functions are simultaneously encoded (Romero et al., 2000; Tang, 2008); a lot of random noise may appear in the final result. In fact, for the simultaneous phase encoding or the Hessian in the encoded source and receiver domain defined by Equation 33, the phase-encoding functions for sources ($\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)$)and receivers ($\beta(\mathbf{x}_r, \mathbf{p}_r, \omega)$) need not be the same. For example, we can use plane-wave phase encoding function to encode the sources but use random phase-encoding function to encode the receivers, or vice visa. Because plane-wave phase encoding functions are very effective in attenuating the cross-talk (Liu et al., 2006; Tang, 2008), while random phase encoding is efficient, by combining those two phase-encoding functions, we are able to balance cost and accuracy.

I apply this idea to a simple constant-velocity model. The acquisition geometry is assumed to be OBC geometry, where all shots share the same receiver array. There are 201 shots from -2000 m to 2000 m with a 10 m sampling; for each shot, there are 201 receivers spanning from -2000 m to 2000 m. Figure 2 shows the diagonal of the Hessian obtained using different methods. Figure 2(a) shows the exact diagonal of the Hessian computed in the original shot-profile domain with Equation 7, which requires pre-computing and saving the Green's functions and is efficient for practical applications. However, since there is no cross-talk in the original shot-profile domain, I use the result as a benchmark to compare the accuracy of other methods. Figure 2(b) is obtained using the most efficient simultaneous random phase-encoding method (only one realization of the random-phase functions has been used), by using Equation 33, with both $\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)$ and $\beta(\mathbf{x}_r, \mathbf{p}_r, \omega)$ being random phase functions. As expected, the result is full of random noise, useful illumination information is greatly distorted, and the result is very far from the exact Hessian. Figure 2(c) shows the result obtained using the mixed phase-encoding scheme, i.e., by using Equation 33, with the weighting function $\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)$ being the plane-wave phase-encoding function and $\beta(\mathbf{x}_r, \mathbf{p}_r, \omega)$ being the random phase-encoding function. A total of 61 plane-wave-encoded source-side Green's functions have been used to generate the result. The cost is the same as a plane-wave source migration with 61 source plane waves. The result looks very similar to the exact Hessian, and the random noise shown in Figure 2(b) has been greatly reduced. For comparison, Figure 2(d) shows the Hessian computed in the encoded receiver domain, i.e., by using Equation 25, where the weighting function $\beta(\mathbf{x}_r, \mathbf{p}_r, \omega)$ is chosen to be a random phase function. The result is also very accurate. However, its cost is the same as a shot-profile migration with 201 shot gathers (Tang, 2008).

Figure 3 and 4 show the Hessian with off-diagonals (with size 21×21) obtained using different methods. Figure 3 illustrates the result at image point $(x = 0, z = 800)$, while Figure 4 illustrates the result at image point $(x = 420, z = 800)$. These results also demonstrate that although simultaneous random phase encoding is efficient, the Hessian operator obtained by this method (Figure 3(b) and Figure 4(b)) suffers a lot from unwanted crosstalk. Encoding only the receiver-side Green's function with random phase functions gives accurate results (Figure 3(d) and Figure 4(d)); however, the cost is similar to a shot-profile migration. In situations where the number of shot gathers is big, this encoding scheme may not be a good choice. In contrast, the mixed phase-encoding scheme gives us very accurate results (Figure 3(c) and Figure 4(c)) but with less cost than the receiver-side encoded Hessian.

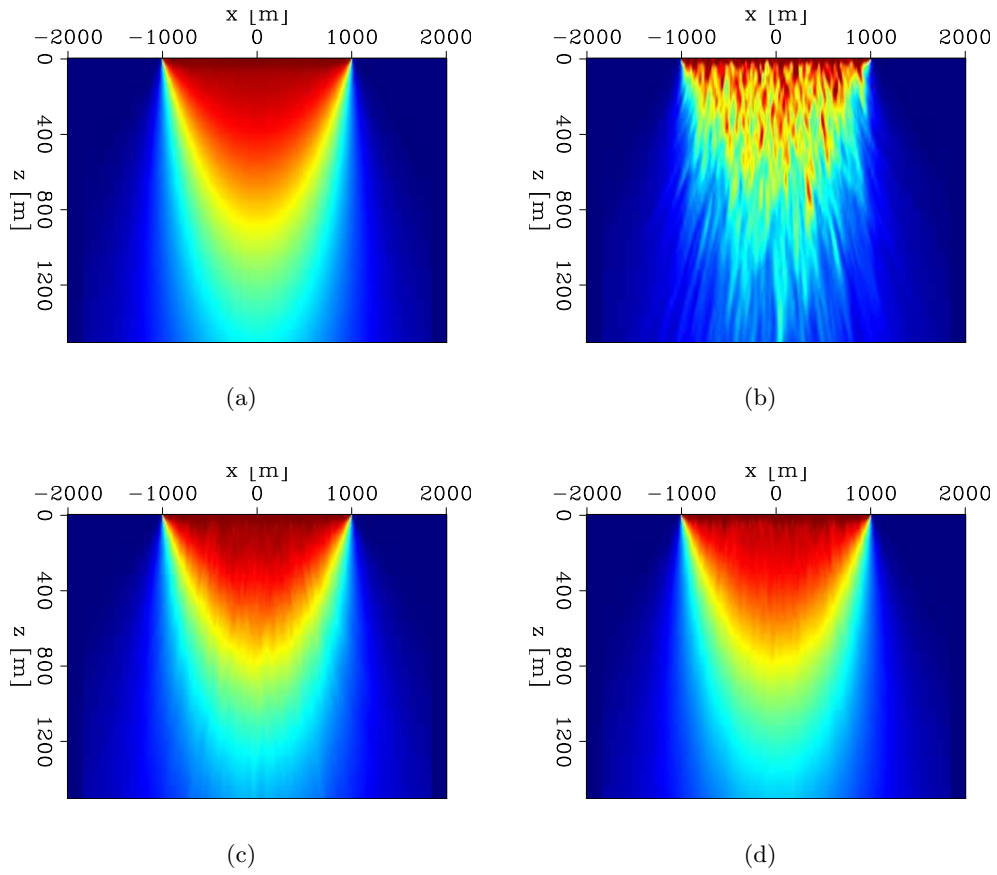
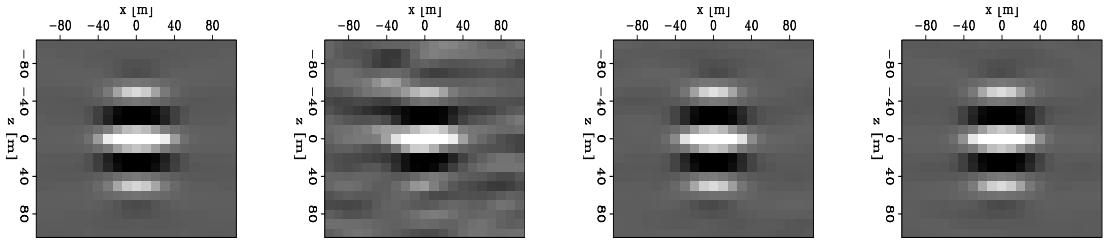


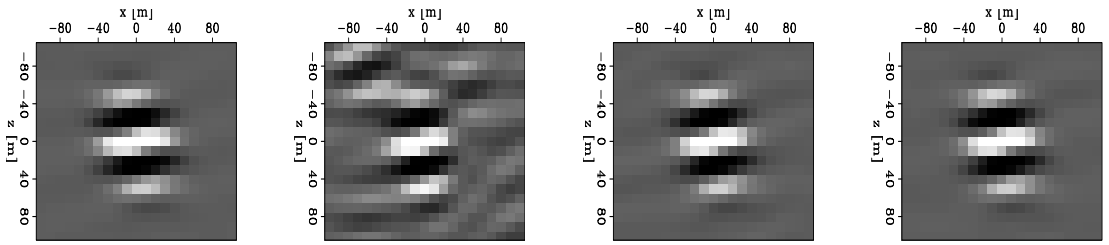
Figure 2: The diagonal of the Hessian obtained with different methods. Panel (a) is the exact diagonal of Hessian computed in the original shot-profile domain. Panel (b) is the result computed in the encoded source and receiver domain, where both the source and receiver Green's functions are randomly encoded. Panel (c) is the result also computed in the encoded source and receiver domain, where the source-side Green's functions are encoded with the plane-wave phase encoding function, while the receiver-side Green's functions are encoded with the random phase functions. Panel (d) is the result computed in the encoded receiver domain, where only the receiver-side Green's functions are randomly encoded. [**CR**]
yaxun2/. hess-exact,hess-simul-random,hess-simul-mixed,hess-random



(a) (b) (c) (d)

Figure 3: The Hessian operator for an image point at $x = 0$ m and $z = 800$ m, with the same acquisition geometry as in Figure 2. The size of the Hessian operator is 21×21. Panel (a) shows the exact Hessian computed in the original shot-profile domain, which is artifact-free; Panel (b) is the result of simultaneous random phase encoding. Note the strong artifact which distorts the useful information; Panel (c) is the result of the mixed phase encoding; the result is very similar to the exact Hessian in (a); Panel (d) is the result computed in the encoded receiver domain, where only the receiver-side Green's functions are randomly encoded. [**CR**]

yaxun2/. hess-exact-offd1,hess-simul-random-offd1,hess-simul-mixed-offd1,hess-random-offd1



(a) (b) (c) (d)

Figure 4: The Hessian operator for an image point at $x = 420$ m and $z = 800$ m, with the same acquisition geometry as in Figure 2. The size of the Hessian operator is 21×21. Panel (a) shows the exact Hessian computed in the original shot-profile domain, which is be artifact-free; Panel (b) is the result of simultaneous random phase encoding. Note the strong artifact which distorts the useful information; Panel (c) is the result of the mixed phase encoding; the result is very similar to the exact Hessian in (a); Panel (d) is the result computed in the encoded receiver domain, where only the receiver-side Green's functions are randomly encoded. [**CR**]

yaxun2/. hess-exact-offd2,hess-simul-random-offd2,hess-simul-mixed-offd2,hess-random-offd2

CONCLUSION

I extend the theory of Born modeling, migration and inversion from the conventional shot-profile domain to the generalized source and receiver domain, and to the plane-wave phase-encoding domain and the random phase-encoding domain in particular. One important advantage of these new domains is that computing the target-oriented Hessian no longer requires storing the Green's functions, which is a major obstacle preventing the Hessian computation in the original shot-profile domain. I also prove that the Hessians obtained in the generalized source and receiver domain are equivalent to those phase-encoded Hessians discussed in Tang (2008), and they both can be derived from the same modeling equations. To balance the accuracy and cost of computing the Hessian, I introduce a new phase encoding domain, i.e., the mixed phase-encoding scheme, which combines plane-wave phase encoding and random phase encoding. My preliminary tests shows that this new phase-encoding scheme achieves high accuracy with relatively low cost.

REFERENCES

Bleistein, N., 2007, Migration/inversion for incident waves synthesized from common-shot data gathers: Geophysics, **72**, W17–W31.

Duquet, B. and P. Lailly, 2006, Efficient 3D wave-equation migration using virtual planar sources: Geophysics, **71**, S185–S197.

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.

Plessix, R.-E. and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophys. J. Int., **157**, 975–987.

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.

Stoffa, P. L., M. K. Sen, R. K. Seifoullaev, R. C. Pestana, and J. T. Fokkema, 2006, Plane-wave depth migration: Geophysics, **71**, S261–S272.

Tang, Y., 2008, Wave-equation Hessian by phase encoding: **SEP-134**, 1–24.

Valenciano, A., 2008, Imaging by Wave-equation Inversion: PhD thesis, Stanford University.

Whitmore, N. D., 1995, An Imaging Hierarchy for Common Angle Plane Wave Seismogram: PhD thesis, University of Tulsa.

Zhang, Y., J. Sun, C. Notfors, S. Grey, L. Chemis, and J. Young, 2005, Delayed-shot 3D depth migration: Geophysics, **70**, E21–E28.

APPENDIX A

This appendix derives the objective function in the encoded source domain. We start with the objective function in the source and receiver domain as follows:

$$J(m(\mathbf{x})) = \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|w(\mathbf{x}_r,\mathbf{x}_s)(d(\mathbf{x}_r,\mathbf{x}_s,\omega)-d_{\text{obs}}(\mathbf{x}_r,\mathbf{x}_s,\omega))|^2. \tag{A-1}$$

From Equation 8, we can also get the inverse phase-encoding transform. For source plane-wave phase encoding, since the forward operator is unitary, the inverse transform can be written as follows:

$$w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega)=|\omega|^2\sum_{\mathbf{p}_s}d(\mathbf{x}_r,\mathbf{p}_s,\omega)e^{-i\omega\mathbf{p}_s\mathbf{x}_s}. \tag{A-2}$$

For random phase encoding, a similar result can also be obtained, because different realizations of random sequences should be approximately orthogonal, provided that those random sequences are "sufficiently" random; thus we have

$$w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega)=\sum_{\mathbf{p}_s}d(\mathbf{x}_r,\mathbf{p}_s,\omega)e^{-i\gamma(\mathbf{x}_s,\mathbf{p}_s,\omega)}. \tag{A-3}$$

Therefore, we can use a more general form to express the inverse phase-encoding transform:

$$w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega)=|c|^2\sum_{\mathbf{p}_s}d(\mathbf{x}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega), \tag{A-4}$$

where for plane-wave phase encoding, $c=\omega$ and $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)=e^{i\omega\mathbf{p}_s\mathbf{x}_s}$; for random phase encoding, $c=1$ and $\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)=e^{i\gamma(\mathbf{x}_s,\mathbf{p}_s,\omega)}$.

Substituting Equation A-4 into A-1 yields:

$$\begin{aligned}
J(m(\mathbf{x})) &= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^4\left|\sum_{\mathbf{p}_s}r(\mathbf{x}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega)\right|^2\\
&= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^4\left(\sum_{\mathbf{p}_s}r(\mathbf{x}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega)\right)'\times\\
&\quad\left(\sum_{\mathbf{p}'_s}r(\mathbf{x}_r,\mathbf{p}'_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)\right)\\
&= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_r}|c|^4\sum_{\mathbf{p}_s}\sum_{\mathbf{p}'_s}r'(\mathbf{x}_r,\mathbf{p}_s,\omega)r(\mathbf{x}_r,\mathbf{p}'_s,\omega)\times\\
&\quad\sum_{\mathbf{x}_s}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega),
\end{aligned} \tag{A-5}$$

where $r(\mathbf{x}_r,\mathbf{p}_s,\omega)$ is defined to be the residual in the encoded source domain:

$$r(\mathbf{x}_r,\mathbf{p}_s,\omega)=d(\mathbf{x}_r,\mathbf{p}_s,\omega)-d_{\text{obs}}(\mathbf{x}_r,\mathbf{p}_s,\omega). \tag{A-6}$$

For plane-wave phase encoding, if the $\mathbf{x}_s$ is sampled densely enough,

$$\sum_{\mathbf{x}_s}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)=\sum_{\mathbf{x}_s}e^{i\omega(\mathbf{p}'_s-\mathbf{p}_s)\mathbf{x}_s}\approx\frac{1}{|\omega|^2}\delta(\mathbf{p}'_s-\mathbf{p}_s). \tag{A-7}$$

For random phase encoding, the following property also holds as long as the random sequences are "sufficiently" random:

$$\sum_{\mathbf{x}_s}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)=\sum_{\mathbf{x}_s}e^{i(\gamma(\mathbf{x}_s,\mathbf{p}'_s,\omega)-\gamma(\mathbf{x}_s,\mathbf{p}_s,\omega))}\approx\delta(\mathbf{p}'_s-\mathbf{p}_s). \tag{A-8}$$

Substituting Equation A-7 or A-8 into A-5, we get the data-misfit function in the encoded source domain:

$$J(m(\mathbf{x})) \approx \frac{1}{2}\sum_{\omega}|c|^2\sum_{\mathbf{x}_r}\sum_{\mathbf{p}_s}|d(\mathbf{x}_r,\mathbf{p}_s,\omega)-d_{\text{obs}}(\mathbf{x}_r,\mathbf{p}_s,\omega)|^2. \tag{A-9}$$

APPENDIX B

This appendix derives the objective function in the encoded receiver domain. We start with the objective function in the source and receiver domain as follows:

$$J(m(\mathbf{x})) = \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|w(\mathbf{x}_r,\mathbf{x}_s)(d(\mathbf{x}_r,\mathbf{x}_s,\omega)-d_{\text{obs}}(\mathbf{x}_r,\mathbf{x}_s,\omega))|^2. \tag{B-1}$$

Similar to the discussion in Appendix A, the general form of the inverse transform of receiver phase encoding can be written as follows:

$$w(\mathbf{x}_r,\mathbf{x}_s)d(\mathbf{x}_r,\mathbf{x}_s,\omega) = |c|^2\sum_{\mathbf{p}_r}d(\mathbf{p}_r,\mathbf{x}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega), \tag{B-2}$$

where for plane-wave phase encoding, $c=\omega$ and $\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)=e^{i\omega\mathbf{p}_r\mathbf{x}_r}$; for random phase encoding, $c=1$ and $\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)=e^{i\gamma(\mathbf{x}_r,\mathbf{p}_r,\omega)}$. Substituting Equation B-2 into B-1 yields:

$$\begin{aligned}
J(m(\mathbf{x})) &= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^4\left|\sum_{\mathbf{p}_r}r(\mathbf{p}_r,\mathbf{x}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega)\right|^2 \\
&= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^4\left(\sum_{\mathbf{p}_r}r(\mathbf{p}_r,\mathbf{x}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega)\right)'\times \\
&\quad \left(\sum_{\mathbf{p}'_r}r(\mathbf{p}'_r,\mathbf{x}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega)\right) \\
&= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}|c|^4\sum_{\mathbf{p}_r}\sum_{\mathbf{p}'_r}r'(\mathbf{p}_r,\mathbf{x}_s,\omega)r(\mathbf{p}'_r,\mathbf{x}_s,\omega)\times \\
&\quad \sum_{\mathbf{x}_r}\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega),
\end{aligned} \tag{B-3}$$

where $r(\mathbf{p}_r,\mathbf{x}_s,\omega)$ is defined to be the residual in the encoded receiver domain:

$$r(\mathbf{p}_r,\mathbf{x}_s,\omega) = d(\mathbf{p}_r,\mathbf{x}_s,\omega)-d_{\text{obs}}(\mathbf{p}_r,\mathbf{x}_s,\omega). \tag{B-4}$$

Similar to the discussion in Appendix A, the inner-most summation in Equation B-3 is approximately a Dirac delta function under certain conditions. Therefore, the data-misfit function in the encoded receiver domain reads as follows:

$$J(m(\mathbf{x})) \approx \frac{1}{2}\sum_{\omega}|c|^2\sum_{\mathbf{p}_r}\sum_{\mathbf{x}_s}|d(\mathbf{p}_r,\mathbf{x}_s,\omega)-d_{\text{obs}}(\mathbf{p}_r,\mathbf{x}_s,\omega)|^2. \tag{B-5}$$

APPENDIX C

This appendix derives the objective function in the simultaneously encoded source and receiver domain. We start with the objective function in the source and receiver domain as follows:

$$J(m(\mathbf{x})) = \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|w(\mathbf{x}_r,\mathbf{x}_s)(d(\mathbf{x}_r,\mathbf{x}_s,\omega)-d_{\rm obs}(\mathbf{x}_r,\mathbf{x}_s,\omega))|^2. \tag{C-1}$$

If we follow a discussion similar to those in Appendices A and B, we obtain the general expression of the inverse transform of the simultaneous encoding:

$$w(\mathbf{x}_r,\mathbf{x}_s,\omega)d(\mathbf{x}_r,\mathbf{x}_s,\omega) = |c|^4\sum_{\mathbf{p}_s}\sum_{\mathbf{p}_r} d(\mathbf{p}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega). \tag{C-2}$$

Substituting Equation C-2 into C-1 yields:

$$\begin{aligned}
J(m(\mathbf{x})) &= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^8\left|\sum_{\mathbf{p}_s}\sum_{\mathbf{p}_r} r(\mathbf{p}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega)\right|^2 \\
&= \frac{1}{2}\sum_{\omega}\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}|c|^8\left(\sum_{\mathbf{p}_s}\sum_{\mathbf{p}_r} r(\mathbf{p}_r,\mathbf{p}_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}_r,\omega)\right)'\times \\
&\quad \left(\sum_{\mathbf{p}'_s}\sum_{\mathbf{p}'_r} r(\mathbf{p}'_r,\mathbf{p}'_s,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega)\right) \\
&= \frac{1}{2}\sum_{\omega}|c|^8\sum_{\mathbf{p}_s}\sum_{\mathbf{p}'_s}\sum_{\mathbf{p}_r}\sum_{\mathbf{p}'_r} r'(\mathbf{p}_r,\mathbf{p}_s,\omega)r(\mathbf{p}'_r,\mathbf{p}'_s,\omega)\times \\
&\quad \sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega),
\end{aligned} \tag{C-3}$$

where $r(\mathbf{p}_r,\mathbf{p}_s,\omega)$ is defined to be the residual in the encoded source and receiver domain:

$$r(\mathbf{p}_r,\mathbf{p}_s,\omega) = d(\mathbf{p}_r,\mathbf{p}_s,\omega) - d_{\rm obs}(\mathbf{p}_r,\mathbf{p}_s,\omega). \tag{C-4}$$

For plane-wave phase encoding, with sampling dense enough in $\mathbf{x}_s$ and $\mathbf{x}_r$, the inner-most summations become Dirac delta functions:

$$\begin{aligned}
&\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega) \\
&\approx \frac{1}{|\omega|^4}\delta(\mathbf{p}'_r-\mathbf{p}_r)\delta(\mathbf{p}'_s-\mathbf{p}_s).
\end{aligned} \tag{C-5}$$

For random phase encoding, we can also approximately have

$$\begin{aligned}
&\sum_{\mathbf{x}_s}\sum_{\mathbf{x}_r}\alpha(\mathbf{x}_s,\mathbf{p}_s,\omega)\beta(\mathbf{x}_r,\mathbf{p}_r,\omega)\alpha'(\mathbf{x}_s,\mathbf{p}'_s,\omega)\beta'(\mathbf{x}_r,\mathbf{p}'_r,\omega) \\
&\approx \delta(\mathbf{p}'_r-\mathbf{p}_r)\delta(\mathbf{p}'_s-\mathbf{p}_s).
\end{aligned} \tag{C-6}$$

Therefore the data-misfit function in the encoded source and receiver domain is

$$J(m(\mathbf{x})) \approx \frac{1}{2}\sum_{\omega}|c|^4\sum_{\mathbf{p}_s}\sum_{\mathbf{p}_r}|d(\mathbf{p}_r,\mathbf{p}_s,\omega)-d_{\rm obs}(\mathbf{p}_r,\mathbf{p}_s,\omega)|^2. \tag{C-7}$$

Salt body segmentation with dip and frequency attributes

Adam Halpert and Robert G. Clapp

ABSTRACT

Image segmentation can automatically locate salt boundaries on seismic sections, an often time-consuming and tedious task when undertaken manually. However, using a single seismic attribute (usually amplitude) is sometimes insufficient to achieve an accurate segmentation. Since any quantifiable measure may be employed as an attribute for segmentation, exploring other possible attributes is an important step in developing a more robust segmentation algorithm. Dip variability within a seismic section is one attribute with many advantages for segmentation, and experimenting with different methods for calculating dips can yield improved results. Determining the frequency content of a seismic image offers other opportunities for improvement. Specifically, instantaneous frequency shows promise as another attribute for segmentation, while employing a continuous wavelet transform to study envelope amplitude at different frequencies can improve the performance of the amplitude attribute.

INTRODUCTION

Automated image segmentation offers a means of quickly and efficiently delineating salt bodies on seismic images. When adapted for seismic purposes (Lomask, 2007; Lomask et al., 2007), the Normalized Cuts Image Segmentation (NCIS) algorithm (Shi and Malik, 2000) provides a global solution to the salt boundary calculation. This helps overcome some of the weaknesses of other methods, such as local horizon trackers, that can fail when the boundary fades or becomes discontinuous. A global determination of salt boundaries can be especially important for building velocity models in complex areas; in such cases, image segmentation may be employed to automatically build or update velocity models, helping to alleviate a major bottleneck for iterative imaging projects (Halpert and Clapp, 2008).

The NCIS algorithm relies on one or more seismic attributes to segment an image; the most straightforward of these attributes is envelope of the amplitude. Often, this single attribute provides an accurate calculation of the salt boundary. Unfortunately, in some instances a single attribute is insufficient, and the algorithm cannot produce a reasonable result. Here, we demonstrate situations in which a single-attribute segmentation fails. We then explore possibilities for useful attributes other than amplitude, most notably dip and instantaneous frequency. Finally, we discuss an ultimate goal of designing an interpreter-guided multi-attribute segmentation scheme.

ATTRIBUTE-BASED SEGMENTATION

The NCIS algorithm (Shi and Malik, 2000) functions by calculating relationships between individual pixel pairs in an image. For seismic image segmentation, each pixel is compared

to a random selection of neighboring pixels (Lomask, 2007), and each pair is assigned a weight value inversely proportional to the likelihood of a salt boundary existing between them. If these weights are placed in a matrix $\mathbf{W}$, a salt boundary path may be calculated via the eigensystem

$$(\mathbf{D} - \mathbf{W})\mathbf{y} = \lambda \mathbf{D}\mathbf{y} \ , \tag{1}$$

where $\mathbf{D}$ is a diagonal matrix whose elements are the sum of each column of $\mathbf{W}$, and λ and $\mathbf{y}$ are an eigenvalue and eigenvector of the system, respectively. Because the smallest eigenvalue will yield only a constant eigenvector, the eigenvector corresponding to the second smallest eigenvalue of the system is used to segment the image. The eigenvector in question ranges in value from -1 to +1 across the estimated boundary path. The image is segmented either by following a constant eigenvector value across the image - usually the zero-value contour; see Lomask (2007) - or by using outside information to follow different eigenvector values across the image (Halpert and Clapp, 2008).

In the examples cited above, the most common measure or attribute used to determine the likely presence of a salt boundary is amplitude of the envelope. For instance, a large amplitude between two pixels suggests a salt boundary is likely, and will result in a relatively small weight assigned to that pixel pair. Amplitude, however, is not the only attribute that can be used for this purpose. In fact, any quantifiable measure that relates to the presence of a salt boundary may be used as an attribute for segmentation.

Limitations of a single attribute

The sources cited above contain several examples of accurate salt boundary picking using only amplitude to segment the image. In some instances, however, such an approach can lead to inaccurate results. Figure 1 shows one such instance. Since the algorithm attempts to find a minimized path across the image, it often tends toward a straight line in areas of uncertainty. In this case, the algorithm incorrectly cuts across a very steep salt canyon in the Sigsbee synthetic model. Even though the outline of the salt top is faintly visible, amplitude information alone is not enough to correctly guide the segmentation process in this instance. This suggests that further information, in the form of attributes other than amplitude, may be an important part of a more robust segmentation algorithm.

DIP AS AN ATTRIBUTE

Estimating dip in seismic images has long been recognized as a useful interpretation technique, especially for 3D volumes (Bednar, 1997). Dip can be a particularly useful attribute when the goal is delineation of salt bodies. Lomask (2007) notes that salt boundaries and nearby sediments will often display different dips; furthermore, the seemingly random dips that often characterize the interior of salt bodies on field seismic data should contrast considerably with more ordered dipping layers outside the salt. Therefore, an attribute that highlights variability in dip throughout a seismic image may be helpful for locating salt boundaries.

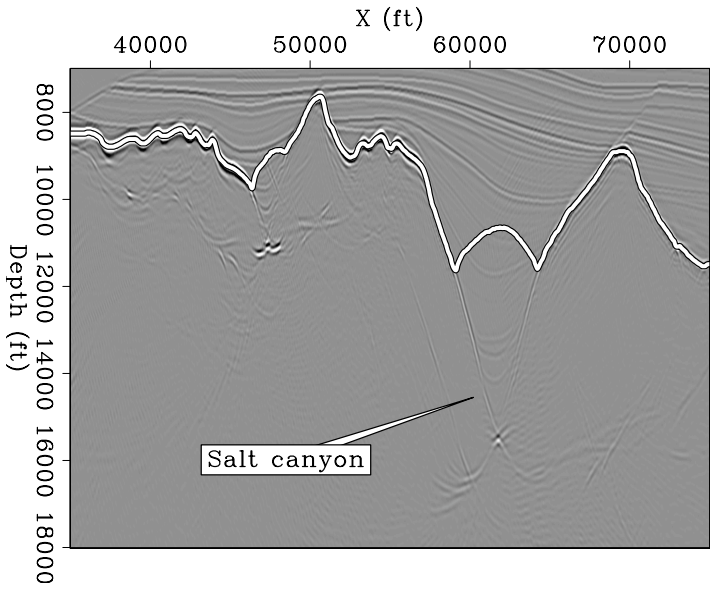


Figure 1: Example of single-attribute segmentation failing to accurately track a salt boundary. The algorithm cuts across the top of the indicated salt canyon. **[CR]** **adam1/. sig-ann**

Dip calculation

Dips on seismic images may be esitmated in several ways. Working from an image processing perspective, van Vliet and Verbeek (1995) describe the process of calculating dips using gradient-square tensors to estimate "orientation" in images. Such methods may easily be adapted to estimate dips of reflectors on seismic images. Claerbout (1992) describes a "plane-wave destructor" mechanism to estimate dips, while Fomel (2002) and Hale (2007) use different plane-wave destructor stencils to implement dip filters. Each of these methods for estimating dips produces a different result, so it is important to determine how these differences can affect the segmentation process.

Figure 2 shows a migrated section of a 2D Gulf of Mexico dataset used for examples throughout this paper; the strong reflector represents the base of a salt body. The first column of Figure 3 displays dip calculations for the three methods metioned above. Panel (a) is the result of applying Claerbout's "puck" method (Claerbout, 1992), which uses a four-point differencing stencil to calculate directional derivatives. In this example, the random character of dips inside the salt body is readily apparent from the chaotic nature of the dip field, although the boundary itself is difficult to discern. To overcome some of the shortcomings of a four-point stencil, the methods used to obtain the results in panels (b) and (c) instead employ a six-point stencil. The stencil used for panel (b) (Fomel, 2002) is

$$\begin{array}{|c|c|}\hline -\frac{(1+\sigma)(2+\sigma)}{12} & \frac{(1-\sigma)(2-\sigma)}{12} \\ \hline -\frac{(2+\sigma)(2-\sigma)}{6} & \frac{(2+\sigma)(2-\sigma)}{6} \\ \hline -\frac{(1-\sigma)(2-\sigma)}{12} & \frac{(1+\sigma)(2+\sigma)}{12} \\ \hline \end{array}, \tag{2}$$

where σ is the slope of the plane wave being "killed." The stencil used for panel (c) (Hale,

2007) was designed to fit on the previous stencil, while allowing for improved handling of very steep dips. It looks like

$$\begin{array}{|c|c|}\hline -2m^2 & 2mp \\ \hline -4mp & 1 \\ \hline -2p^2 & 2mp \\ \hline \end{array}, \tag{3}$$

where $m = \frac{1}{2}\left[\cos(\tan^{-1}\sigma) + \sin(\tan^{-1}\sigma)\right]$ and $p = \frac{1}{2}\left[\cos(\tan^{-1}\sigma) - \sin(\tan^{-1}\sigma)\right]$. Because of the six-point stencil in use, both methods do a better job of recovering the coherent dips along the salt boundary; panel (c) is the best result, as it performs more accurately even at the steepest dips.

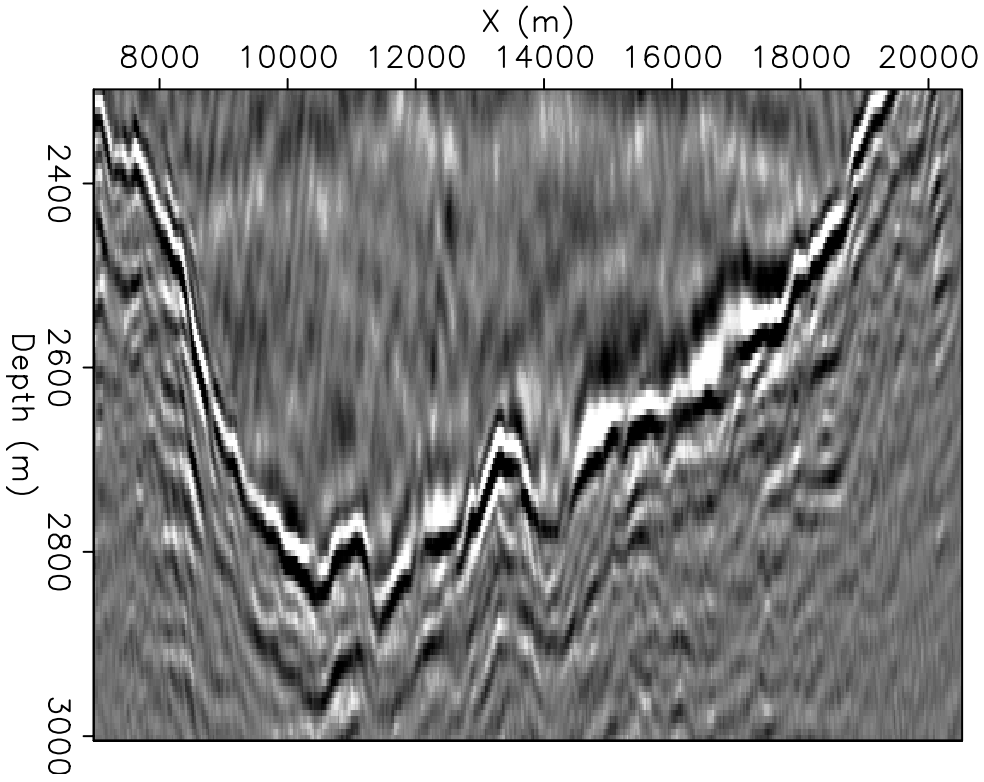


Figure 2: A 2D seismic section used throughout this paper. The strong reflector is the base of a salt body. **[ER]** **adam1/. gulf**

Segmentation using dips

To prepare a dip volume for segmentation, additional steps are necessary. To highlight changes in the dip (an abrupt change is indicative of a salt boundary), a roughener should be applied; here, the helical derivative (Claerbout, 2005) is used to facilitate extension to three dimensions. Finally, calculating the envelope of this volume will produce an image suitable for the segmentation process. The second column of figure 3 shows envelopes for the three dip volumes seen in the left column. Panel (c) shows the clearest salt boundary path, so we expect this image to produce the superior segmentation eigenvector. Figure 4, which

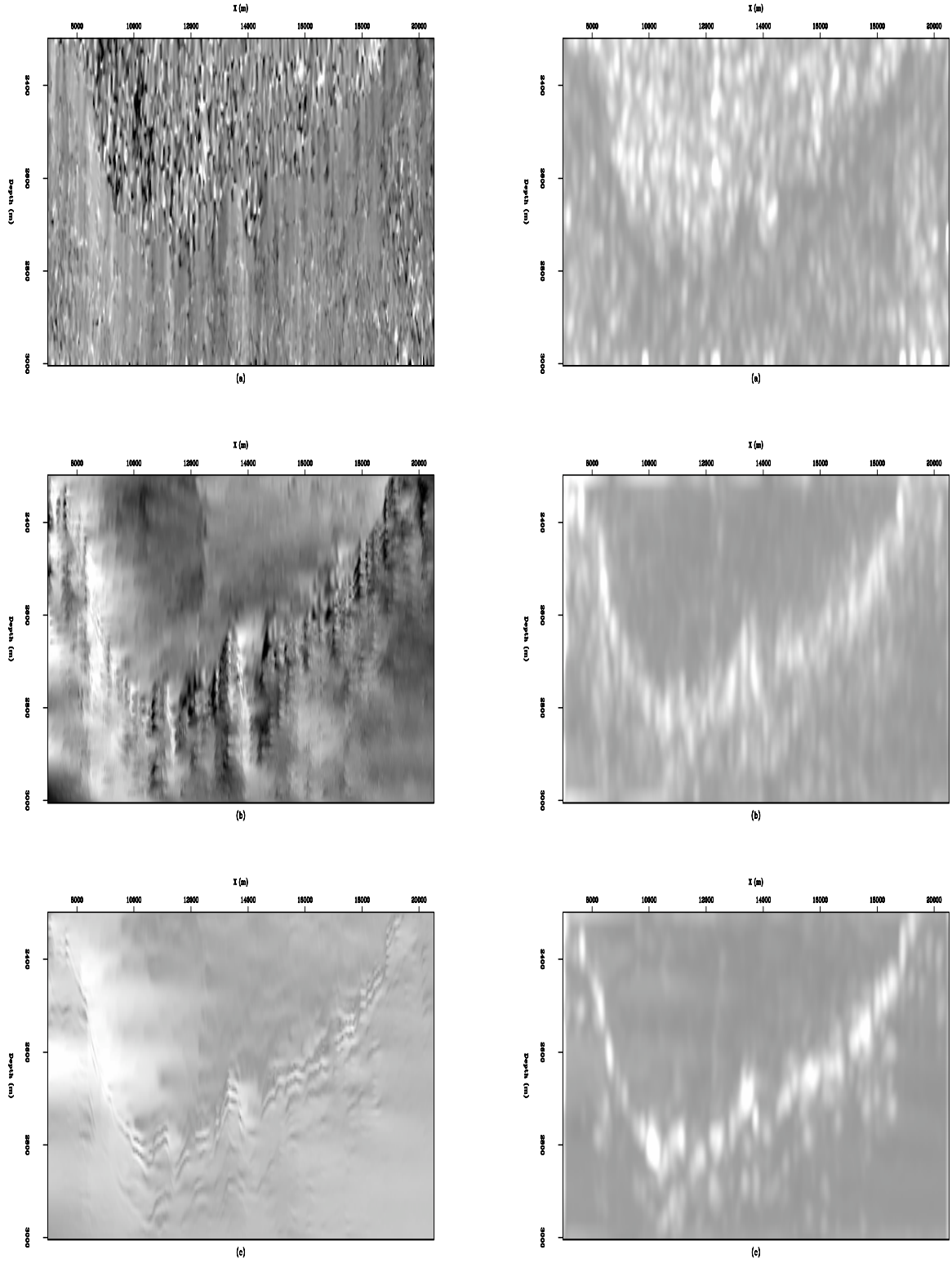


Figure 3: Dip calculations (left column) and their corresponding envelope volumes (right column) for the section in Figure 2. Row (a) uses Claerbout's puck method, row (b) uses Fomel's plane-wave destruction filtering stencil, and row (c) uses Hale's filtering stencil. **[ER]** **adam1/. gulf-er**

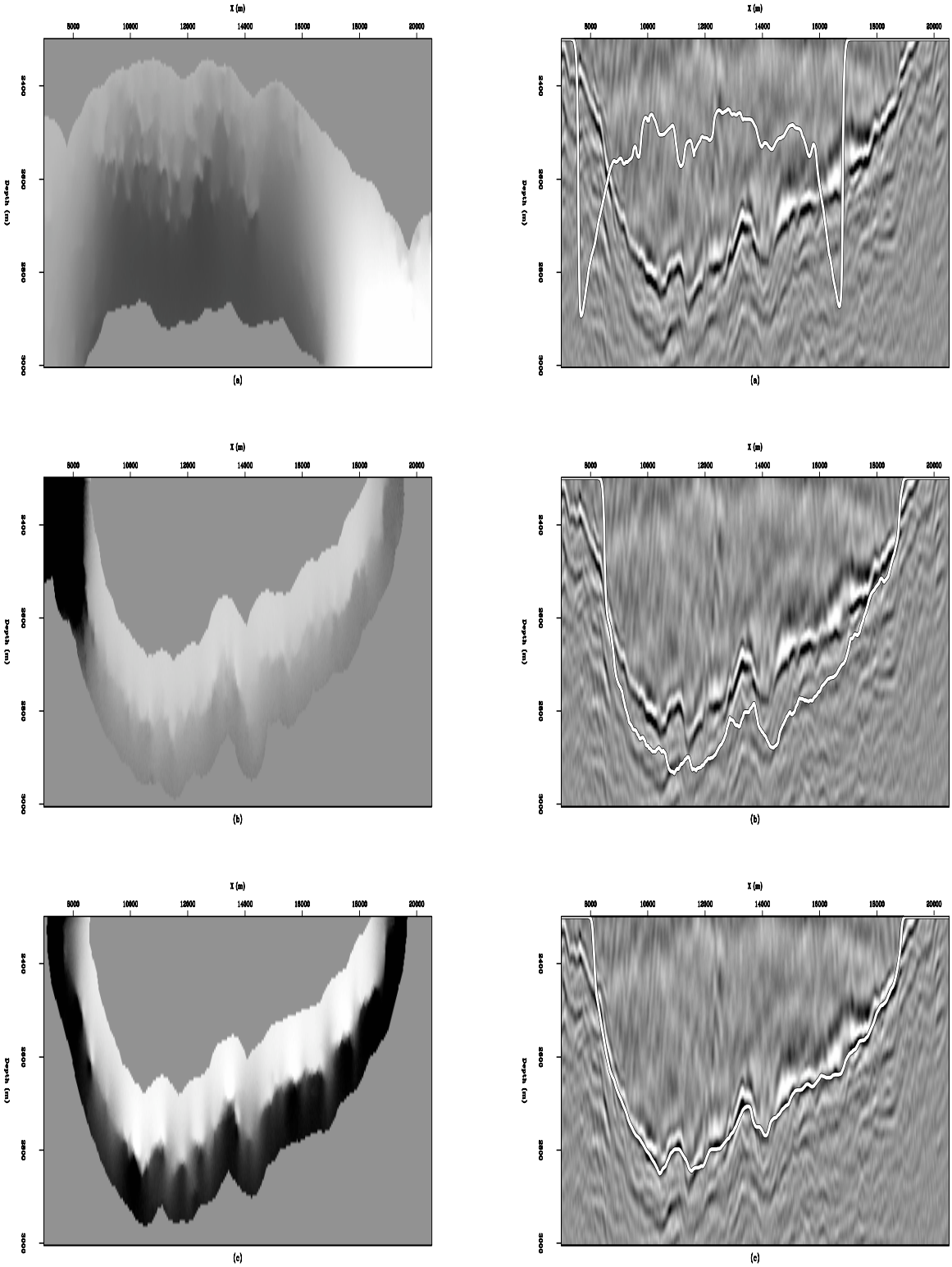


Figure 4: Eigenvectors (left column) and zero-contour boundaries (right column) derived from the dip envelope volumes in Figure 3. **[CR]** **adam1/. gulf-cr**

displays segmentation results - eigenvectors in the left column, and the corresponding zero-contour boundaries on the right - confirms this expectation. The eigenvector in panel (c) shows the clearest transition from light to dark along the boundary. Panel (b) is skewed by a strong anomaly in the upper left, and panel (a) rarely exhibits a clearly defined boundary. The corresponding zero-contour boundary picks shown in the second column differ greatly in terms of accuracy. Clearly, using Hale's stencil as part of a dip attribute segmentation algorithm produces the most accurate result in this case.

FREQUENCY AS AN ATTRIBUTE

Seismic waves behave differently inside salt bodies than they do when traveling through sediment layers. One way to take advantage of this fact is to calculate *instantaneous frequency* (Taner et al., 1979) in different parts of an image. In the complex representation of a seismic trace

$$A(t)e^{i\phi(t)}, \tag{4}$$

$A(t)$ is the amplitude of the envelope (the most common segmentation attribute), and $\phi(t)$ is the instantaneous phase. The first derivative of the instantaneous phase yields instantaneous frequency. A rapid change in instantaneous frequency may be indicative of a salt boundary, especially if it occurs in a coherent manner across an image. Figure 5 shows the instantaneous frequency calculation for the same 2D section used previously. It is apparent that frequency behavior inside the salt differs noticeably from behavior outside the salt. An eigenvector and boundary derived from this volume can be seen in Figure 6.

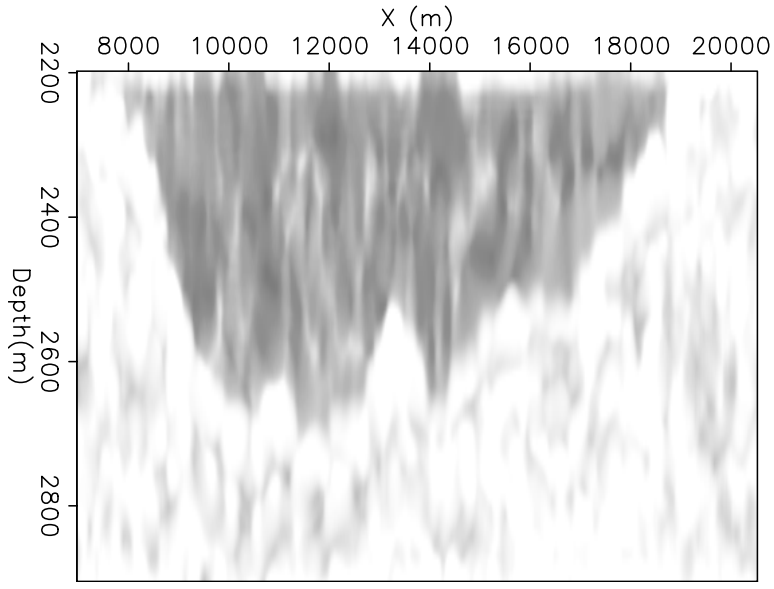


Figure 5: Instantaneous frequency calculation for the seismic section. There is a clear difference in character inside and outside the salt body. **[ER]** **adam1/. gulf-ifreq**

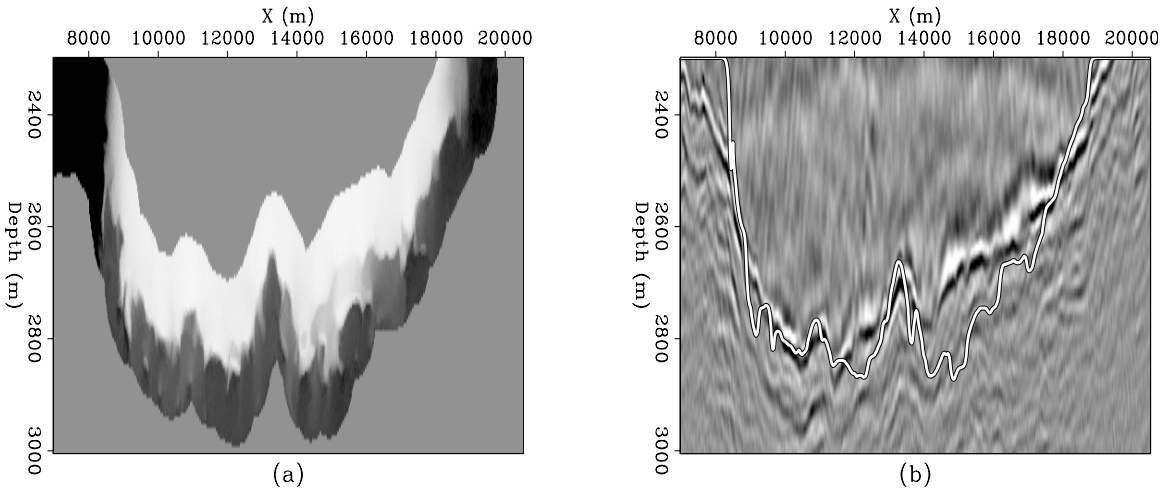


Figure 6: Eigenvector (a) and boundary calculation (b) using the instantaneous frequency attribute from Figure 5. **[CR]** **adam1/. gulf-ifreq-seg**

Continuous wavelet transform

Another method for investigating frequency properties of an image is to employ the continuous wavelet transform (Sinha et al., 2005), or CWT, to explore an image's "response" to different frequencies. Because the interior of salt bodies is often characterized by random, high-frequency noise, one application of the CWT is to improve the effectiveness of the *amplitude* attribute by eliminating this noise, thereby emphasizing the amplitude contributions of the salt boundary. Figure 7 compares envelope amplitudes for both the original seismic section in panel (a), and for the same section after a CWT has been used to isolate and stack over only the low frequencies in panel (b). The right-hand image features a much clearer boundary path, and less "noise" both inside and outside the salt body. Such an image allows for an easier segmentation process with less uncertainty.

TOWARD MULTI-ATTRIBUTE SEGMENTATION

As a way of comparing segmentation results from different attributes, Figure 8 displays three boundaries overlain on the seismic section. Each boundary is the result of segmenting with a different attribute: amplitude of the envelope, dip (from Hale's filtering stencil), or instantaneous frequency. Clearly, the results using amplitude and dip are extremely similar, and both closely track the strong reflector known to be the salt base across nearly the entire image. The amplitude boundary is more accurate in the upper left corner, while the dip boundary appears to follow a more likely path at around $x = 17000m$. This result suggests that these attributes can function either independently, or more importantly, in a complementary manner if used together in a multi-attribute segmentation scheme. While the boundary obtained using the instantaneous frequency attribute does not as accurately track the salt interface at all points of the image, it does seem to provide some useful information that may also be important in such a scheme.

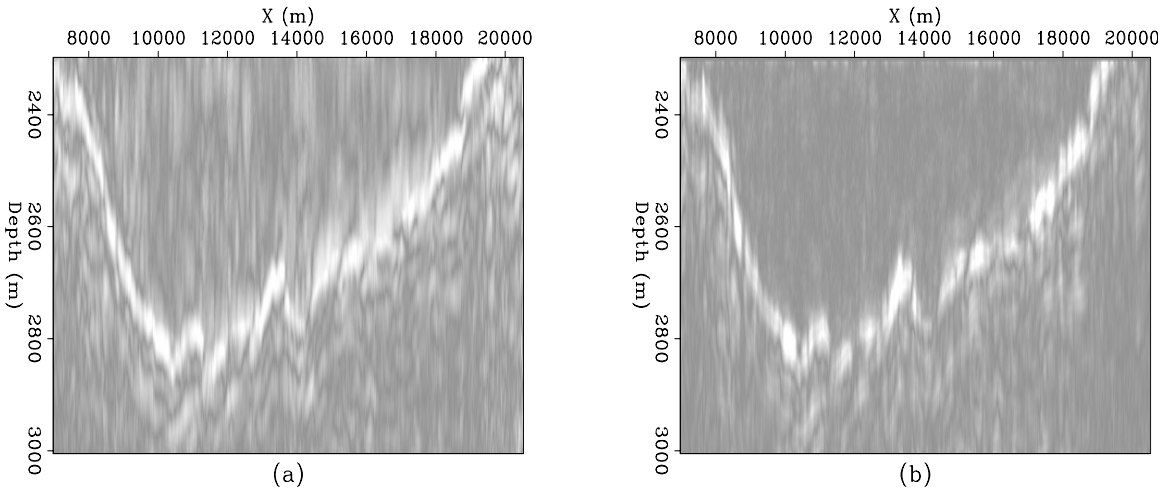


Figure 7: Amplitude envelope when all frequencies are included (a), and after a CWT is used to isolate and stack over only low frequencies (b). [**ER**] **adam1/. cwtamp**

Interpreter-guided segmentation

The purpose of this area of research is to find multiple attributes that may be useful for seismic image segmentation. This raises the important question of how to determine which attribute(s) should be used for segmentation when several are available, especially since not all attributes may be appropriate for a given situation. Here, a human interpreter's input is vital to the process. An ultimate goal for seismic image segmentation is to create an algorithm that can "learn" from the interpreter. For example, if the interpreter picks salt boundaries on a small number of 2D lines from a 3D survey, an inversion algorithm could determine which specific attributes were most important for the manual interpretation. Once this determination is made, the entire 3D volume could be segmented based on the information provided by the human interpreter. This scheme would take advantage both of humans' abilities to accurately pick boundaries on 2D sections, and computers' superior abilities to "see" in three dimensions.

CONCLUSIONS

Amplitude is only one possible seismic attribute that may be used for image segmentation. Because the NCIS algorithm does not always succeed using amplitude alone, other attributes are necessary to achieve higher accuracy. One such attribute is dip; both the nature of salt boundaries and random character of salt body interiors on seismic data make the dip variability attribute useful for segmentation. Frequency content of an image is another important attribute; calculating instantaneous frequencies throughout an image shows promise as a boundary-detection technique. Examining wavelet properties via a continuous wavelet transform also yields potentially useful information for segmentation, especially when used in combination with other attributes like amplitude of the envelope. Ultimately, these attributes could be an important part of an interpreter-aided seismic segmentation algorithm that robustly and accurately picks salt boundaries on 3D seismic volumes.

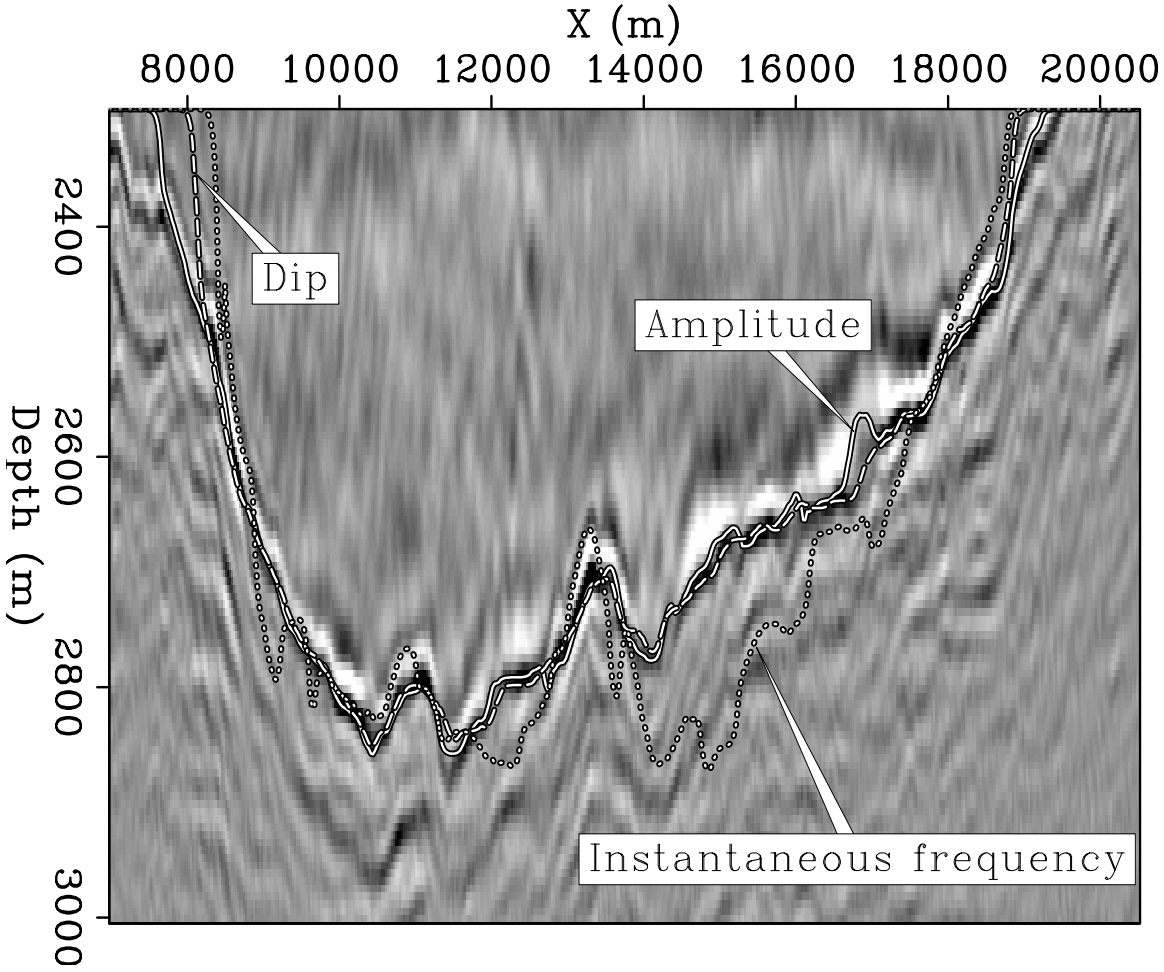


Figure 8: A comparison of boundaries calculated using three different attributes: envelope amplitude (solid line), dip from Hale's filtering stencil (dashed line), and instantaneous frequency (dotted line). **[CR]** **adam1/. bnd-ann**

ACKNOWLEDGMENTS

We would like to thank SMAART JV for the Sigsbee synthetic dataset, and WesternGeco for providing the real data example in this paper.

REFERENCES

Bednar, J. B., 1997, Least squares dip and coherency attributes: SEP-Report, **95**, 219–225.

Claerbout, J., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications.

———, 2005, Image estimation by example: Stanford University.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

Hale, D., 2007, Local dip filtering with directional laplacians: CWP-Report, **567**.

Halpert, A. and R. G. Clapp, 2008, Image segmentation for velocity model construction and updating: 2008, **134**, 159–170.

Lomask, J., 2007, Seismic volumetric flattening and segmentation: PhD thesis, Stanford University.

Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3d salt boundaries: Geophysics, **72**, P47–P56.

Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence, **22**, 838–905.

Sinha, S., P. S. Routh, P. D. Anno, and J. P. Castagna, 2005, Spectral decomposition of seismic data with continuous-wavelet transform: Geophysics, **70**, P19–P25.

Taner, M. T., F. Koehler, and R. E. Sheriff, 1979, Complex seismic trace analysis: Geophysics, **44**, 1041–1063.

van Vliet, L. J. and P. W. Verbeek, 1995, Estimators for orientation and anisotropy in digitized images, *in* Proceedings, Conference of the Advanced School for Computing and Imaging, 442–450.

Hyercube viewer: New displays and new data-types

Robert G. Clapp and Nelson Nagales

ABSTRACT

No single way to view seismic data is effective in all cases. Rather than building separate tools for each viewing approach, we added functionality to SEP's existing hypercube viewing tool. In addition to other functionality improvements, we added the capability to view wiggle traces, contours, out-of-core datasets, and datasets with different number of dimensions and size.

INTRODUCTION

Seismic viewing tools can be broken into two categories: batch viewing programs and interactive viewers. SEP has a long history of both program types. Static viewing programs can produce line graphs, wiggle traces, hidden line plots, contours, and raster images. These programs, in combination with utility programs that allow windowing and transposing of arrays offer the ability to create effective static graphics. For viewing and understanding multi-dimensional volumes, these tools are not as useful.

A series of SEP's interactive viewers (Ottolini, 1982, 1983, 1988, 1990; Clapp, 2001; Chen and Clapp, 2006) have also been developed at SEP. These viewers are more effective in viewing multi-dimensional volumes but have been limited in three key ways. The first shortcoming is their inability to produce high quality graphics and reproduce a given view of a dataset. Second, they were restricted to viewing a single (later a few) datasets that had to be identical in size and fit in memory. Finally, they only display data in raster format. Clapp et al. (2008) made an initial attempt in addressing the first shortcoming by recording, and allowing, replaying of mouse and keyboard actions.

In this paper we attempt to address the viewer's dataset and display limitations. We describe changes to the way the program interacts with datasets. These changes allow the viewer to handle datasets with different sampling and different number of dimensions, and even operate in an out-of-core mode. We show the new display options that allow the user to view the data as contour and wiggle plots. Finally we describe additional changes in the UI to the viewer and briefly discuss what might be added in the future.

DISPLAY OPTIONS

There are several possible ways to plot regular fields. SEP has generally used raster plots to display data. These plots have the advantage of being able to display relatively dense data within wide spatial frequency range. In some cases raster is not the ideal plotting option.

When examining waveforms it is often useful to display traces in a wiggle format. Wiggle format allows you to see more clearly the actual waveform recorded by the sensor. Figure 1

demonstrates hypercube's wiggle capabilities. In this case we are overlaying a velocity model on top of a migrated volume. In order to see the waveform each trace must take a few pixels of screen width, we chose to use a minimum of 8 pixels. If $8n > w$, where n is the number of traces and w is the width in pixels of the display we subsample the data volume until this criteria is met.

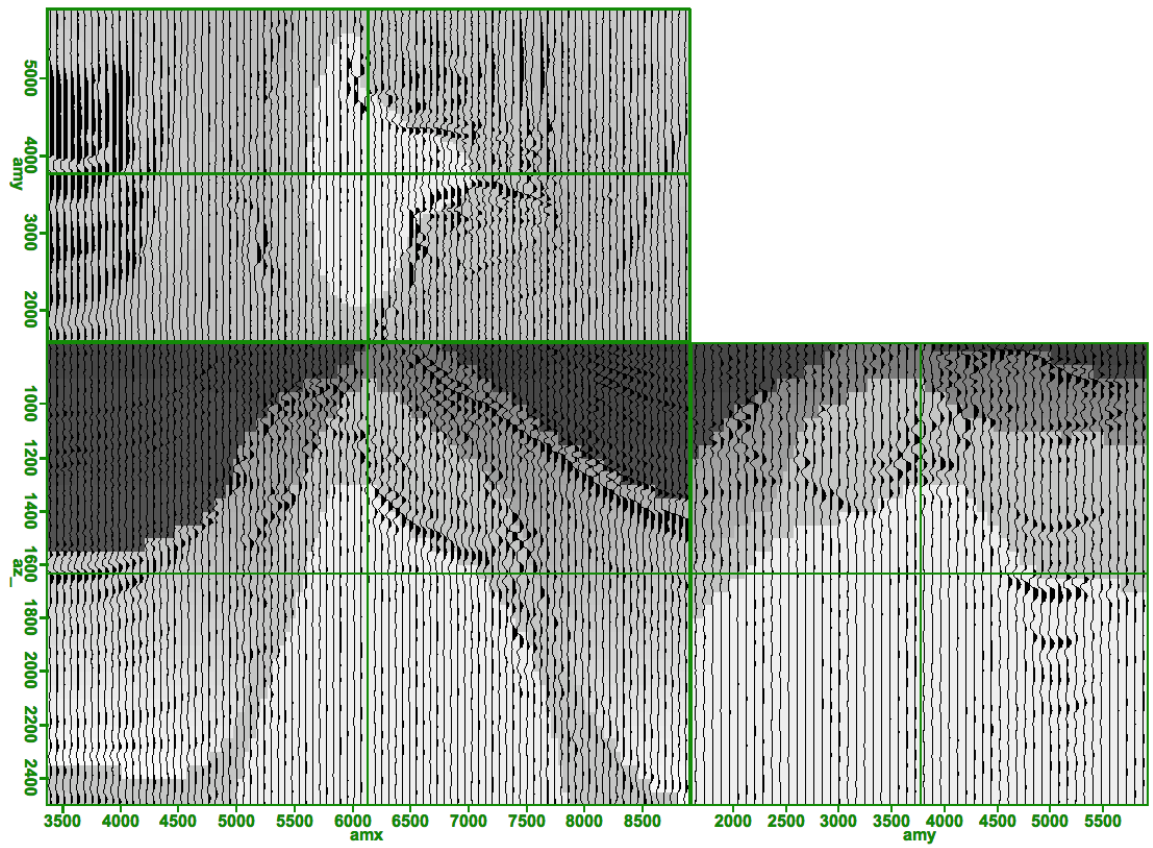


Figure 1: Velocity model overlain by a wiggle plot of a migrated volume. [**NR**] bob1/. wiggle

For fields that vary smoothly it is often useful to contour the data rather than display directly each cells value. Figure 2 demonstrates this concept by drawing velocity contours on top of migrated volume. The program attempts to separate contour labels by some distance. Note the jagged nature of the contour lines. This is due to storing the velocity data as a series of bytes rather floats.

DATASET DEFINITION

How to define a dataset poses a challenge and is where using an object language proves most beneficial. The hypercube viewer uses a grid concept similar to SEPlib and SEP3D. The user, either directly, or indirectly, describes an n-dimensional grid in which the viewer operates. Each dataset is then described in terms of this overlying grid. Each dataset must fit within this grid description. A `io_func` class is responsible for reading the dataset from

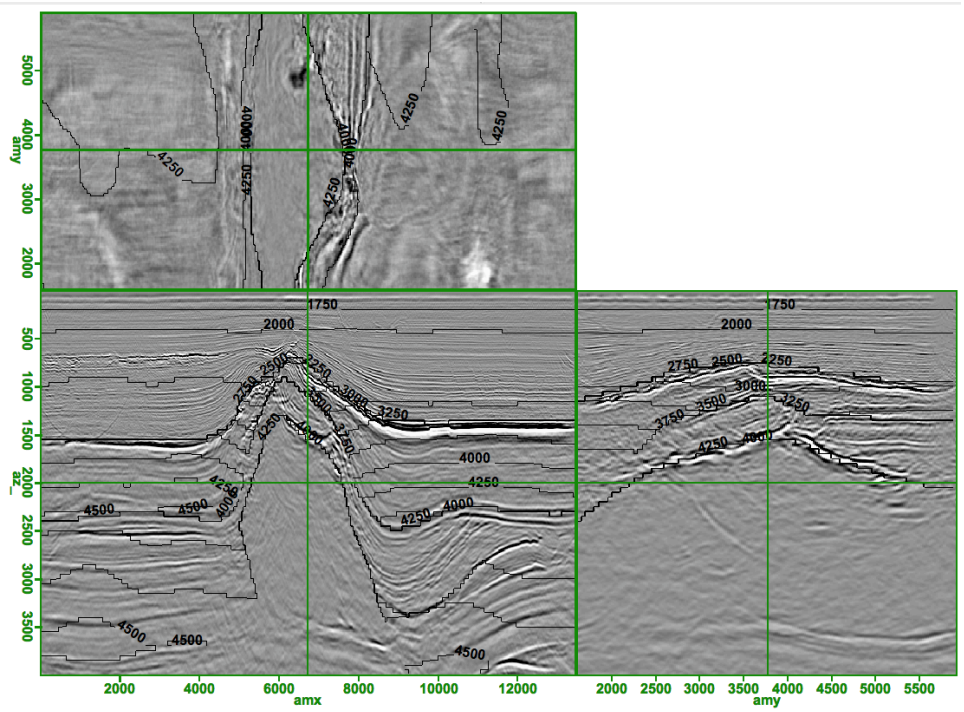


Figure 2: Migrated volume overlain by contours. Note the jagged nature of the contours due to the velocity volume being stored as bytes. [**CR**] bob1/. contour

disk or creating the dataset. A `buffer` class is used to store the data, and the `dataset` class is responsible for handling requests from other portions of the program.

IO

For conventional datasets the `io_func` class is responsible for reading from disk a subsection of the controlling grid. It first converts from the grid coordinate system to the local dataset's coordinate system. This amounts to honoring the range of each axis requested. For example, imagine the grid is defined by the table below.

Axis	**n**	**o**	**d**
1	100	0	1
2	100	0	1
3	100	0	1

And the dataset is only 2-D (an example is a grid that define time, offset, and midpoint while the dataset contains only time and midpoint) with the following sampling.

Axis	**n**	**o**	**d**
1	50	0	.2
2	1	0	1
3	80	0	1

If there is a read request for the entire volume it will be converted into a read request for the 2-D subsection of the grid in which the dataset exists. Currently the `io_func` module can read SEPlib, RSF, SU, SEG-Y, and SeisPak formats from disk in this manner.

The `io_func` does not necessarily have to read from disk, and the data isn't necessarily static. For example imagine doing interactive NMO on 2-D dataset. Three datasets can exist: the original CMP, a semblance panel, and NMO corrected gather. In this case the grid contains 4 axes: time, offset, midpoint, and velocity. None of the three datasets contain all four axes. The latter two change depending user interaction. When the midpoint location changes semblance is recomputed. When a new velocity is selected on the semblance panel the NMOed data changes. To support this type of functionality the `io_func` has a `changed` boolean.

Buffer

The `buffer` class has two main responsibilities. It holds a subsection of the data and can return a 2-D slice of the subsection it holds in memory. Currently `buffer` class stores data as either a series of bytes of floats in a regular mesh, but it is not limited to this storage mechanism. The data could be stored in number of compressed formats. A dataset can be composed of several different buffers.

Dataset

A dataset is initialized by an `io_func` and creates buffers as it needs them. Currently there are four basic buffer types. The simplest is inherited from `incore_data`. As the name implies this data type is stored completely in memory and has the same number of axes as the controlling grid. The `created_data` class is for data that changes depending on the user's interaction with the program. An example of this is the semblance and NMO corrected datasets in the velocity analysis example. If a dataset does not contain all of the axes of the grid, it should be stored using the `partial_data` class.

The `outcore_data` class, as the name implies, is for datasets that are not stored completely in memory. This class reads in a 3-D subset of the domain based on the order the axes are being displayed in a given view. For example, given a 4-D volume and a view that is displaying the first, third, and fourth axis, a 3-D subset of the data will be read centered at the current position along the second axis. As long as the position is only being changed in these three dimension no further read requests are necessary. Several of these out-of-core buffers are created for each dataset allowing multiple views display multiple subsets and allowing switching back and forth between positions without having constant data read delays.

USER INTERFACE CHANGES

Several changes where made to the user interface beyond the plotting options discussed above. In this section we will highlight the most significant changes.

Mouse/status

As discussed in Clapp et al. (2008) how the hypercube viewer interacts with the mouse can be altered. The `main` tab, submenu `info` the mouse controls can be modified. Currently there are three options: zoom, navigate, and pick. The options refer to the action of the left mouse button. The effect of the middle and right button can be displayed in the status bar.

Status bar

The status bar is located at the bottom portion of the view window. It starts off displaying the current mouse controls. The `main tab`, submenu `info` changes the status bar to display either the current position, in grid coordinates, or the actual sample value that has been selected.

Overlay/display options

The `view` tab, submenu `data` has been modified so that both the primary displayed data and the data (if any) you wish to overlay can be changed easily.

How you wish to display the primary and overlaid data can be modified in the `Display` tab, `general` submenu. For each dataset you can choose to display using various color maps, in wiggle format, or with contours. The `Display` tab contains two additional submenus, `Wiggle` and `Contour`. The `Wiggle` submenu allows you to change the line color and whether, and what color, to fill positive and negative portion of the wiggle trace. The `Contour` submenu allows you to set the number of contours, the initial contour value, and the contour interval.

FUTURE DIRECTIONS

There are three active areas of development for the viewer. The first is the ability to partition space in N-D. This ability will enable the second area, interactive processing. Many processes need human guidance but that guidance often is more complex then modifying a single 1-D line. The final area is the ability to store additional data types. The ability to store data in a compressed form, such as curvelets, would enable very large volumes to be manipulated effectively.

CONCLUSION

We expanded SEP's hypercube viewer by addressing two of its weaknesses. First we add the ability to view contour and wiggle plots. Second we added the ability to view multiple datasets which exist in the same space but might not have the sampling, the same number of dimensions, or even be possible to completely hold in memory.

REFERENCES

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**.

Clapp, R., 2001, Ricksep: Interactive display of multi-dimensional data: SEP-Report, **110**, 163–172.

Clapp, R. G., D. M. Chen, and S. Luo, 2008, Hypercube viewer: 2008, **134**, 179–192.

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.

Many-core and PSPI: Mixing fine-grain and coarse-grain parallelism

Ching-Bih Liaw and Robert G. Clapp

ABSTRACT

Many of today's computer architectures are supported for fine-grain, rather than coarse-grain, parallelism. Conversely, many seismic imaging algorithms have been implemented using coarse-grain parallelization techniques. Sun's Niagara2 uses several processing threads per computational core, therefore the amount of memory per thread makes a strict coarse-grain approach to problems impractical. A strictly fine-grain parallelism approach can be problematic in algorithms that require frequent synchronization. We use a combination of fine-gain and coarse-grain parallelism in implementing a downward continuation based migration algorithm on the Niagara2. We show the best performance can be achieved by mixing these two programing styles.

INTRODUCTION

Seismic imaging problems lend themselves well to coarse-grain parallelism. Kirchoff migration can be parallelized by splitting the image space (and/or data space) over many processing units. Downward continuation based migration can be parallelized over frequency. Flavors of downward continuation and reverse time migration can be further parallelized over shot or plane wave. All of these parallelism methods can be described as 'coarse-gained'. Coarse-grained parallelism fits well the cluster computing of the last decade. Several exciting new architectures including Nvidia's Grahic's Precision Unit (GPU), IBM's cell, Field Programable Gate Arrays (FPGA), and Sun's Niagara platform are more aimed at a fine-grained parallelism model. These platforms can have threads in the 10s-100s often making coarse-grain parallelism impractical because of memory constraints. Early results (Pell et al., 2008) on these architecture's are promising but implementation can be challenging.

Downward-continuation based migration (Claerbout, 1995) is a more challenging imaging algorithm to implement on a fine-grained parallel machine. The challenge in the implementation comes from the 2-D (shot-profile, plane-wave) 3-D (common-azimuth), or 4-D (narrow-azimuth, full-azimuth) FFT. The implicit-transpose and the non-uniform data access pattern does not easily port to FPGA and GPU solutions. The multi-thread per core approach of the Sun Niagara2 offers an easier parallelism route.

In this paper we demonstrate that the optimal solution for PSPI migration on the Niagara2 is by mixing the coarse-grained and fine-grained parallelism models. We begin by presenting an overview of the Niagara2 architecture and the PSPI algorithm. We show how some portions of the PSPI algorithm benefit from Niagara's multiple threads per core while others show only minimal improvement. We conclude by discussing the bottlenecks to further efficiency improvements.

NIAGARA2 OVERVIEW

Niagara2 is the second generation innovative CMT, Chip Multi-Threading, CPU design from Sun Microsystems, Inc. It has eight computation cores with 4 Megabytes of shared L2 cache and 4 dual channel FBDIMM memory controller. It also contains integrated networking units, PCI-Express unit, embedded wire-speed cryptography coprocessor, and built-in virtualization supports. Each core has two integer execution pipes and one floating point execution pipe shared by eight fine-grained hardware threads. In all, Niagara2 sports 64 hardware threads and combines all major server and network functions on a single chip and is well suited for power efficient secure data-center and thread level parallel computing applications.

The idea behind the chip design is that most applications are memory bound, most of the time is waiting to retrieve memory from the either cache or main memory. By having several (in this case eight) simultaneous tasks attached to each processing unit you can hide the memory latency. Figure 1 illustrates this concept. The 'M' shows a thread waiting for a memory request while the 'C' shows computation. At each clock cycle computation is being performed and the time associated with memory requests are hidden.

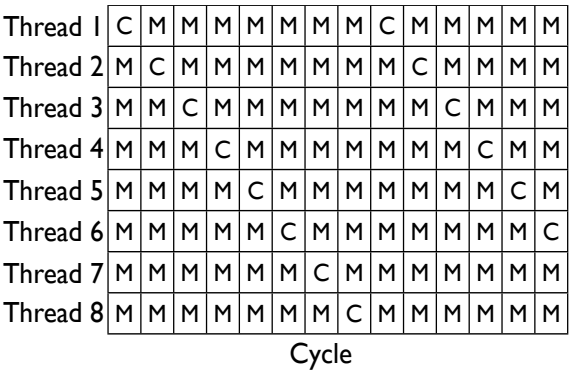


Figure 1: The idea behind the Niagara architecture. The 'M' shows a thread waiting for a memory request while the 'C' shows computation. At each clock cycle computation is being performed and the time associated with memory requests are hidden. [**NR**] bob2/. niagara2

The Niagara2 platform performs well on an application when two requirements are met. First, that the problem is truly memory bound. This is a function of memory access speed, memory hierarchy, and the compute engines of a given core. Second, the parallelism granularity of the application cannot require a siginificant level of synchronization.

PSPI MIGRATION

Downward continued migration comes in various flavors including Common Azimuth Migration (Biondi and Palacharla, 1996), shot profile migration, source-receiver migration, plane-wave or delayed shot migration, and narrow azimuth migrations. For downward continued based migration there are four potential computational bottlenecks that vary depending on the flavor of the downward continuation algorithm. The Phase-Shift Plus Interpolation (PSPI) method is one of the easier methods to implement. The computational cost is dominated by the cost of downward propagating a wavefield at a given frequency w, a given depth step z. Within this loop the wavefield is Fourier transformed, a correction term in the FX domain is applied, and the wavefield is downward continued in the FK domain. Pseudo code for the algorithm takes the following form,

```
Loop over w{  !CORASE
  Loop over z{
    Loop over source/receiver{
     Loop over v{
      FX    !FINE
      IFFT  !FINE
      FK    !FINE
    }
    FFT     !FINE
  }
  }
}
```

In many cases the dominant cost is the FFT step. The dimensionality of the FFT varies from 1-D (tilted plane-wave migration (Shan and Biondi, 2007)) to 4-D (narrow azimuth migration (Biondi, 2003)). The FFT cost is often dominant due its $nlog(n)$ cost ratio, n being the number of points in the transform, and the non-cache friendly nature of multi-dimensional FFTs. The FK step, which involves evaluating a square root function and performing complex exponential is a second potential bottleneck. The high operational count per sample can eat up significant cycles. The FX step, which involves a complex exponential, or sine/cosine multiplication, has a similar, but computationally less demanding, profile.

RESULTS

For this test we used a Common Azimuth Migration (CAM) variant of the PSPI algorithm discussed above. The FK, FX, and FFT are all performed on a 3-D field. We began from a code that used coarse-grain parallelization over frequency. We used a relatively small domain size (574x256x52) which is well beyond the L2 cache of the system but still allowed a large series of tests to be run in a reasonable amount of time. Figure 2 shows the normalized performance of the entire algorithm as a function of coarse-grain threads. Note how we achieve linear speed up all the way to 9 threads. Going beyond 9 coarse threads was not possible given the machine's memory.

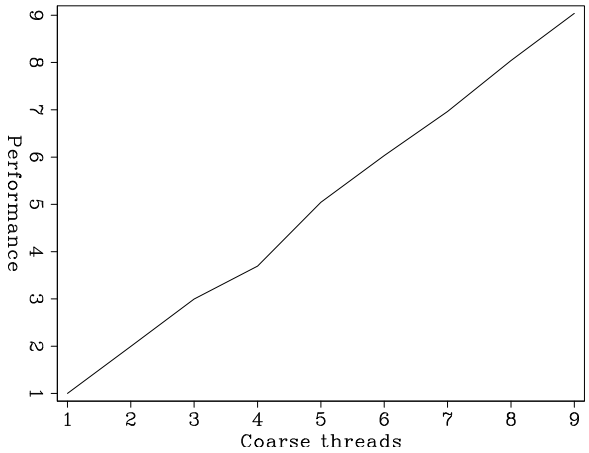


Figure 2: Performance as a function of the number of coarse-grained threads. [**NR**] bob2/. coarse

We then parallelized the `FX, FK`, and `FFT` routines. The `FK` and `FX` routines are sample by sample operations well suited to fine-grain parallelism and generally trivial to parallelize using the pthreads library. For the `FFT`, we used Sun's prime factor FFT rather than FFTW. The single-thread performance of the Sun's library was nearly double FFTW's performance. Figure 3 shows the normalized performance as the number of fine-grain threads increase. Note how we achieve nearly no performance gain after 32 threads. Figure 4 explains the lack of improvement. It shows the performance of the FFTW, FK, and FX steps portion of the algorithm. After 20-25 threads the FFT shows no performance improvements. This is not surprising due to the synchronization inherent in the FFT algorithm.

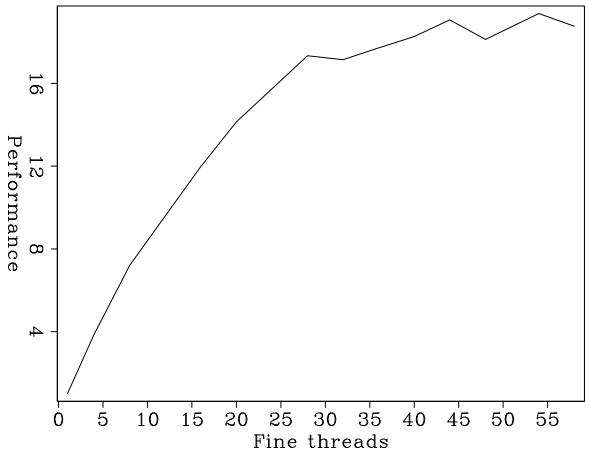


Figure 3: Normalized performance as a function of the number of fine-grained threads. Note that little performance gain is achieved after 32 fine-grain threads. [**NR**] bob2/. fine

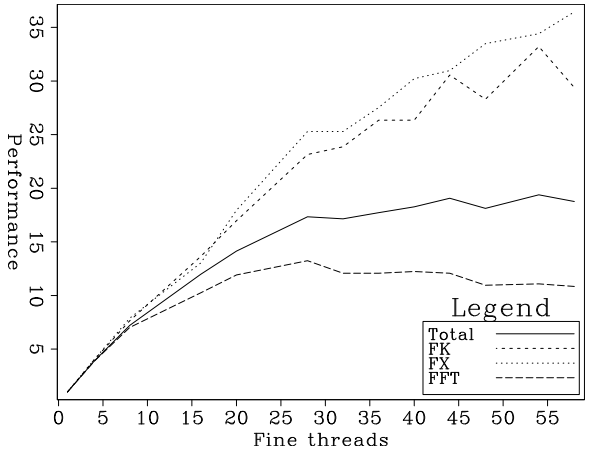


Figure 4: Performance as a function of the number of fine-grained threads for different parts of the algorithms. Not the nearly linear speed up of the FK and FX steps while performance peaks for the FFT at 20 threads. [**NR**] bob2/. part

As a final test we combined the coarse-grained and fine-grained approaches. Figure 5 shows the maximum performance as a function of coarse-grain threads n_c. For each coarse-grain thread we used n fine-grain threads where $n = floor(63/n_c)$, maximizing the available threads on the machine. Note that the graph is normalized by the single thread performance. Peak performance was achieved using 6 or more coarse-grain threads. Figure 6 shows the break down by function. Not surprisingly, the FK and FX step show nearly constant performance independent of the number coarse-grain vs. fine-grain threads. On the other hand the FFT benefits from less fine-grained parallelism bringing up the overall total performance of the algorithm.

As a comparison we run the same code on 4-core 1.8GHz, dual processor intel machine.

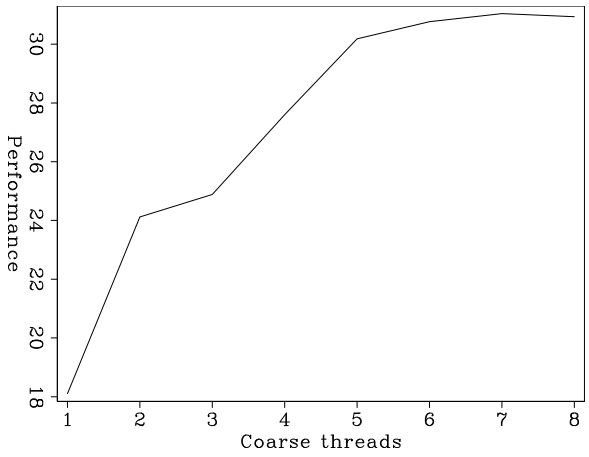


Figure 5: Performance as a function of number of coarse-grain threads. The number of fine-grain n_f threads per coarse-grain n_c threads is $n_f = floor(63/n_c)$. [**NR**] bob2/. best

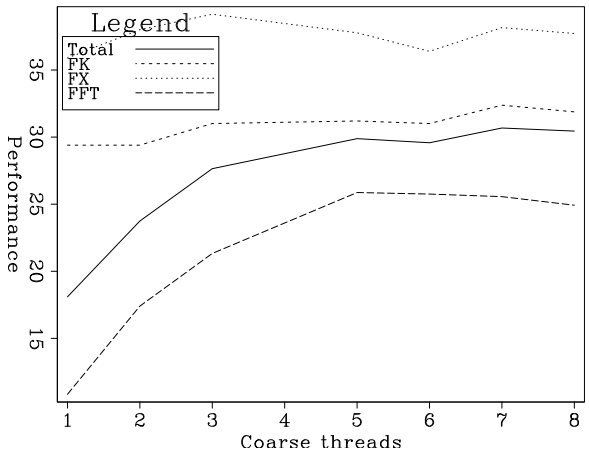


Figure 6: Performance of different routines with the PSPI algorithm as a function of the number of coarse-grain threads. Note how the FFT benefits the most from more coarse-grain parallelism. [**NR**] bob2/. partb

Running 8 coarse-grain threads and normalizing in terms of the the Niagara2 single thread results we found:

Segment	**Relative speed**
FFT	23.9
FX	10.9
FK	61.8
Total	23.9

In general the code scaled linearly with number of processors.The `FX` ratio was low compared to the Niagara because of the significant memory requests (due to the velocity correction), something that the Niagara2 archetecture is well designed for. The `FK` number was large due to the vector nature of the computation and its high floating point operation count.

Improving the floating point and vector potential of the Niagara architecture could have a large impact on these results. Both the FFT and the FK steps involve significant floating point computations. The synchronization requirements of the FFT algorithm limits effective scaling to approximately 16 threads even for large volumes.

CONCLUSIONS

We implemented PSPI migration on the Sun Niagara2 by combining coarse-grained and fine-grained prallelism. We showed that the multi-thread per core model leads to significant uplift in performance over a single thread approach. Compared to a strictly fine grain parallelism we achieved a 60% uplift. Compared to a coarse grain approach the improvement was 5X. Improved floating point/vector performance could lead to signficant uplift for this aglorithm.

REFERENCES

Biondi, B., 2003, Narrow azimuth migration of marine streamer data: SEG, Expanded Abstracts, **22**, 897–900.

Biondi, B. and G. Palacharla, 1996, 3-d prestack migration of common-azimuth data: Geophysics, 1822–1832, Soc. of Expl. Geophys.

Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.

Pell, O., T. Nemeth, J. Stefani, and R. Ergas, 2008, Design space analysis for the acoustic wave equation implementation on fpga circuits: European Association of Geoscientists and Engineers, Expanded Abstracts, 1406–1409.

Shan, G. and B. Biondi, 2007, Angle domain common image gathers for steep reflectors: 2007, **131**, 33–46.

Reverse time migration: Saving the boundaries

Robert G. Clapp

ABSTRACT

The need to save or regenerate the source or receiver wavefield is one of the computational challenges of Reverse Time Migration (RTM). The wavefield at each time step can be saved at the edge of the damping/boundary condition zone. The wave equation can be run in reverse, re-injecting these saved points to regenerate the wavefield. I show that this a better choice than checkpoint schemes as the domain grows larger and if the computation is performed on a streaming architecture.

INTRODUCTION

Reverse time migration (Baysal et al., 1983; Etgen, 1986) is quickly becoming the high-end imaging method of choice in complex geology. One of the computational challenges of RTM is that the source wavefield is propagated forward in time (0 to $t_{\max}$) while the receiver wavefield is propagated backwards in time ($t_{\max}$ to 0), yet the imaging step requires these two fields to be correlated at each time t. Storing one of the wavefields in memory is impractical for 3-D problems. The most obvious solution is to store the wavefield at each imaging step on disk. This requires significant disk storage on each node and can cause the problem to quickly become Input/Output (IO) bound. Symes (2007) proposed a checkpointing scheme where a smaller number of snapshots are stored to disk and intermediate wavefield are regenerated. Another approach, alluded to in Dussaud et al. (2008), is to save the wavefield at boundaries of the computational domain and to re-inject them.

In this paper, I demonstrate how to implement a boundary re-injection scheme. I discuss the computational tradeoffs of boundary re-injection vs. a checkpoint scheme and conclude that, as the computational domain increases in size, the boundary method proves superior.

REVERSE TIME MIGRATION REVIEW

Reverse time migration is an attractive imaging method because it does not suffer from the limitations of the Kirchhoff and downward continuation based imaging approaches. A single, or even multiple, travel-time can not accurately describe wave propagation in complex media, making Kirchhoff methods ineffective under salt. The angle limitation, and difficulty/inability to handle multiple bounces limits the effectiveness of downward continuation based methods with complex structures.

Algorithm

The basic idea behind RTM is fairly simple: we are reversing the propagation experiment. Given data d recorded from $1...nt$ with sampling dt, we begin by re-injecting into our earth

model that data recorded at nt. We then propagate this data back into the earth a length in time dt. We can imagine storing the current status of the receiver wavefield W_r into the last elements of 4-D volume in terms of x, y, z, t. We then re-inject at the surface data recorded at $nt - 1$. We follow this procedure until we have re-injected to the first time sample.

We also need to simulate the source portion of the experiment. We input at the source location of the earth model the first sample of the source pulse. We then propagate the wavefield dt, insert the next sample of the source pulse and continue until we have reached $dt * nt$ time. At each dt we store in a second wavefield array W_s the wavefield at each time sample. The zero offset migrated image I is formed at each iz, ix, iy location by

$$I(ix, iy, iz) = \sum_{it=0}^{nt} W_s(ix, iy, iz, it) W_r(ix, iy, iz, it). \tag{1}$$

.

Cost

The major disadvantage of RTM is that the cost is generally thought to be an order of magnitude more expensive than wave equation or Kirchoff alternative. Calculating the cost of the various imaging methods is a tricky proposition. In general people use two different metrics for comparing cost. The first is to simply count operations; the second counts the number of memory access requests. Which method is appropriate depends on whether you are saturating the memory bus or computation units. Kirchhoff migration is more likely to be limited by memory access, while downward continuation is usually limited by the computational units.

In today's world of multi-core, many-core, streaming architectures a third metric needs to be applied: how parallelizable is the method? This is in many ways a more difficult assessment because it strongly depends on the type of hardware and the type of parallelization. Many-core and streaming architectures are most effective with problems that are compute-bound and can have a high level of fine-grained parallelism. **?** showed how downward-continuation based methods can be limited by the data-dependency of the FFT.

The main cost of RTM is advancing the wavefield in time. This is usually done by solving the wave-equation using finite-differences. The basic idea is to start from some version of the wave equation, for example the acoustic version where u is the wavefield,

$$\frac{\partial^2 u}{\partial t^2} = v^2 (\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}). \tag{2}$$

Each second derivative is approximated by doing a Taylor (or similar method) expansion. The difference approximation affects the accuracy of the propagation along with the required sampling of the wavefield in time and space to maintain stability and avoid dispersion. In general, a much higher order accuracy approximation is used in the space domain than the time domain, often from 6th to 14th order. The filter implied by this approximation can be quite large, from 19 to 43 points in size, and is the dominant computational cost. On the other hand, the structure of the computation is quite simple with virtually no data-dependency, and amenable to a very high level of fine-gain parallelism. As a result it is

ideal for many-core and streaming architectures. The speed advantage of FFT and sparse matrix (Kirchoff) approaches diminishes significantly with these platforms.

As the speed of the wavefield propagation increases, a new bottleneck appears in a standard RTM implementation. First, the size of W_r and W_s are well beyond a conventional system's memory. In addition, note how W_r and W_s are filled in reverse order. The receiver wavefield stores $nt, nt-1, ...1$ while the source is filled $1, 2, ...nt-1, nt$. The obvious solution is to store one of the two fields (from now on I will choose the receiver wavefield) to disk. This requires a large but feasible amount of storage. The problem is that writing to disk is orders of magnitude slower than accessing memory. With an optimized implementation of the propagator, reading and writing the wavefield becomes the bottleneck.

The most common solution to this problem is to save a subset of the receiver wavefields. **?** describes an 'optimal' checkpointing scheme which minimizes the total number of imaging wavefields that need to be recomputed by storing a series of checkpoints either in main memory or on disk. A similar approach is linear checkpointing which stores every jt imaging steps where jt is a function of the amount of memory on the system.

SAVING THE BOUNDARY

Another approach is regenerating the wavefield, taking advantage of the reversibility of the wave equation. For example, imagine using a second order in time finite-difference scheme. Given the wavefield at the current time w_t and previous time $w_{t-\Delta t}$ we can find the wavefield at the next time $w_{t+\Delta t}$ through

$$w_{t+\Delta t} = w_{t-\Delta t} + w_t + Lw_t, \tag{3}$$

where L calculates the second derivative. We can reverse $w_{t-\Delta t}$ and w_t in equation 4 to find thew wavefield at $w_{t-2\Delta t}$,

$$w_{t-2\Delta t} = w_t + w_{t-\Delta t} + Lw_{t-\Delta t}. \tag{4}$$

Figure 1 demonstrates this property. Panel 'A' shows a wavefield at time t, panel 'B' shows the wavefield after $2t$ at which stage the calculation is reversed. Panel 'C' shows the regenerated data, again at time t. Panel 'D' shows the difference between the regenerated wavefield and the original wavefield (with the clip at 1/10th panels 'A','B', and 'C'. Note that even energy that has hit the boundary has been handled correctly.

The problem comes when we attempt to kill energy entering the boundary. No longer is our time reversal scheme valid, as equation 4 does not fully describe what is being applied to the wavefield. Figure 2 demonstrates this concept. In this figure a damping boundary condition has been applied. Note how the wavefield and the reconstructed wavefield vary significantly.

Figure 3 demonstrates a way to solve this problem. Imagine our first time reversal step. The grey area represents cells where a boundary condition is being applied. Any locations where the filter implied by $\mathbf{L}$ touches a grey region will lead to incorrect reverse propagation. The area in black in Figure 3 show regions which uses cells where the boundary condition has been applied. If we save the wavefield from the forward propagation in the black region, we can substitute them in when doing the reverse propagation. Figure 4 shows the result

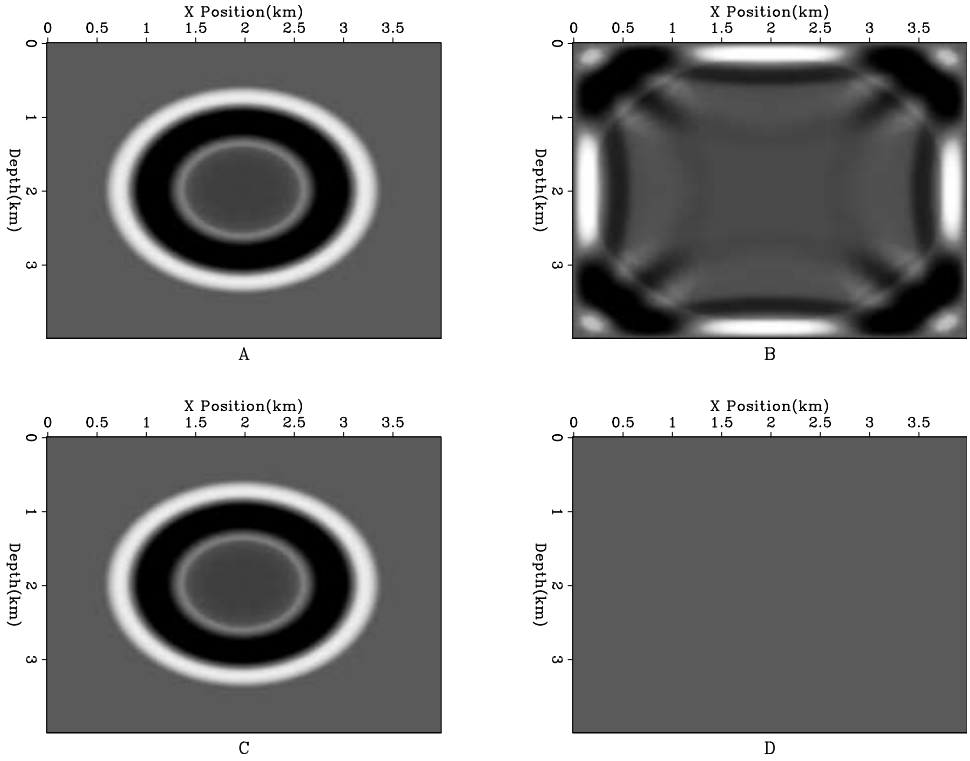


Figure 1: Panel 'A' shows a wavefield at time t, panel 'B' shows the wavefield after $2t$ at which stage the calculation is reversed. Panel 'C' shows the regenerated data, again at time t. Panel 'D' shows the difference between the regenerated wavefield and the original wavefield. [**ER**] bob4/. nobound

of this save and replace scheme. Note how we achieve a perfect result everywhere except the damping zone.

Note how the size of the black region is a single sample, while as previously noted, we tend to use a high order approximation in the space domain (which would make the black area larger). For this example I reduced the derivative approximation as I approached starting from 10th order solution and going to a second order at the boundary between the damped and undamped region. As a result only a single point along the boundary sees the damped region.

BOUNDARY VS. CHECK-POINTING

There are several metrics in comparing the 'cost' of using checkpointing vs. boundary saving approach. These are the amount of disk IO, IO throughput required, the amount of main memory needed for an optimal solution, and the ease of implementation. I will compare the linear checkpointing vs. boundary saving approach, but the optimal approach would behave similarly.

To setup the comparison let's assume that n is the length of the domain in x, y, and z. We have nt propagation steps, ji is the the number of propagation steps between imaging

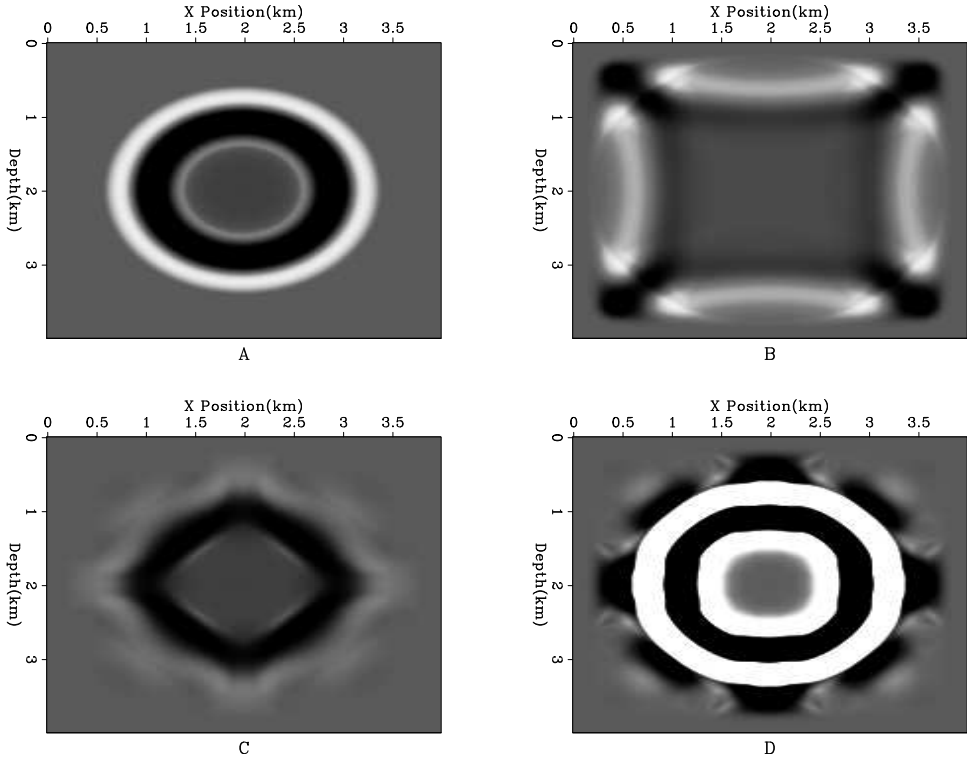


Figure 2: Panel 'A' shows a wavefield at time t, panel 'B' shows the wavefield after $2t$ at which stage the calculation is reversed. Panel 'C' shows the regenerated data, again at time t. Panel; 'D' shows the difference between the regenerated wavefield and the original wavefield. In this case a damping boundary condition has been applied around the edges of the domain. [**ER**] bob4/. bound

steps, and for the checkpointing scheme we will save every jc imaging time. Each approach does the same amount of propagation step $nt * 3$. The amount of disk required for the checkpointing m_c scheme is

$$m_c = \frac{2 * nt * n^3}{ji * jc}. \tag{5}$$

For saving the boundaries we need to save slices around the edge of the cube rather than the entire cube, but we need to save at every time step. The memory requirement m_b is then

$$m_b = 6nt * n^2. \tag{6}$$

Checkpointing requires less disk when we save less than every $\frac{n}{3*jt*jc}$ steps. Put another way, the larger the migration aperture the more advantageous saving the boundaries.

Checkpointing requires larger volumes, but less frequent reads; buffering read requests are a necessity with a checkpoint approach, but add to memory requirements. For true streaming hardware this can still be problematic because the entire buffer must be re-passed to the streaming engine. The boundary approach requires a much smaller volume to be passed continuously.

The memory requirement of the two systems is significantly different. The checkpointing

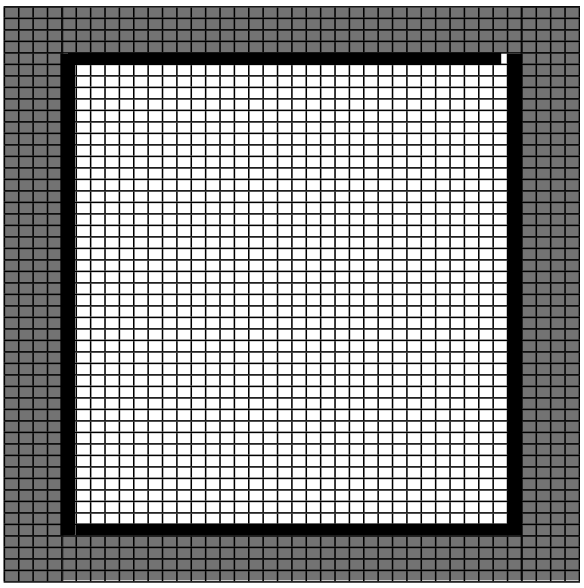


Figure 3: Computational grid for propagation. Grey area is cells effected by applying boundary conditions. Black area is cells that are saved and re-injected. [**ER**] bob4/. box

scheme must redo the propagation in the same direction or suffer problems at the boundaries. This means that we must save jc copies of wavefield volume in memory, exchanging disk space for memory space. Again, this can be problematic if the imaging step is done with a hardware accelerator. It requires the volume to be read from disk and passed to the accelerator and each imaging step to be passed back to main memory, then sent back to that accelerator for producing the image. In the boundary saving approach the computations are done in the same directions, greatly reducing communication requirements.

Neither scheme is difficult to implement on a conventional CPU. Checkpointing benefits more from smart overlapping IO and compute but this doesn't add significant complexity. On accelerators the checkpointing scheme, with its significant additional data movement, is significantly more difficult to implement and optimize.

CONCLUSIONS

RTM lends itself well to fine grain parallelism and hardware acceleration technologies. Disk IO becomes the bottleneck with faster wavefield propagation. Checkpointing and saving boundary regions can reduce the IO cost while increasing the compute cost. For large problem saving the boundary is the more efficient mechanism.

REFERENCES

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Dussaud, E., W. W. Symes, L. Lemaistre, P. Singer, B. Denel, and A. Cherrett, 2008, Computational strategies for reverse-time migration: 78th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, SPMI 3.3.

Etgen, J., 1986, Prestack reverse time migration of shot profiles: SEP-Report, **50**, 151–170.

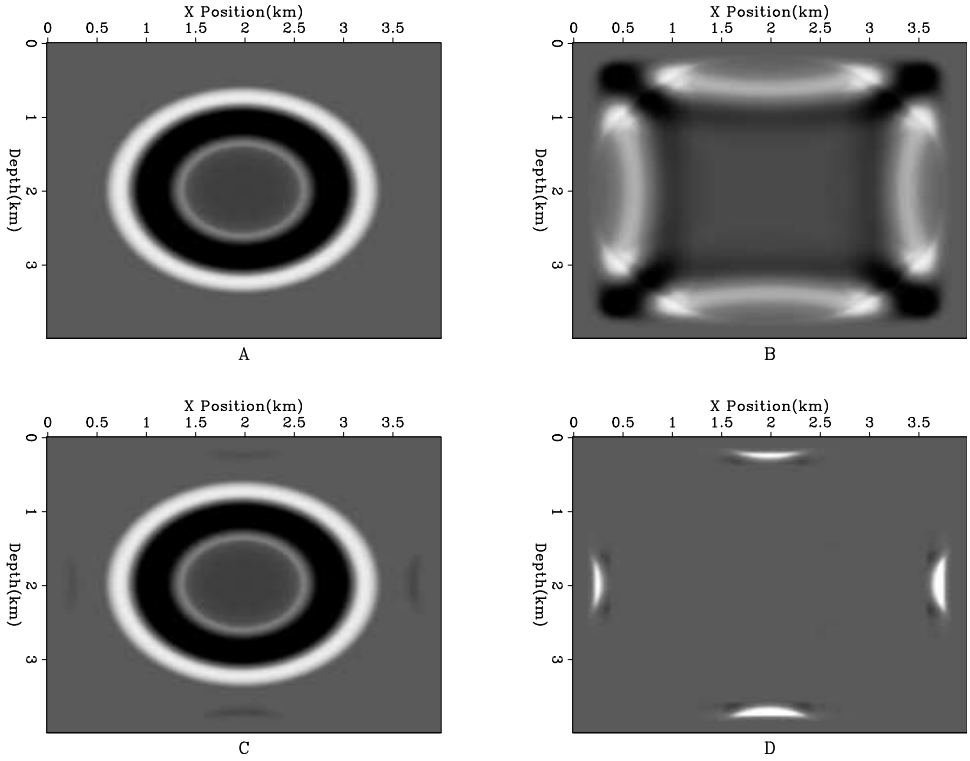


Figure 4: Panel 'A' shows a wavefield at time t, panel 'B' shows the wavefield after $2t$ at which stage the calculation is reversed. Panel 'C' shows the regenerated data, again at time t. Panel 'D' shows the difference between the regenerated wavefield and the original wavefield. In this case a damping boundary condition has been applied around the edges of the domain and the wavefield within this region has been re-injected. [**ER**] bob4/. inject

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–SM221.

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 2D coordinate systems. I develop an expression linking the definition of reflection opening angle to various generalized geometric factors. I demonstrate that generalized coordinate ADCIGs can be calculated directly using Fourier-based offset-to-angle approaches for coordinate systems satisfying the Cauchy-Riemann differentiability criteria. The canonical examples of tilted Cartesian, polar, and elliptic coordinates are used to illustrate the ADCIG theory. I compare analytically and numerically generated image volumes for a set of elliptically shaped reflectors. Experiments with a synthetic data set illustrate that elliptic-coordinate ADCIGs better-resolve the reflection opening angles of steeply dipping structure, relative to conventional Cartesian image volumes, due to improved large-angle propagation and enhanced sensitivity to steep structural dips afforded by coordinate system transformations.

INTRODUCTION

Angle-domain common-image gathers (ADCIGs) are used increasingly in seismic imaging to examine migration velocity model accuracy (**?**). The key idea is that migrating with the correct velocity model leads to flat angle gathers that shift neither vertically nor horizontally as a function of reflection opening angle. Migrating with an incorrect velocity, though, leads to inconsistent angle-domain reflectivity and generates residual curvature in the ADCIG volume. ADCIGs are thus an effective velocity analysis tool and have been incorporated in wave-equation-based inversion schemes to update velocity profiles (**??**).

Wave-equation imaging techniques generate ADCIGs in straightforward manners for both shot-profile (**???**) and shot-geophone (**??**) migration approaches. In shot-profile migration, one first generates 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 using, for example, post-imaging Fourier-based operators (**?**).

Conventional ADCIG theory usually assumes horizontal wavefield shifts, largely because wavefield extrapolation and imaging are most commonly performed in Cartesian coordinates. However, a number of studies have noted that these ADCIG results degrade for steeply dipping structures, such as salt flanks (**??**). Although this is partially due to problems associated with inaccurate large-angle extrapolation, ADCIGs calculated using horizontal wavefield shifts become decreasingly sensitive for increasingly steep structural dips. **?** demonstrate that this problem can be addressed by generating ADCIGs with vertical subsurface-offset-domain common-image gathers (VODCIGs); however, this approach

is less desirable computationally because it requires storing the larger wavefield volumes required to calculate the VODCIGs in memory.

The introduction of shot-profile migration in more general coordinate systems [e.g. tilted Cartesian (**?**) and elliptic meshes (**?**)] presents an opportunity to circumvent problems associated with generating ADCIGs for steeply dipping structure. In particular, migration domains can be oriented such that geologic structures with steep dips in Cartesian meshes have relatively gentle dip in generalized coordinate systems, thus improving the robustness of the ADCIG calculation. Developing an ADCIG theory capable of handling more arbitrary coordinate meshes, though, requires proper treatment of the effects of non-Cartesian geometries. For example, wavefield extrapolation in non-Cartesian coordinate systems induces local wavenumber stretches, rotations and/or shearing (**?**). Similarly, non-uniform wavefield sampling can lead to anisotropic angle-domain stretching. These effects can be corrected using Jacobian change-of-variable transformations.

The goal of this paper is to extend ADCIG theory to non-Cartesian geometries. I demonstrate that ADCIG theory, as developed in a differential sense (**?**), remains valid for arbitrary geometries provided that the corresponding derivative operators are properly specified. Non-Cartesian coordinates do, however, introduce space-domain geometric factors that can render Fourier-based offset-to-angle methods unsuitable. However, I show that ADCIGs can be calculated directly in the Fourier domain for all coordinate systems satisfying the Cauchy-Riemann differentiability criteria (**?**). Moreover, ADCIGs can be calculated in all situations using the slant-stack approaches discussed in **?**.

I begin by discussing how to generate subsurface offsets and ADCIGs in Cartesian coordinates. I then provide an extension to generalized coordinate systems based on Jacobian change-of-variable arguments. I examine two canonical coordinate systems, tilted Cartesian and elliptic meshes, where the reflection angle can be explicitly calculated using Fourier-based methods, and a third, polar coordinates, where it cannot. I test the generalized ADCIG theory analytically and numerically using a set of elliptic reflectors, and demonstrate how computing angle gathers in elliptic coordinates can lead to improvements relative to Cartesian coordinates, especially for steeply dipping structure.

ADCIG THEORY

The ADCIG theory presented in this section draws from that presented in **?**. In the ensuing development, $\boldsymbol{x} = [x_1, x_3]$ denotes the Cartesian variables and $\boldsymbol{\xi} = [\xi_1, \xi_3]$ represents a generalized Riemannian coordinate system. I also use a convention where the extrapolation axis is oriented in the x_3 (ξ_3) direction for Cartesian (Riemannian) coordinates. Coordinates $x_2 = \xi_2 = 0$ are assumed throughout.

Generating subsurface offsets axes

Shot-profile migration in Cartesian coordinates consists of completing a recursive two-step procedure. The first step involves propagating the source and receiver wavefields, S and R,

from depth level $x_3 - \Delta x_3$ to x_3 using an extrapolation operator $E_{x_3}[\cdot]$

$$\begin{aligned} E_{x_3}[S(x_3 - \Delta x_3, x_1|\omega)] &= S(x_3, x_1|\omega), \\ E^*_{x_3}[R(x_3 - \Delta x_3, x_1|\omega)] &= R(x_3, x_1|\omega), \end{aligned} \tag{1}$$

where $*$ denotes the conjugate operator, ω is angular frequency, and Δx_3 is the depth step. A subsurface image, I, is subsequently computed at each extrapolation step by evaluating an imaging condition

$$I(x_3, x_1, h_{x_1}) = \sum_{\omega} S^*(x_3, x_1 - h_{x_1}|\omega) R(x_3, x_1 + h_{x_1}|\omega), \tag{2}$$

where the subsurface-offset axis, h_{x_1}, is generated by correlating the source and receiver wavefields at various relative shifts in the x_1 direction. Finally, the ADCIG volume is computed using an offset-to-angle transformation operator, $\mathbf{T}_{h_{x_1} \to \gamma}$

$$I(x_3, x_1, \gamma) = \mathbf{T}_{h_{x_1} \to \gamma} I(x_3, x_1, h_{x_1}), \tag{3}$$

where γ is the reflection opening angle shown in Figure 1.

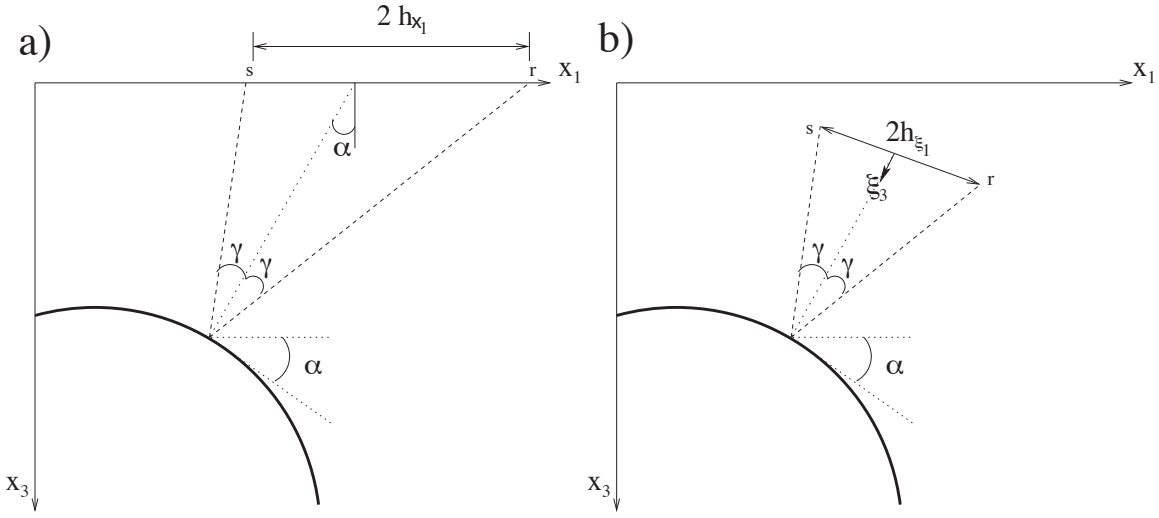


Figure 1: Cartoon illustrating the geometry of the ADCIG calculation. Parameter γ is the reflection opening angle, α is geologic dip. a) Cartesian geometry using coordinates x_1, x_3 andh_{x_1}. b) Generalized geometry using coordinates ξ_1, ξ_3 and h_{ξ_1}. Adapted from **?**. jeff1/. rays

Imaging in generalized coordinate systems follows the same two-step procedure. However, because of the different migration geometry in the $\boldsymbol{\xi}$-coordinate system, new extrapolation operators, $E_{\xi_3}[\cdot]$, must be used to propagate wavefields. I specify these operators using Riemannian wavefield extrapolation (RWE). I do not discuss RWE herein, and refer readers interested in additional information to **?** and **?**.

The first generalized coordinate imaging step is performing wavefield extrapolation

$$\begin{aligned} E_{\xi_3}[S(\xi_3 - \Delta\xi_3, \xi_1|\omega)] &= S(\xi_3, \xi_1|\omega), \\ E^*_{\xi_3}[R(\xi_3 - \Delta\xi_3, \xi_1|\omega)] &= R(\xi_3, \xi_1|\omega), \end{aligned} \tag{4}$$

where $\Delta\xi_3$ is the extrapolation step increment. Generalized coordinate images are then constructed by evaluating an imaging condition

$$I(\xi_3, \xi_1, h_{\xi_1}) = \sum_{\omega} S^*(\xi_3, \xi_1 + h_{\xi_1}|\omega) R(\xi_3, \xi_1 - h_{\xi_1}|\omega), \tag{5}$$

where h_{ξ_1} is the $\boldsymbol{\xi}$-coordinate equivalent of Cartesian subsurface offset axis h_{ξ_1}. The generalized coordinate ADCIG volume is generated by applying an offset-to-angle transformation $\mathbf{T}_{h_\xi \to \gamma}$

$$I(\xi_3, \xi_1, \gamma) = \mathbf{T}_{h_{\xi_1} \to \gamma} I(\xi_3, \xi_1, h_{\xi_1}). \tag{6}$$

Conventional ADCIG volumes can be recovered by sinc interpolating each $I(\xi_3, \xi_1, \gamma)$ image computed via equation 6 to the final Cartesian coordinate volume.

Figure 2 illustrates this process using the elliptic coordinate system. Panel 2a shows the BP synthetic velocity model (**?**) with an elliptic mesh overlain. Note that the salt flanks to the right-side of the model are nearly vertical in Cartesian coordinates. Panel 2b shows the velocity model in panel 2a interpolated to the elliptic coordinate system. Importantly, the aforementioned salt flanks in the elliptic coordinate system are nearly horizontal, which should lead to ADCIG calculations more robust than in Cartesian coordinates. However, proving this assertion requires understanding the differences, if any, between the Cartesian and generalized coordinate offset-to-angle operators, $\mathbf{T}_{h_{x_1} \to \gamma}$ and $\mathbf{T}_{h_{\xi_1} \to \gamma}$, in equations 3 and 6, respectively.

Cartesian coordinate ADCIGs

For constant velocity media in conventional Cartesian geometry, 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 these variables.

Mathematically, these relationships are

$$\begin{bmatrix} \frac{\partial t}{\partial h_{x_1}} \\ \frac{\partial t}{\partial x_3} \end{bmatrix} = 2\, s \cos\alpha \begin{bmatrix} \sin\gamma \\ \cos\gamma \end{bmatrix}, \tag{7}$$

where s is slowness, α is reflector dip, and γ is the reflection opening angle. The right-hand-side of equations 7 are derived by **?**. Equations 7 can be rewritten as

$$-\left.\frac{\partial x_3}{\partial h_{x_1}}\right|_{x_1,t} = \frac{\partial t}{\partial h_{x_1}} \Big/ \frac{\partial t}{\partial x_3} = \tan\gamma, \tag{8}$$

where the negative sign derives from use of the implicit functions theory (**?**). **?** note that Cartesian ADCIGs become pathogenically degenerate in situations where $\frac{\partial t}{\partial x_3} \to 0$ (i.e. for steeply dipping structures where $\alpha \to 90^\circ$ in Figure 1). However, vertically oriented structures are, generally, not well imaged in Cartesian coordinates because of limited steep-angle propagation in downward extrapolation.

Finally, because equation 7 has no explicit geometric-dependence, Fourier-based methods can calculate the reflection opening angle directly in the wavenumber domain

$$\tan\gamma = -\frac{k_{h_{x_1}}}{k_{x_3}}, \tag{9}$$

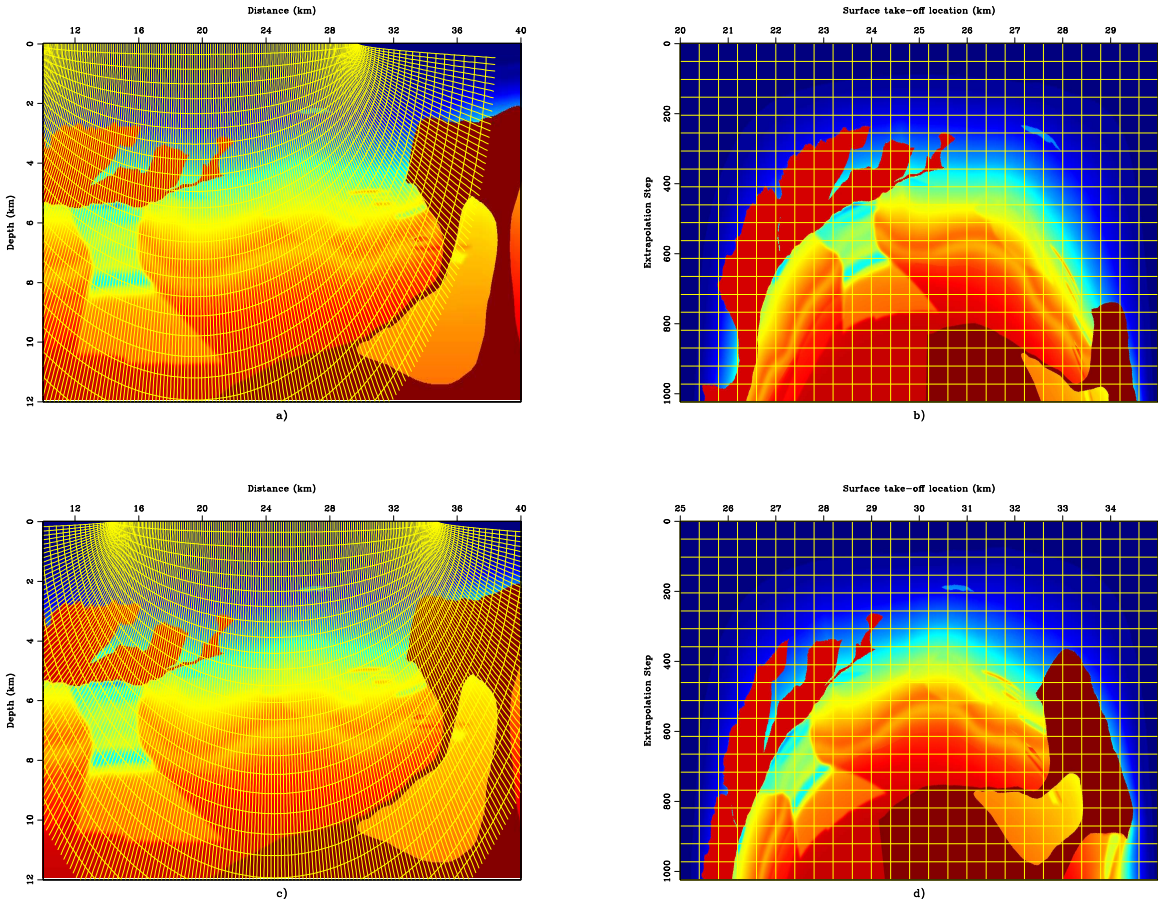


Figure 2: Prestack migration test in elliptic coordinates. a) Benchmark synthetic velocity model with an overlying elliptic coordinate system. b) Effective slowness model in the transformed elliptic coordinate system in a). c) Benchmark synthetic velocity model with a different overlying elliptic coordinate system. d) Effective elliptic coordinate slowness model for the coordinate system in c). jeff1/. RC2

where $k_{h_{x_1}}$ and k_{x_3} are the wavenumbers in the h_{x_1} and x_3 directions, respectively.

Generalized coordinate ADCIGs

Figure 1b illustrates a scenario similar to that in Figure 1a, but for a more general coordinate system. The reflection opening angle, γ, and the reflector dip, α, obviously remain unchanged in the subsurface; 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?

Answering this question requires properly formulating the derivative operators, $\frac{\partial}{\partial x_3}$ and $\frac{\partial}{\partial h_{x_1}}$, in equations 7 in the generalized coordinate system variables $\boldsymbol{\xi} = [\xi_1, \xi_3]$ and $\boldsymbol{h_\xi} = [h_{\xi_1}, h_{\xi_3}]$. Appendix A shows how these derivatives can be specified using Jacobian change-of-variable arguments. Assuming that the subsurface-offset axes are formed by uniform wavefield shifts, Appendix A derives the following expression for generalized coordinate ADCIGs:

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t} = \frac{\partial t}{\partial h_{\xi_1}} \Big/ \frac{\partial t}{\partial \xi_3} = \tan\gamma \, \frac{\left(\frac{\partial x_1}{\partial \xi_1}\cos\alpha - \frac{\partial x_3}{\partial \xi_1}\sin\alpha\right)}{\left(\frac{\partial x_3}{\partial \xi_3}\cos\alpha + \frac{\partial x_1}{\partial \xi_3}\sin\alpha\right)}. \tag{10}$$

Note that if the $\boldsymbol{\xi}$-coordinate system satisfies the Cauchy-Riemann differentiability criteria (**?**)

$$\frac{\partial x_1}{\partial \xi_1} = \frac{\partial x_3}{\partial \xi_3} \quad \text{and} \quad \frac{\partial x_3}{\partial \xi_1} = -\frac{\partial x_1}{\partial \xi_3}, \tag{11}$$

equation 10 then reduces to

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t} = \tan\gamma. \tag{12}$$

This is the generalized coordinate equivalent of the Cartesian expression in equation 7. A physical meaning of the criteria in equations 11 is that the coordinate system must behave isotropically (i.e. dilatationally and rotationally) in the neighborhood of every grid point. Three canonical examples, two of which satisfy equations 11, are discussed in the following section.

Similar to Cartesian coordinates, elliptic coordinate ADCIGs become insensitive where structural dips cause $\frac{\partial t}{\partial \xi_3} \rightarrow 0$. However, this insensitivity can be minimized when using generalized coordinate systems, because structural dips appear at different angles in different translated elliptic meshes. Figures 2c-d illustrate this by showing a different coordinate shift for a different shot-location than that presented in panels 2a-b. Note the changes in structural dip in the right-hand-side of the elliptic coordinate panels. Thus, while ADCIGs calculated on one elliptic grid may be insensitive to certain structure locally, mesh translation ensures that ADCIGs are sensitive globally. Imaging steep dips in elliptic coordinates, though, is limited by the accuracy of wide-angle one-way wavefield extrapolation.

Finally, one may calculate reflection opening angles in the wavenumber domain for coordinate systems satisfying equations 11

$$\tan\gamma = -\frac{k_{h_{\xi_1}}}{k_{\xi_3}}, \tag{13}$$

where $k_{h_{\xi_1}}$ and k_{ξ_3} are the wavenumbers in the h_{ξ_1} and ξ_3 directions, respectively. While some non-orthogonal coordinate systems might satisfy equations 11, most practical applications will have orthogonal $k_{h_{\xi_1}}$ and k_{ξ_3}.

CANONICAL EXAMPLES

This section presents three canonical examples that illustrate the generalized ADCIG theory: tilted Cartesian, polar, and elliptic coordinate systems. Figure 3 presents schematic examples of these three coordinate systems.

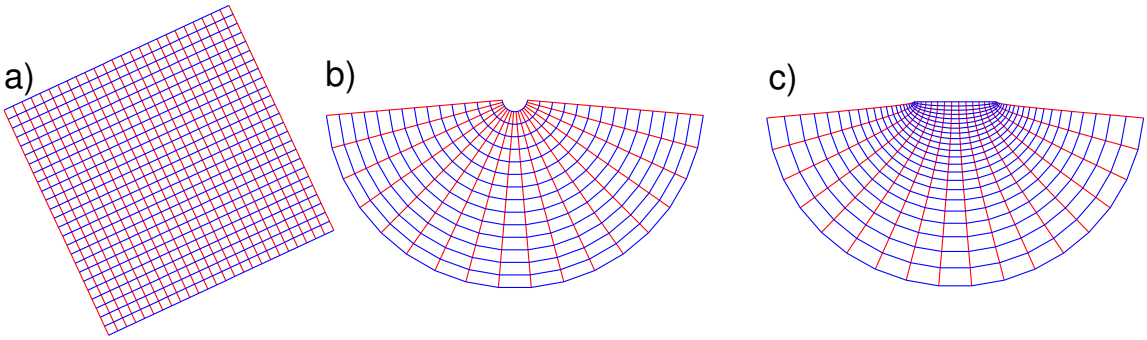


Figure 3: Canonical coordinate system examples. a) Tilted Cartesian coordinates. b) Polar coordinates. c) Elliptic coordinates. jeff1/. COORDS

Tilted Cartesian coordinates

Tilted Cartesian coordinates are a useful generalized migration coordinate system (see Figure 3a). **?** use this mesh in a plane-wave migration scheme where the coordinate system is oriented toward the plane-wave take-off angle to improve large-angle propagation accuracy. A tilted Cartesian mesh is defined by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_3 \end{bmatrix}, \tag{14}$$

where θ is the tilt angle. The partial derivative transform matrix is

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{15}$$

which leads to the following ADCIG equation:

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t} = \tan\gamma \frac{(\cos\theta\cos\alpha + \sin\theta\sin\alpha)}{(\cos\theta\cos\alpha + \sin\theta\sin\alpha)} = \tan\gamma. \tag{16}$$

Thus, calculating ADCIGs in tilted Cartesian coordinates directly recovers the correct reflection opening angle. Note that setting $\theta = 0^\circ$ recovers the Cartesian expression in equation 8.

Polar coordinates

The polar coordinate system (see Figure 3b), where the extrapolation direction is oriented along the radial direction, is appropriate for generating 2D Green's function estimates. The polar coordinate system is defined by

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} a\,\xi_3 \cos\xi_1 \\ a\,\xi_3 \sin\xi_1 \end{bmatrix}. \tag{17}$$

The partial derivative transformation matrix is

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_1} & \frac{\partial x_3}{\partial \xi_3} \end{bmatrix} = \begin{bmatrix} -a\,\xi_3 \sin\xi_1 & a\cos\xi_1 \\ a\,\xi_3 \cos\xi_1 & a\sin\xi_1 \end{bmatrix}, \tag{18}$$

which leads to the following ADCIG equation:

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t} = \xi_3 \tan\gamma\, \frac{(-\sin\xi_1\cos\alpha - \cos\xi_1 \sin\alpha)}{(\sin\xi_1 \cos\alpha + \cos\xi_1 \sin,\alpha)} = -\xi_3 \tan\gamma. \tag{19}$$

Thus, one cannot calculate ADCIGs directly with Fourier-based methods in polar coordinates because of the spatial geometric dependence on ξ_3. However, polar-coordinate ADCIGs can be calculated using slant-stack operators (**?**), because the geometric factor ξ_3 is no more than a local weight applied to the velocity model used to calculate the angle gathers.

Elliptic coordinates

Elliptic coordinates (see Figure 3c) are a useful coordinate system for performing 2D shot-profile migration (**?**). An elliptic mesh is defined by

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

The partial derivative transformation matrix is

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

which leads to the following ADCIG equation:

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t} = \tan\gamma\, \frac{(\cosh\xi_3 \sin\xi_1 \cos\alpha - \sinh\xi_3 \cos\xi_1 \sin\alpha)}{(\cosh\xi_3 \sin\xi_1 \cos\alpha - \sinh\xi_3 \cos\xi_1 \sin\alpha)} = \tan\gamma. \tag{22}$$

Thus, calculating ADCIGs in elliptic coordinates with Fourier-based methods will directly recover the true reflection opening angle.

NUMERICAL EXAMPLES

This section presents numerical tests of the generalized theory by comparing the ADCIG volumes for elliptic and Cartesian coordinate systems. I generated the results using a shot-profile migration algorithm altered to account for elliptic geometry by replacing the Cartesian extrapolator, $E_{x_3}[\cdot]$, with an elliptic coordinate operator, $E_{\xi_3}[\cdot]$. In both coordinate systems, the implemented extrapolation operators were accurate to roughly 80° (Lee and Suh, 1985) with respect to the extrapolation axis. The reader is referred to **?** for further implementation details on shot-profile migration in elliptic coordinates.

I calculated ADCIG image volumes for each shot-profile by following a three-step procedure: 1) generate image volume $I(\xi_1, \xi_3, h_{\xi_1})$ by computing the image for 64 subsurface shifts in h_{ξ_1} at each point in every extrapolation step; 2) calculate ADCIG volume $I(\xi_1, \xi_3, \gamma)$ using the procedure described in **?**; and 3) output the image $I(x_1, x_3, \gamma)$ by sinc-based interpolation of single-shot ADCIGs to the global volume. Steps 1-3 were repeated for all shot-profile sections contributing to the final image.

Test 1: Elliptic Isochrons

The first test demonstrates the impulse response of the ADCIG imaging operator. I do this by imaging the elliptic isochronal responses of two offset point sources in a constant velocity medium (see Figure 4). The source wavefield impulse is at time $t = 0$ s and at $x = -1.12$ km, while the receiver wavefield impulses are at $t = 2.0$, 2.5, 3.0, and 3.5 s and $x = 1.12$ km for a total source-receiver offset of $2h = 2.24$ km. Correlating these two wavefield volumes leads to four elliptic isochrons specified by

$$x_1(t,h) = \frac{vt}{4}\left(\cos\theta_1 + \cos\theta_2\right) \tag{23}$$

$$x_3(t,h) = \frac{vt}{4}\left(1 - \frac{h^2}{v^2t^2}\right)\left(\sin\theta_1 + \sin\theta_2\right), \tag{24}$$

where, given the source take-off angle θ_1, the receiver take-off angle θ_2 is obtained by

$$\theta_2 = \cos^{-1}\left(\frac{vt}{2h}\left[\frac{1}{\frac{2}{1-\frac{4h^2}{v^2t^2}} - \frac{1}{1-\frac{2h\cos\theta_1}{vt}}} - 1\right]\right). \tag{25}$$

The reflection opening angle, given by $\gamma = 90 - \frac{(\theta_1+\theta_2)}{2}$, is shown color-coded on the scatterplot in Figure 4.

Figure 5 shows the ADCIG volumes calculated in both (a single) elliptic and Cartesian coordinate system. To generate this image, I first calculated an elliptic coordinate (EC) volume $I(\xi_1, \xi_3, h_{\xi_1})$ by correlating the source and receiver wavefields at 64 subsurface shifts in h_{ξ_1} at each point in every extrapolation step. I then input the ODCIG volume to a Fourier-based offset-to-angle transformation operator to generate the EC ADCIG volume, $I(\xi_1, \xi_3, \gamma)$, which I interpolated to Cartesian coordinates to generate the desired image volume, $I(x_1, x_3, \gamma)$. For the ADCIG transformation, I choose to limit the maximum opening angle to $\gamma = \pm 60°$. Thus, the 64 image shifts lead to ADCIGs with an angular range between $-60° < \gamma < 60°$ with a sampling increment of $\Delta\gamma = 1.875°$.

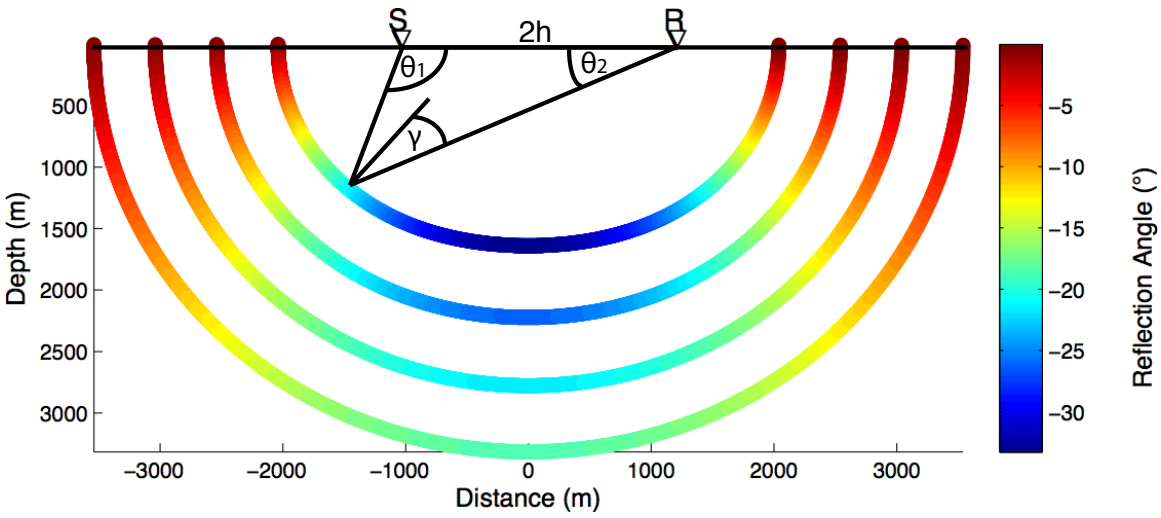


Figure 4: Theoretical results for an elliptic isochron for four travel times in a constant velocity medium. The elliptic surfaces are color-coded according to reflection opening angle. jeff1/. Theory

Panel 5a presents the elliptic coordinate image extracted at $\gamma = -24°$ from the ADCIG volume $I(x_1, x_3, \gamma)$. The ADCIG volumes consist of elliptically shaped reflectors (panels 5a) that are ideally localized in the angle domain (panels 5b-c and f-h). The analytical ADCIG locations are represented by black dots. The analytically and numerically generated results are well matched. Figure 5a also shows three vertical lines indicating the locations from left to right of the three ADCIGs in panels 5b-d. Again, the analytic and numerical ADCIGs are well matched, though less so at shallower depths due to the increased smearing about the image point in the $I(\xi_1, \xi_3, h_{\xi_1})$ domain (**?**).

The Cartesian coordinate (CC) ADCIGs are presented in panels 5e-h. Panel 5e shows the Cartesian image again extracted at $\gamma = -24°$ in the angle domain. Panels 5f-h present three ADCIGs at the same locations as in panels 5b-d. The Cartesian image volumes are well-matched to the elliptic coordinate examples, and good agreement between the theoretical results and the wavefield volume is observed in both images. Energy is focused in the neighborhood of the correct locations. The angle gathers are not always centered relative to the true location, though, which is more noticeable at shallower depths where the Cartesian and elliptic ADCIG volumes both overestimate the reflection opening angle.

Figure 6 presents the results of a test similar to that shown in Figure 5, but with the velocity model rescaled by 0.98. Again, the black dots show the location of the true image point (assuming a true velocity model). Note that the image points in each ADCIG remain well-focused, but shift nearer to the surface and to wider angles. Thus, imaging with an overly slow velocity model will generate, as expected, reflectors that exhibit upward curvature at wider angles.

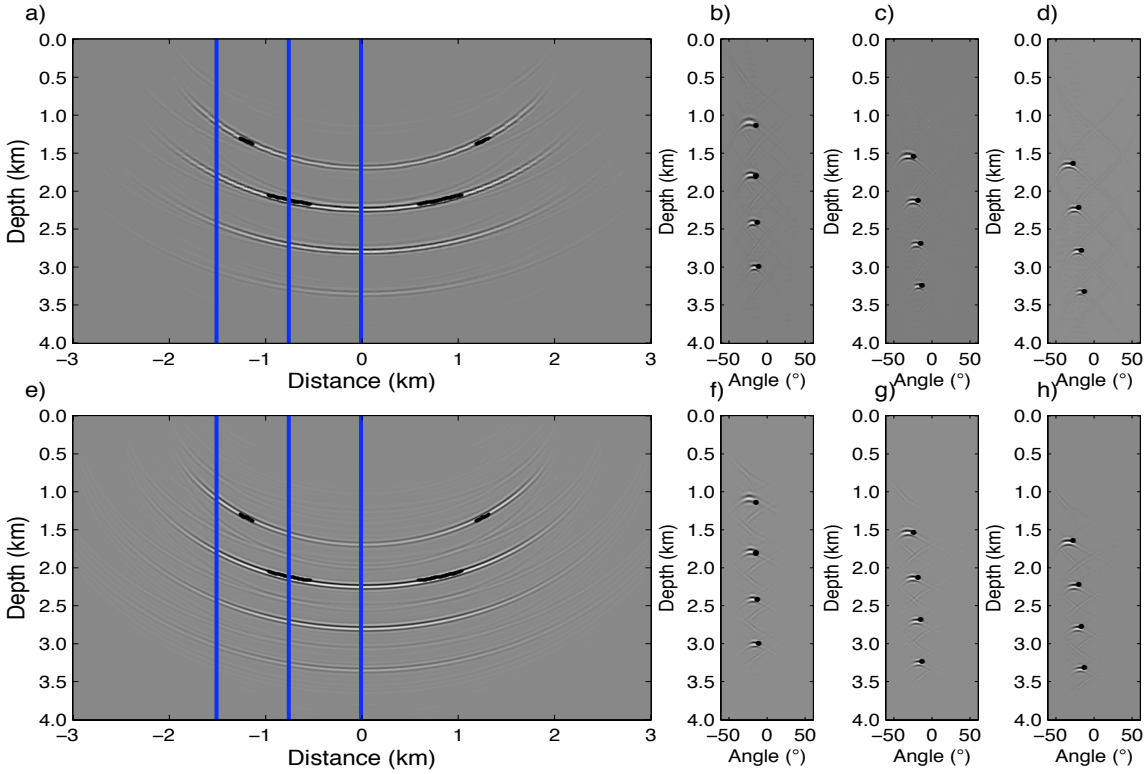


Figure 5: Elliptic reflector comparison tests between analytically (black bullets) and numerically generated ADCIG volumes. Panels a-d are computed in elliptic coordinates (EC), while panels e-h are in Cartesian coordinates (CC). a) EC image extracted at the -24° reflection angle. b) EC angle gather at -1.5 km. c) EC gather at 0.75 km. d) EC gather at 0.0 km. e) CC image extracted at the -24° reflection angle. f) CC angle gather at -1.5 km. g) CC gather at 0.75 km. h) CC gather at 0.0 km. jeff1/. EllipticTest

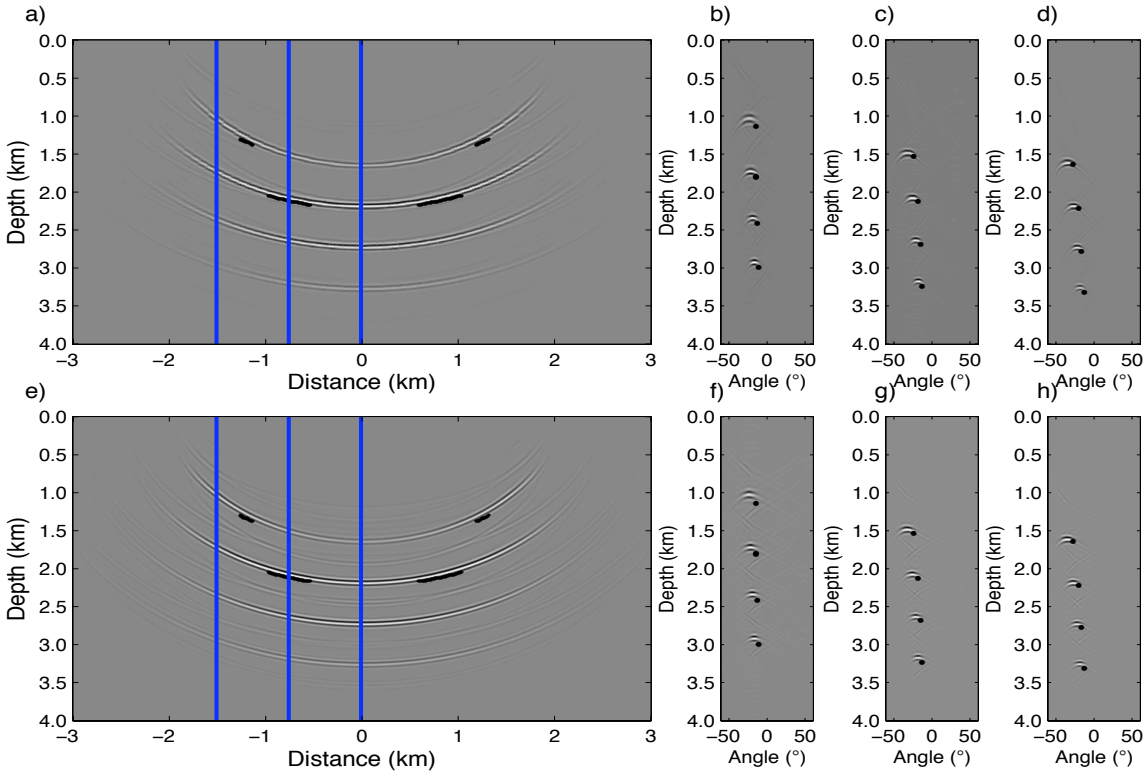


Figure 6: Elliptic reflector comparison tests between analytically (black bullets) and numerically generated ADCIG volumes using a velocity scaled by factor 0.98. Panels a-d are computed in elliptic coordinates (EC), while panels e-h are in Cartesian coordinates (CC). a) EC image extracted at the -24° reflection angle. b) EC angle gather at -1.5 km. c) EC gather at 0.75 km. d) EC gather at 0.0 km. e) CC image extracted at the -24° reflection angle. f) CC angle gather at -1.5 km. g) CC gather at 0.75 km. h) CC gather at 0.0 km. jeff1/. WrongEllipticTest

Test 2: BP velocity model

The second test compares elliptic and Cartesian coordinate ADCIG volumes computed for the BP synthetic velocity model. Images computed in elliptic coordinates used only one coordinate system per shot. For the one-sided data set, I used a (surface) migration aperture of 12km and located the source and farthest offset receiver (at 8km) points 2km in from the edges of the computational mesh. (Note that the migration aperture effectively expands during wavefield extrapolation because the coordinate mesh expands outward.) I found this initial migration geometry to produce the best results for the BP synthetic model through iterative testing. Generally, the optimal elliptic coordinate migration geometry is controlled by the velocity model.

Figure 7 shows slices all clipped at the 99^{th} percentile from the corresponding elliptic and Cartesian ADCIG image volumes. Panel 7a shows an elliptic coordinate image with

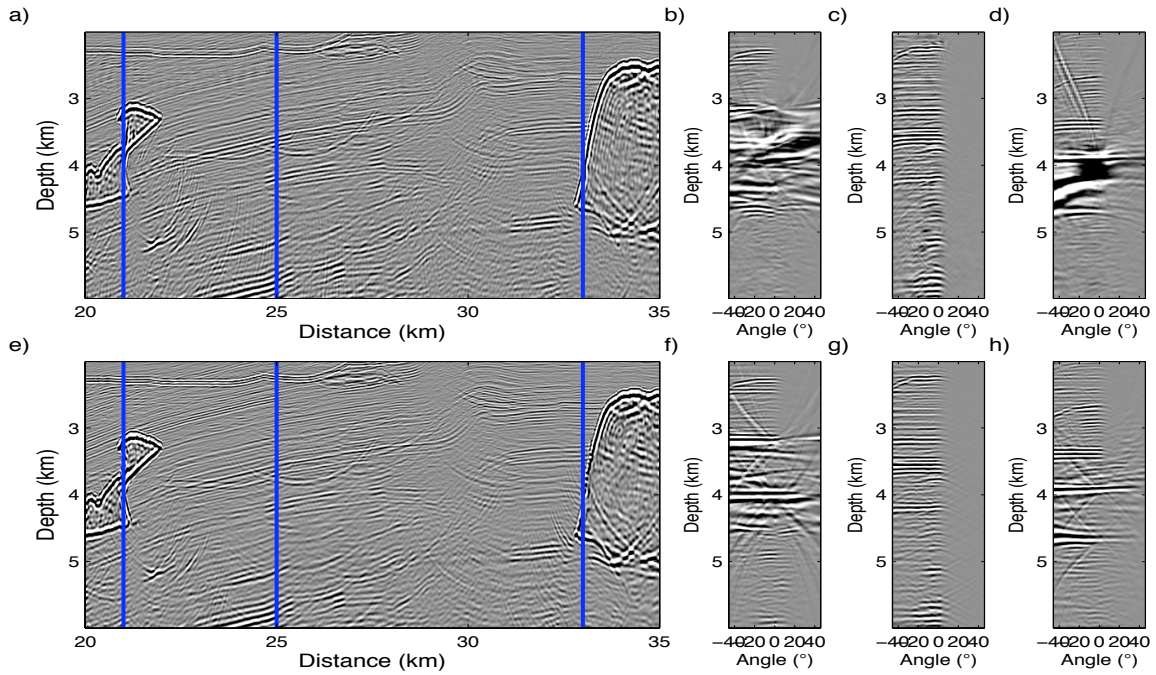


Figure 7: Vertical elliptic and Cartesian ADCIGs slices using the correct migration velocity model. a) Elliptic coordinate image with three vertical lines showing the locations of ADCIG gathers from left to right in panels b-d. e) Cartesian coordinate image with three vertical lines showing the locations of ADCIG gathers from left to right in panels f-h. jeff1/. GOOD

three vertical lines indicating the angle-gather locations from left to right in Figures 7b-d. The three panels show predominantly one-sided reflectivity, which is to be expected because the input migration data were not in a split-spread geometry. (This statement holds for all subsequent images calculated using this data set.) The only significant exceptions occur in panel 7b within salt where energy is exhibited for both positive and negative reflection angles. I attribute this to the reversal of source and receiver wavefield orientations within the salt.

The image in panel 7d has a wide reflection zone between 3.75-4.25 km in depth, which

occurs because the shown angle gather is a vertical slice through the nearly vertical salt flank. This creates the appears of low-frequency noise, which is the appropriate response for a near-vertical reflector. Panel 7e shows the Cartesian image for the same location as panel 7a, while panels 7f-h are extracted from the same locations as panels 7b-d. The Cartesian angle gathers look similar to those in elliptic coordinates, except for the salt flanks to the right-hand-side of panel 7h.

A final observation from Figure 7 is that ADCIGs calculated via subsurface correlations will generate artifacts at locations near salt-sediment interfaces - whether in an elliptic or a Cartesian coordinate system. This geologic setting leads to situations where a wavefield sample inside a salt body is correlated with another sample located in the sediment with a significantly different velocity. This velocity difference violates one of the theoretical ADCIG assumptions, namely that the velocity remains constant across the correlation window. Hence, one must be careful not to interpret ADCIG artifacts as signal useful for migration velocity analysis.

Figure 8 shows horizontal slices that better resolve the vertical salt flank. Panel 8a presents the elliptic coordinate image, with three horizontal lines showing the ADCIG slice locations from top to bottom. The right-hand sides of panels 8b-d display the well-focused vertical salt-flank reflector. This demonstrates the robustness of the ADCIG calculation in elliptic coordinates. Panel 8e shows the Cartesian coordinate image with three horizontal lines showing the locations of the ADCIG slices. The right-hand salt-flank reflector in panel 8f is similarly well-resolved, largely because the structural dip is relatively low. However, the salt-flank images in panels 8g-h are somewhat blurred out. I attribute this to the combined effects of inaccurate large-angle extrapolation and insensitivity of the ADCIG calculation to steep structural dip.

An additional test examines how the ADCIG volumes change when introducing an incorrect migration velocity profile. Figure 9 presents ADCIG volumes similar to those shown in Figure 7 after using a migration velocity profile rescaled by 98%. Both images are poorly focused and have residual curvature indicating an incorrect migration velocity. Because the reflectors are near vertical, though, the sensitivity of horizontal gathers is weak. This low sensitivity is greatly improved when examining the horizontal slices in Figure 10 taken at the same locations as in Figure 8. The elliptic angle gathers in panels 10b-d, and especially to the right-hand side in panel 10d, show much greater residual curvature. This indicates that the elliptic coordinate horizontal ADCIGs have greater sensitivity to velocity error for near-vertical structures than Cartesian coordinate horizontal ADCIGs. The imaging enhancements afforded by elliptic coordinates should improve any migration velocity analysis approach that uses residual curvature in steeply dipping reflectors to compute velocity model updates.

Discussion

Extending the above theory of generalized coordinate ADCIGs to 3D coordinate systems is fairly straightforward, though more difficult to implement numerically. **?** presents a theory for 3D Cartesian coordinates that specifies the differential travel-time expressions required to express the reflection opening angle, γ, in 3D Cartesian ADCIGs [see equation 16 in **?**]. Applying Jacobian change-of-variable transformations to these equations should yield a 3D

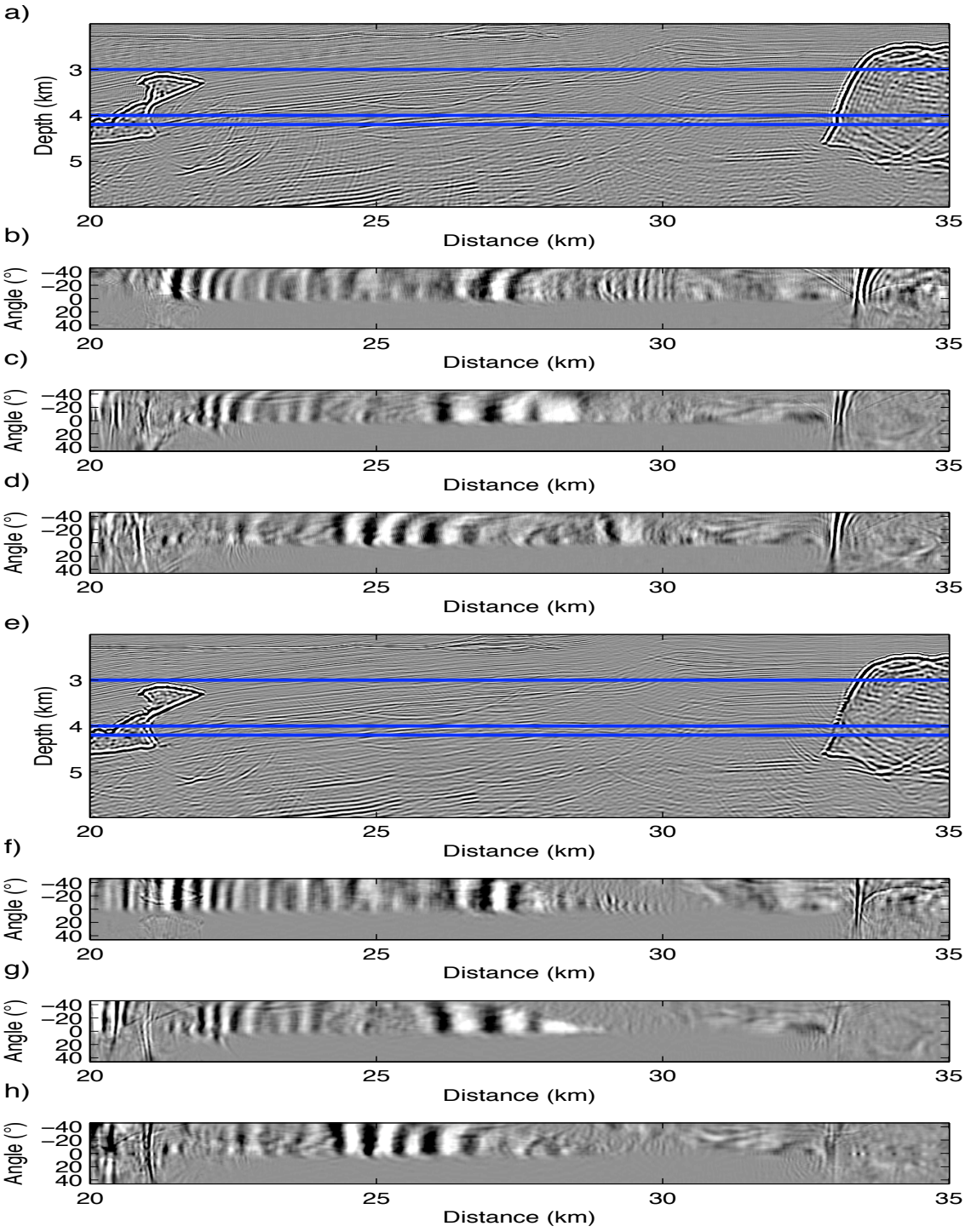


Figure 8: Horizontal elliptic and Cartesian ADCIGs slices using the correct migration velocity model. a) Elliptic coordinate image with three horizontal lines showing the locations of horizontal ADCIG gathers from top to bottom in panels b-d. e) Cartesian coordinate image with three horizontal lines showing the locations of horizontal ADCIG gathers from top to bottom in panels f-h. jeff1/. GOODZ

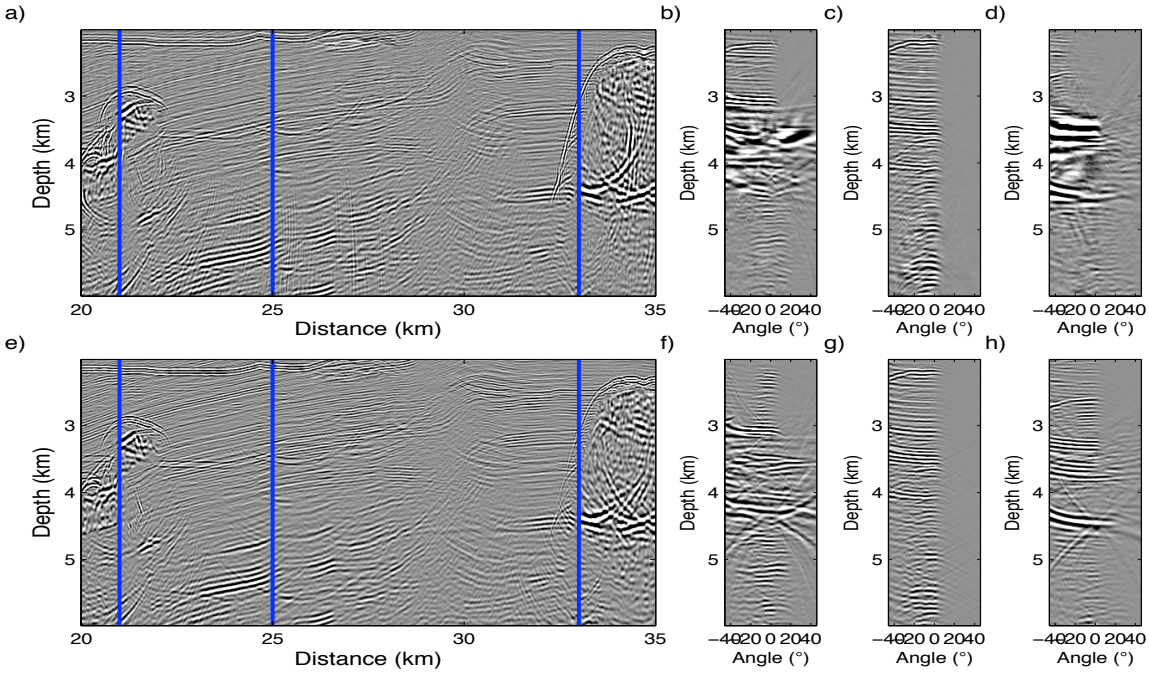


Figure 9: Vertical elliptic and Cartesian ADCIGs slices using an incorrect migration velocity model. a) Elliptic coordinate image with three vertical lines showing the locations of vertical ADCIG gathers from left to right in panels b-d. e) Cartesian coordinate image with three vertical lines showing the locations of vertical ADCIG gathers from left to right in panels f-h. jeff1/. BAD

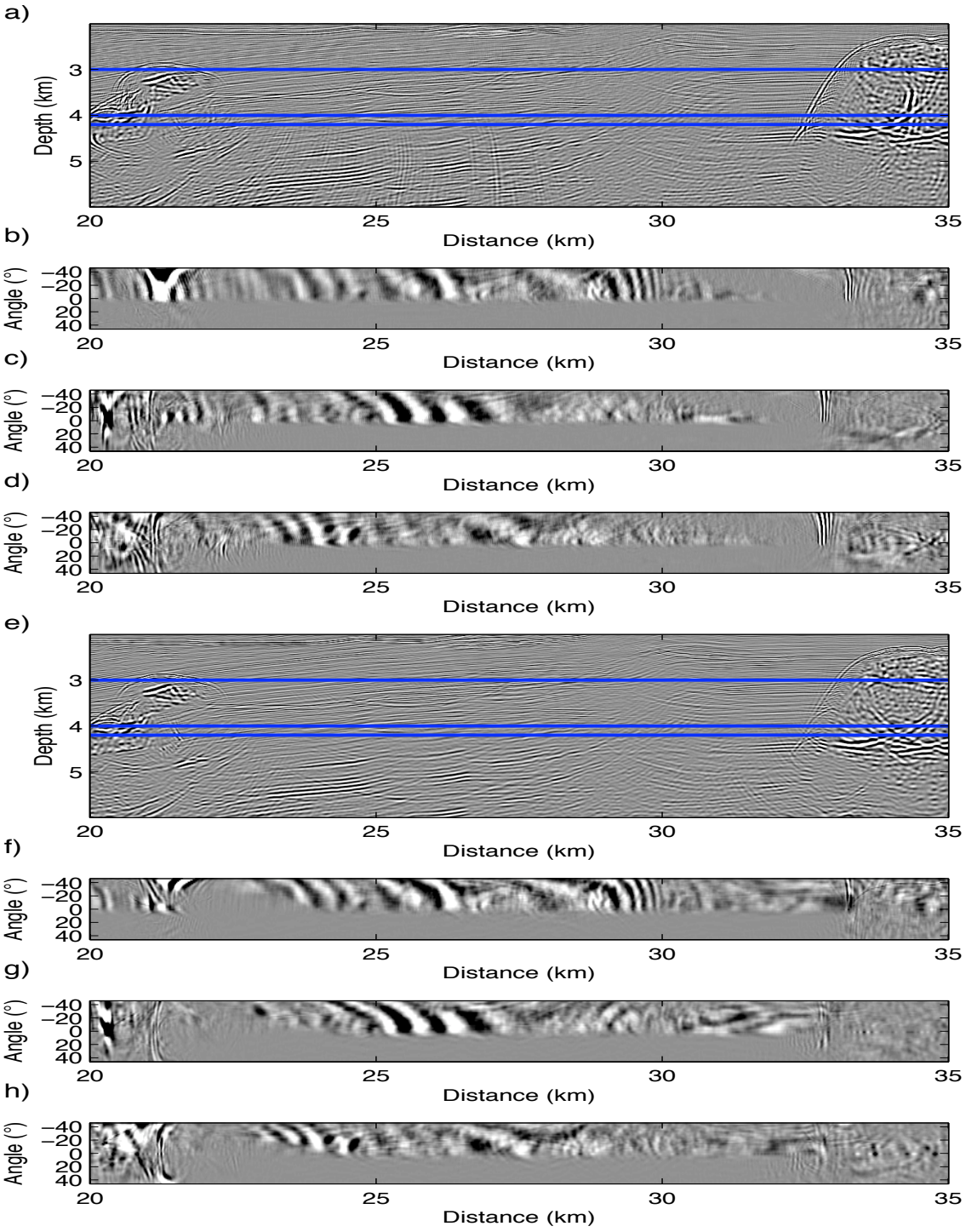


Figure 10: Horizontal elliptic and Cartesian ADCIGs slices using an incorrect migration velocity model. a) Elliptic coordinate image with three horizontal lines showing the locations of horizontal ADCIG gathers from top to bottom in panels b-d. e) Cartesian coordinate image with three horizontal lines showing the locations of horizontal ADCIG gathers from top to bottom in panels f-h. jeff1/. BADZ

expression for reflection angle. Similar to 3D Cartesian coordinates, though, this quantity will depend on geologic dips and need to be computed by one of the two algorithms suggested by **?**.

Given that a 3D expression can be formulated, there are a number of coordinate systems well-suited to imaging steep geologic dips where 3D ADCIG volumes could be a good diagnostic tool for velocity analysis. **?** discusses how a judicious choice of 3D coordinate system depends greatly on the acquisition geometry and the desired migration geometry. For example, the migration geometries employed in shot-profile migration of wide-azimuth data sets are well-matched with 3D ellipsoidal meshes that enable high-angle and turning-wave propagation in all directions. Evaluating ADCIG image focussing in such a 3D coordinate geometry would then provide information on velocity model accuracy for steeply dipping reflectors - such as salt flanks. These somewhat speculative extensions, though, are beyond the scope of this paper and remain an active area of research.

CONCLUSIONS

I extend the Cartesian ADCIG theory to 2D generalized coordinate systems. The generalized ADCIG expressions related the reflection opening angle to differential traveltime operators and spatially varying weights derived from the non-Cartesian geometry. I show that these geometric expressions cancel out for coordinate systems satisfying the Cauchy-Riemann differentiability criteria, which include tilted Cartesian and elliptic meshes. The procedure for calculating ADCIGs in elliptic coordinates is very similar to that in Cartesian coordinates. I validate the approach by comparing analytically and numerically generated ADCIG volumes, and with tests on the BP synthetic data set. ADCIGs calculations are more robust where computed in elliptic coordinates than in Cartesian coordinate. I assert that this result is due to improved large-angle propagation and enhanced sensitivity to steep structural dips afforded by the coordinate transforms. Finally, the imaging advantages afforded by elliptic coordinates should improve the procedure of any migration velocity analysis approach that uses residual ADCIG curvature on steeply dipping reflectors to compute velocity model updates.

ACKNOWLEDGMENTS

I would like to thank Biondo Biondi, Alejandro Valenciano, Antoine Guitton and Brad Artman for helpful discussions, and BP for releasing the synthetic data used in the examples. I thank Tamas Nemeth, Samuel Gray, and an anonymous reviewer for helpful comments on the early stages of this manuscript. I acknowledge the sponsors of the SEP consortium for their continuing support.

REFERENCES

Lee, M. W. and S. Y. Suh, 1985, Optimization of one-way wave-equations (short note): Geophysics, **50**, 1634–1637.

APPENDIX A

ADCIG COORDINATE TRANSFORM

This appendix addresses how to express operators $\frac{\partial}{\partial x_3}$ and $\frac{\partial}{\partial h_{x_1}}$ in generalized coordinate systems to derive equation 10. I first assume that generalized coordinate systems are related to the Cartesian variables through a bijection (i.e., one-to-one mapping)

$$x_1 = f(\xi_1, \xi_3) \quad \text{and} \quad x_3 = g(\xi_1, \xi_3) \tag{A-1}$$

with a non-vanishing Jacobian of coordinate transformation, $J_{\boldsymbol{\xi}}$. The bijection between a generalized and Cartesian coordinate system allows us to rewrite the left-hand-sides of equations 7 as (**?**)

$$\frac{\partial t}{\partial x_1} = \frac{1}{J_{\boldsymbol{\xi}}} \frac{\partial(t, x_3)}{\partial(\xi_1, \xi_3)} \quad \text{and} \quad \frac{\partial t}{\partial x_3} = \frac{1}{J_{\boldsymbol{\xi}}} \frac{\partial(x_1, t)}{\partial(\xi_1, \xi_3)}. \tag{A-2}$$

Expanding the Jacobian notation leads to

$$\begin{bmatrix} \frac{\partial t}{\partial \xi_1}\frac{\partial x_3}{\partial \xi_3} - \frac{\partial t}{\partial \xi_3}\frac{\partial x_3}{\partial \xi_1} \\ \frac{\partial t}{\partial \xi_3}\frac{\partial x_1}{\partial \xi_1} - \frac{\partial t}{\partial \xi_1}\frac{\partial x_1}{\partial \xi_3} \end{bmatrix} = 2\, J_{\boldsymbol{\xi}}\, s \cos\gamma \begin{bmatrix} \sin\alpha \\ \cos\alpha \end{bmatrix}. \tag{A-3}$$

The right-hand-sides of equations A-3 are analogous to those derived by **?**. Cross-multiplying the expressions by factors $\frac{\partial x_1}{\partial \xi_3}$ and $\frac{\partial x_3}{\partial \xi_3}$

$$\begin{bmatrix} \frac{\partial x_1}{\partial \xi_3}\left(\frac{\partial t}{\partial \xi_1}\frac{\partial x_3}{\partial \xi_3} - \frac{\partial t}{\partial \xi_3}\frac{\partial x_3}{\partial \xi_1}\right) \\ \frac{\partial x_3}{\partial \xi_3}\left(\frac{\partial t}{\partial \xi_3}\frac{\partial x_1}{\partial \xi_1} - \frac{\partial t}{\partial \xi_1}\frac{\partial x_1}{\partial \xi_3}\right) \end{bmatrix} = 2\, J_{\boldsymbol{\xi}}\, s \cos\gamma \begin{bmatrix} \frac{\partial x_1}{\partial \xi_3}\sin\alpha \\ \frac{\partial x_3}{\partial \xi_3}\cos\alpha \end{bmatrix} \tag{A-4}$$

and adding the two expressions results in

$$\frac{\partial t}{\partial \xi_3}\left(\frac{\partial x_3}{\partial \xi_3}\frac{\partial x_1}{\partial \xi_1} - \frac{\partial x_1}{\partial \xi_3}\frac{\partial x_3}{\partial \xi_1}\right) = 2\, J_{\boldsymbol{\xi}}\, s \cos\gamma \left(\frac{\partial x_1}{\partial \xi_3}\sin\alpha + \frac{\partial x_3}{\partial \xi_3}\cos\alpha\right). \tag{A-5}$$

A similar argument can be used to construct the equations for the subsurface-offset axis. The bijection between the generalized coordinate and Cartesian subsurface-offset axes allows for the left-hand-side of equations 7 to be rewritten as

$$\frac{\partial t}{\partial h_{x_1}} = \frac{1}{J_{\boldsymbol{h}}} \frac{\partial(t, h_{x_3})}{\partial(h_{\xi_1}, h_{\xi_3})} \quad \text{and} \quad \frac{\partial t}{\partial h_{x_3}} = \frac{1}{J_{\boldsymbol{h}}} \frac{\partial(h_{x_1}, t)}{\partial(h_{\xi_1}, h_{\xi_3})}, \tag{A-6}$$

where $J_{\boldsymbol{h}}$ is the subsurface-offset Jacobian of transformation. Expanding the Jacobian notation leads to

$$\begin{bmatrix} \frac{\partial t}{\partial h_{\xi_1}}\frac{\partial h_{x_3}}{\partial h_{\xi_3}} - \frac{\partial t}{\partial h_{\xi_3}}\frac{\partial h_{x_3}}{\partial h_{\xi_1}} \\ \frac{\partial t}{\partial h_{\xi_3}}\frac{\partial h_{x_1}}{\partial h_{\xi_1}} - \frac{\partial t}{\partial h_{\xi_1}}\frac{\partial h_{x_1}}{\partial h_{\xi_3}} \end{bmatrix} = 2\, J_{\boldsymbol{h}}\, s \sin\gamma \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}. \tag{A-7}$$

The right-hand-side of equations A-7 are again analogous to those given by **?**. Cross-multiplying the expressions by factors $\frac{\partial h_{x_1}}{\partial h_{\xi_1}}$ and $\frac{\partial h_{x_3}}{\partial h_{\xi_1}}$

$$\begin{bmatrix} \frac{\partial h_{x_1}}{\partial h_{\xi_1}}\left(\frac{\partial t}{\partial h_{\xi_1}}\frac{\partial h_{x_3}}{\partial h_{\xi_3}} - \frac{\partial t}{\partial h_{\xi_3}}\frac{\partial h_{x_3}}{\partial h_{\xi_1}}\right) \\ \frac{\partial h_{x_3}}{\partial h_{\xi_1}}\left(\frac{\partial t}{\partial h_{\xi_3}}\frac{\partial h_{x_1}}{\partial h_{\xi_1}} - \frac{\partial t}{\partial h_{\xi_1}}\frac{\partial h_{x_1}}{\partial h_{\xi_3}}\right) \end{bmatrix} = 2\, J_{\boldsymbol{h}}\, s \sin\gamma \begin{bmatrix} \frac{\partial h_{x_1}}{\partial h_{\xi_1}}\cos\alpha \\ \frac{\partial h_{x_3}}{\partial h_{\xi_1}}\sin\alpha \end{bmatrix}, \tag{A-8}$$

and subtracting the two expressions above yields

$$\frac{\partial t}{\partial h_{\xi_1}}\left(\frac{\partial h_{x_1}}{\partial h_{\xi_1}}\frac{\partial h_{x_3}}{\partial h_{\xi_3}}-\frac{\partial h_{x_1}}{\partial h_{\xi_3}}\frac{\partial h_{x_3}}{\partial h_{\xi_1}}\right)=2\,J_{\boldsymbol{h}}\,s\sin\gamma\left(\frac{\partial h_{x_1}}{\partial h_{\xi_1}}\cos\alpha-\frac{\partial h_{x_3}}{\partial h_{\xi_1}}\sin\alpha\right). \tag{A-9}$$

An expression for ADCIGs can be obtained by dividing equation A-9 by equation A-5

$$\frac{\frac{\partial t}{\partial h_{\xi_1}}}{\frac{\partial t}{\partial \xi_3}}\frac{\left(\frac{\partial h_{x_1}}{\partial h_{\xi_1}}\frac{\partial h_{x_3}}{\partial h_{\xi_3}}-\frac{\partial h_{x_1}}{\partial h_{\xi_3}}\frac{\partial h_{x_3}}{\partial h_{\xi_1}}\right)}{\left(\frac{\partial x_3}{\partial \xi_3}\frac{\partial x_1}{\partial \xi_1}-\frac{\partial x_1}{\partial \xi_3}\frac{\partial x_3}{\partial \xi_1}\right)}=\tan\gamma\,\frac{J_{\boldsymbol{h}}}{J_{\boldsymbol{\xi}}}\,\frac{\left(\frac{\partial h_{x_1}}{\partial h_{\xi_1}}\cos\alpha-\frac{\partial h_{x_3}}{\partial h_{\xi_1}}\sin\alpha\right)}{\left(\frac{\partial x_3}{\partial \xi_3}\cos\alpha+\frac{\partial x_1}{\partial \xi_3}\sin\alpha\right)}. \tag{A-10}$$

One question arising from the geometric factors in equation A-10 is what do the terms $\frac{\partial h_{x_1}}{\partial h_{\xi_1}}$, $\frac{\partial h_{x_3}}{\partial h_{\xi_1}}$, $\frac{\partial h_{x_1}}{\partial h_{\xi_3}}$ and $\frac{\partial h_{x_3}}{\partial h_{\xi_3}}$ represent? I assume that the subsurface offset axes are generated by uniform wavefield shifting such that the following equations are valid:

$$\begin{bmatrix} h_{x_1} \\ h_{x_3} \\ h_{\xi_1} \\ h_{\xi_1} \end{bmatrix}=\begin{bmatrix} x_1 \\ x_3 \\ \xi_1 \\ \xi_3 \end{bmatrix}\quad \text{such that}\quad \begin{bmatrix} \frac{\partial h_{x_1}}{\partial h_{\xi_1}} \\ \frac{\partial h_{x_3}}{\partial h_{\xi_1}} \\ \frac{\partial h_{x_1}}{\partial h_{\xi_3}} \\ \frac{\partial h_{x_3}}{\partial h_{\xi_3}} \end{bmatrix}=\begin{bmatrix} \frac{\partial x_1}{\partial \xi_1} \\ \frac{\partial x_3}{\partial \xi_1} \\ \frac{\partial x_1}{\partial \xi_3} \\ \frac{\partial x_3}{\partial \xi_3} \end{bmatrix}. \tag{A-11}$$

If the subsurface offset axes were generated by anything other than uniform shifting (e.g. $h_{x_1}=x_1^2$), then the assumptions behind equations A-11 would not be honored.

Using these identities in equation A-5 reduces equation A-10 to

$$-\left.\frac{\partial \xi_3}{\partial h_{\xi_1}}\right|_{\xi_1,t}=\frac{\partial t}{\partial h_{\xi_1}}\Big/\frac{\partial t}{\partial \xi_3}=\tan\gamma\,\frac{\left(\frac{\partial x_1}{\partial \xi_1}\cos\alpha-\frac{\partial x_3}{\partial \xi_1}\sin\alpha\right)}{\left(\frac{\partial x_3}{\partial \xi_3}\cos\alpha+\frac{\partial x_1}{\partial \xi_3}\sin\alpha\right)}, \tag{A-12}$$

where the two Jacobian transformations are equivalent (i.e. $J_{\boldsymbol{\xi}}=J_{\boldsymbol{h}}$). This completes the derivation of equation 10.

Seismic interferometry versus spatial auto-correlation method on the regional coda of the NPE

Sjoerd de Ridder

ABSTRACT

A seismic recording of the non-proliferation experiment (NPE) contains the first break of the regional P phases followed by a three minute long coda. The frequency-domain result of seismic interferometry is studied. This procedure is analogous to the spatial auto-correlation (SPAC) method, devised for studying microtremors by Aki (1957). Cross-correlating two receiver stations retrieves, under favorable circumstances, an approximation of the Green's function between these two stations. To first order, this Green's function consists of a direct event traveling between the receivers. In the frequency-domain, the lowest mode in the Green's function is a weighted and scaled zero-order Bessel function of the first kind, J_0. The cross-spectrum from the coda of the NPE is estimated using multitaper spectral analysis. The retrieved Green's functions are fitted to damped J_0 functions to recover phase velocity and estimates of the attenuation coefficients. Only energy between 1-4 Hz can be fitted unambiguously with J_0 functions, because higher frequencies contain too much spurious energy. This result shows the equivalence of the SPAC method and seismic interferometry for the lowest mode in the Green's function. This study also demonstrates that the coda of a regional event, seemingly unfavorably positioned, can contain energy useful for seismic interferometry.

INTRODUCTION

In an efort to ban all nuclear tests, a large majority at the United Nations General Assembly in New York passed the comprehensive test ban treaty (CTBT) in September, 1996. Although 180 states signed the treaty, only 145 ratified it, prohibiting the treaty from entering into force (CTBTO, 2008). The US Department of Energy detonated a 1.5 kiloton chemical explosive charge at the Nevada test site on September 22$^{\text{nd}}$, 1993. The experiment, named the non-proliferation experiment (NPE), was conducted in anticipation of the CTBT. The explosion was recorded by over 50 broadband seismic stations in 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 well known recording was made by the Subsurface Exploration Company of Pasadena, CA, which operated a 610-channel petroleum-exploration seismic array approximately 200 km distant. The array was oriented east-west in Railroad Valley, Nevada, which is located north of the Nevada Test Site, see Figure 1. This sign-bit equipment was activated at midnight and recorded an extraordinary coda over 5 minutes in length. The incoming waves are spatially coherent at early times, when all energy comes in as a single plane wave (see Figure 2a). Very rapidly after the first break, the wavefield becomes more

chaotic, and at later arrivals the higher frequencies are lost, as observed in Figures 2b - 2d. At later times, recorded arrivals are incident for a large range of apparent slownesses. This phenomenon was studied by de Ridder (2008), who concluded that all admissible slownesses define a cone in the frequency-wavenumber domain, as seen in Figure 3. The slope of the cone is determined by the event with the slowest possible apparent velocity; a surface-wave traveling purely along the array.

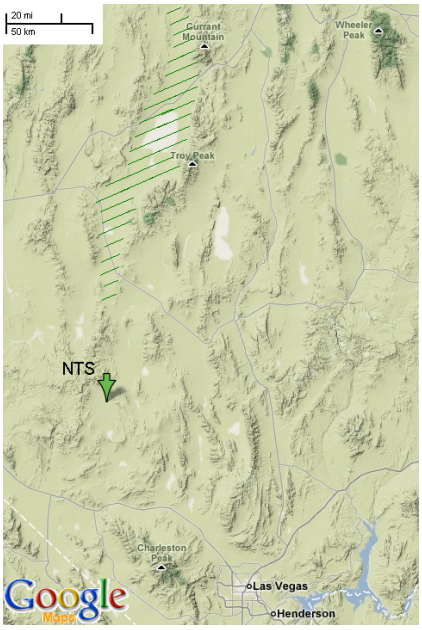


Figure 1: Map showing the location of the Nevada Test Site (arrow) and Railroad Valley (shaded area). [**NR**] sjoerd1/. map

In this report, the coda of the NPE is studied using cross-correlation techniques. Two approaches are contrasted: the seismic interferometry (SI) method which potentially retrieves the full impulse response of the earth from the recorded background field (Wapenaar, 2004) and the spatial auto-correlation (SPAC) method, which retrieves dispersion curves of the fundamental mode of the surface-waves in a horizontally layered medium (Aki, 1957).

SEISMIC INTERFEROMETRY AND THE SPATIAL AUTO-CORRELATION METHOD

Seismic interferometry refers to the principle of generating new seismic responses through cross-correlations of recorded seismic wavefields at receivers (Schuster, 2001). The first to derive this principle for deterministic wavefields in 1D media was Claerbout (1968), who showed that the reflection response of a 1D medium could be synthesized from the auto-correlation of the transmission response. Later derivations include many different approaches based upon: diffusivity of the wavefields (Weaver and Lobkis, 2001; Roux et al., 2005; Sánchez-Sesma et al., 2006; Sánchez-Sesma and Campillo, 2006), stationary phase analysis (Schuster et al., 2004; Snieder, 2004) and propagation invariants and reciprocity theorems (Claerbout, 1976; Weaver and Lobkis, 2004; Wapenaar, 2004; Wapenaar and Fokkema, 2006; van Manen et al., 2005).

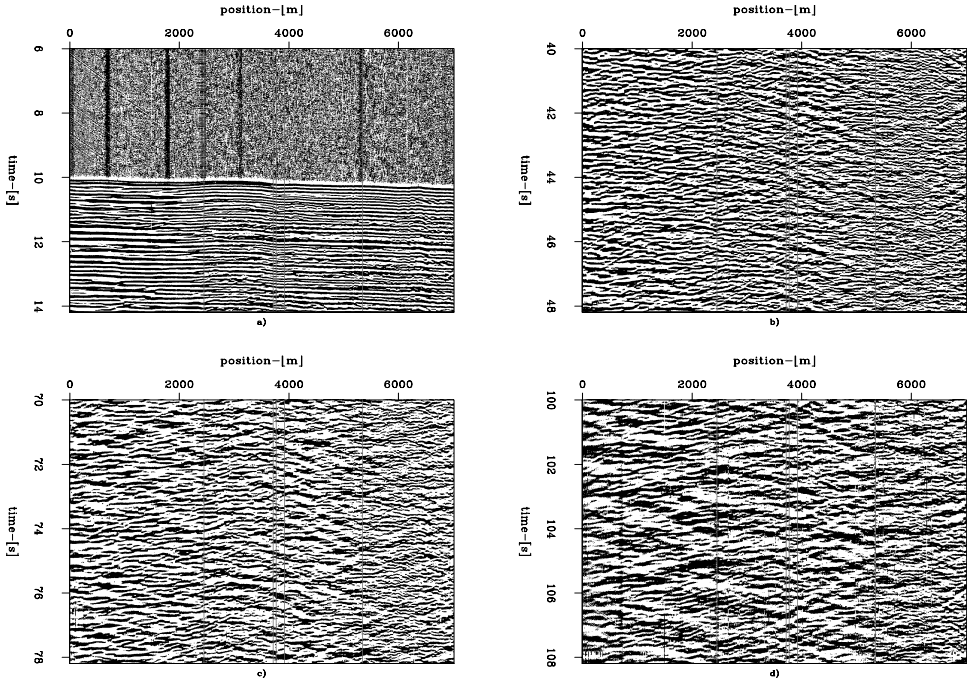


Figure 2: Four segments of the first NPE recording; a) segment containing the first arrival; b) segment 30 seconds after the first arrival; c) segment 60 seconds after the first arrival; d) segment 90 seconds after the first arrival. [**ER**] sjoerd1/. shot1frames

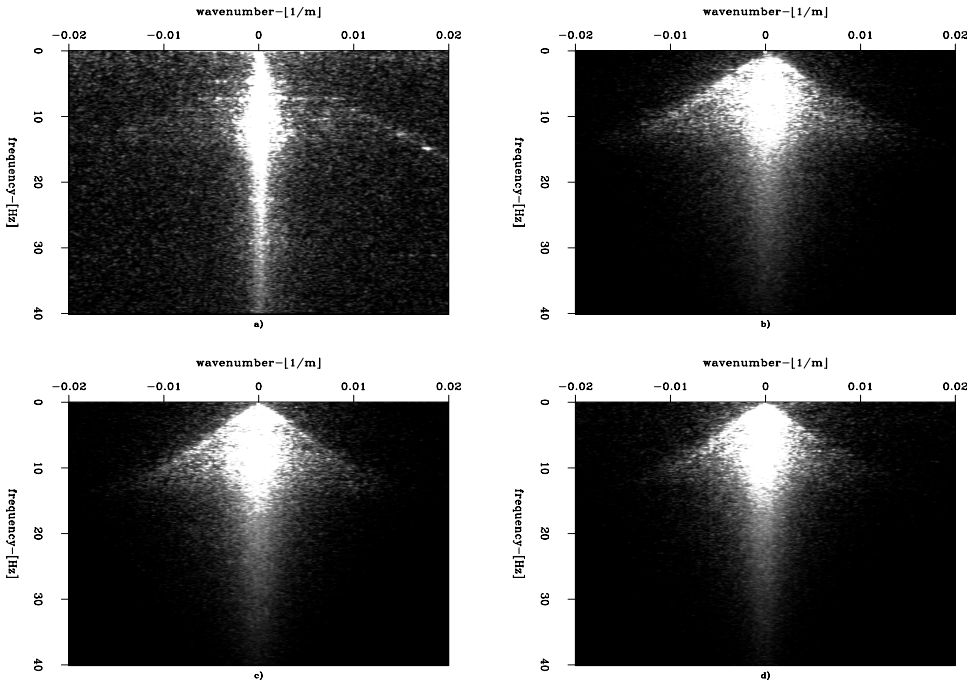


Figure 3: Frequency-wavenumber spectra of the NPE recording, subfigures a, b, c and d display each the frequency-wavenumber spectra for the corresponding subfigures in Figure 2. [**ER**] sjoerd1/. shot1WKframes

Wapenaar and Fokkema (2006) derives an interferometric representation of the Green's function from energy principles. Under certain conditions, this representation can be simplified to a direct cross-correlation between two receiver stations. Consider a domain $\mathbf{D}$ in an arbitrary, inhomogeneous medium enclosing two points $\mathbf{x}_A$ and $\mathbf{x}_B$, bounded by an arbitrarily shaped surface $\partial\mathbf{D}$ with outward pointing normal vector $\mathbf{n}$. The interferometric representation of the Green's function for the vertical component of particle velocity measured at $\mathbf{x}_A$ in response to a vertical force-impulse point-source acting at $\mathbf{x}_B$ is given in the frequency-domain as follows (Wapenaar and Fokkema, 2006):

$$2\Re\{\hat{G}^{\upsilon,f}_{3,3}(\mathbf{x}_A,\mathbf{x}_B,\omega)\} = \tag{1}$$
$$-\oint_{\partial\mathbf{D}}\left[\hat{G}^{\upsilon,h}_{3,ij}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^* + \hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,h}_{3,ij}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^*\right] n_j \mathrm{d}^2\mathbf{x},$$

where the asterisk denotes complex conjugation, and ω denotes angular frequency. The notation convention for Green's functions is that superscripts denote the receiver (first) and source type (second), and subscripts denote the components of the source (first) and receiver (second) fields. The fields and sources in the elastodynamic system are particle velocity $\boldsymbol{\upsilon}$, stress tensor $\boldsymbol{\tau}$ (used below), external volume force density $\mathbf{f}$, and external deformation rate density $\mathbf{h}$. Einstein's summation convention is applied on all repeated subscripts.

The interferometric integral in equation 1 represents the real part of the elastodynamic Green's function between two receiver stations located at A and B, as a summation of cross-correlations of independent measurements at the two receiver stations. (Independent measurements of responses of various source components and types located on a surface enclosing both receivers are required to evaluate this integral.) The integral can be modified to reflect the field configuration of the NPE, where the receivers are located just below the traction-free surface, that has $\mathbf{n} = (0, 0, 1)$. The domain integral is split into two segments, $\partial\mathbf{D}_0$ and $\partial\mathbf{D}_1$, which are the parts of the domain boundary that coincide with the traction-free surface and the remainder, respectively. Thus the interferometric representation, equation 1, can be split into two parts:

$$2\Re\{\hat{G}^{\upsilon,f}_{3,3}(\mathbf{x}_A,\mathbf{x}_B,\omega)\} = \tag{2}$$
$$-\oint_{\partial\mathbf{D}_1}\left[\hat{G}^{\upsilon,h}_{3,ij}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^* + \hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,h}_{3,ij}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^*\right] n_j \mathrm{d}^2\mathbf{x}$$
$$-\oint_{\partial\mathbf{D}_0}\left[\hat{G}^{\upsilon,h}_{3,i3}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^* + \hat{G}^{\upsilon,f}_{3,i}(\mathbf{x}_B,\mathbf{x},\omega)\left\{\hat{G}^{\upsilon,h}_{3,i3}(\mathbf{x}_A,\mathbf{x},\omega)\right\}^*\right] \mathrm{d}^2\mathbf{x} \quad ,$$

where $\mathbf{n} = (0, 0, 1)$ has been substituted into the integral segment over the traction-free surface, $\partial\mathbf{D}_0$. According to source-receiver reciprocity, the required response $\hat{G}^{\upsilon,h}_{3,i3}$ at the traction-free surface satisfies $\hat{G}^{\upsilon,h}_{3,i3}(\mathbf{x}_b,\mathbf{x},\omega) = \hat{G}^{\tau,f}_{i3,3}(\mathbf{x},\mathbf{x}_B,\omega) = 0$. Thus the second integral on the right hand side of equation 2 is equal to zero.

Following Wapenaar and Fokkema (2006), consider the situation when the medium outside $\mathbf{D}$ is homogeneous and when the wavefield is generated by many mutually uncorrelated sources located on $\partial\mathbf{D}_1$, acting simultaneously with a weighted power spectrum, $w(\mathbf{x})\hat{S}(\omega)$. The integral over $\partial\mathbf{D}_1$ in equation 2, can be evaluated by a direct cross-correlation between recordings of particle velocity at receiver stations A and B:

$$2\Re\{\hat{G}^{\upsilon,f}_{3,3}(\mathbf{x}_A,\mathbf{x}_B,\omega)\}\hat{S}(\omega) \approx \frac{2}{\rho c_p}\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}^*_3(\mathbf{x}_B,\omega)\,\right\rangle_x, \tag{3}$$

where c_p is the P-wave velocity. Note that the weighting factor $w(\omega)$ has disappeared. The weighting factor depends on the local medium parameters and source types, as is discussed at length by Wapenaar and Fokkema (2006). The wavefield required to make this simplification is named equipartitioned or diffuse (Hennino et al., 2001; Sánchez-Sesma and Campillo, 2006). The spatial ensemble average $\langle\cdot\rangle_x$ over sources is usually evaluated using a sufficiently long recording. Secondary scattering can render the coda of the NPE sufficiently equipartitioned. But in the coda of the NPE, most of the coherent energy between the receivers resides in the surface-wave mode (de Ridder, 2008) and does not provide significant energy for imaging.

A similar situation occurs for earthquake tremor. Aki Aki (1957) developed a technique named the spatial auto-correlation method. The close relationship between SI and SPAC was reported by Yokoi and Margaryan (2008). Their steps are briefly repeated here to derive from equation 3 a relationship used in the SPAC method. For a wavefield dominated by the surface modes, the frequency-domain Green's function for the vertical component of particle velocity measured at $\mathbf{x}_A$ in response to a vertical force-impulse point-source acting at $\mathbf{x}_B$ is

$$\hat{G}^{\upsilon,f}_{3,3}(\mathbf{x}_A,\mathbf{x}_B,\omega) \approx -\omega \sum_{n=0}^{\infty} \hat{m}_2(k_n,x_{3,A})\hat{m}_2(k_n,x_{3,B})J_0(k_n\,|\mathbf{x}_A-\mathbf{x}_B|), \tag{4}$$

where $\hat{m}_2$ is a normalized eigenfunction, and J_0 is a zero-order Bessel function of the first kind (Yokoi and Margaryan (2008) from Aki and Richards (2002)). Substituting this Green's function into the left side of equation 3 expands it to

$$2\Re\left\{-\omega \sum_{n=0}^{\infty} \hat{m}_2(k_n,x_{3,A})\hat{m}_2(k_n,x_{3,B})J_0(k_n\,|\mathbf{x}_A-\mathbf{x}_B|)\right\}\hat{S}(\omega) \approx \frac{2}{\rho c_p}\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_B,\omega)\,\right\rangle_x. \tag{5}$$

Normalizing by the auto-correlation of the recording at station A gives

$$\frac{2\Re\left\{-\omega\sum_{n=0}^{\infty}\hat{m}_2(k_n,x_{3,A})\hat{m}_2(k_n,x_{3,B})J_0(k_n\,|\mathbf{x}_A-\mathbf{x}_B|)\right\}}{2\Re\left\{-\omega\sum_{n=0}^{\infty}\left\{\hat{m}_2(k_n,x_{3,A})\right\}^2 J_0(0)\right\}} \approx \frac{\frac{2}{\rho c_p}\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_B,\omega)\,\right\rangle_x}{\frac{2}{\rho c_p}\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_A,\omega)\,\right\rangle_x}. \tag{6}$$

When the fundamental surface-wave dominates, and both receivers are located at equal depth ($x_{3,A} = x_{3,B}$), the higher-order terms can be neglected, and equation 6 simplifies to

$$J_0(k_0\,|\mathbf{x}_A-\mathbf{x}_B|) \approx \frac{\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_B,\omega)\,\right\rangle_x}{\left\langle\, \hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_A,\omega)\,\right\rangle_x} \approx \phi(\mathbf{x}_A,\mathbf{x}_B,\omega), \tag{7}$$

where $\phi(\mathbf{x}_A,\mathbf{x}_B,\omega)$ is defined as the azimuthally averaged auto-correlation coefficient. The wavenumber of the fundamental surface-wave mode is given by a specific dispersion curve, $k_0 = \frac{\omega}{c(\omega)}$, where $c(\omega)$ is phase velocity.

Notice from equation 7 how the cross-spectra in frequency *and space* are predicted to obey Bessel functions, with oscillations determined by the phase velocity. The Bessel function of the first kind is real-valued. The cross-spectrum on the right hand-side is complex-valued, but if the conditions that lead to equation 7 are fulfilled, the imaginary component vanishes (Asten, 2006), ($\hat{v}_3(\mathbf{x}_A,\omega)\hat{v}_3^*(\mathbf{x}_A,\omega)$ is always real). The real part of the cross-spectrum is retrieved as the zero-lag temporal cross-correlation, i.e., a spatial auto-correlation coefficient. It should be noted that the close relationship between SI and SPAC

seems to hold only for surface-waves in horizontally stratified media. The SPAC method as commonly applied involves fitting Bessel functions to the computed auto-correlation coefficient with frequency, with explicit directional averaging of the wavefield in all directions. In the case of isotropic wavefields, this averaging is unnecessary (Aki, 1957; Okada, 2003). The coda of the NPE quickly becomes isotropic after the first break, thus this relationship seems suitable for the cross-spectra calculated from NPE data. To introduce an estimation of attenuation, the Green's function, equation 4, is supplemented with an exponential attenuation factor $Q(\omega)$, (Aki and Richards, 2002). The final model for the frequency-domain spatial auto-correlation coefficient $\phi(\mathbf{x}_A, \mathbf{x}_B, \omega)$ becomes

$$\phi(\mathbf{x}_A, \mathbf{x}_B, \omega) = J_0\left(\frac{\omega}{c(\omega)}\left|\mathbf{x}_A - \mathbf{x}_B\right|\right) \exp\left\{-\frac{\omega}{c(\omega)}\frac{1}{2Q(\omega)}\left|\mathbf{x}_A - \mathbf{x}_B\right|\right\}. \tag{8}$$

DATA ANALYSIS

The two records analyzed are of approximately 131 seconds duration (2^{14} samples at a 125 Hz sampling frequency) at 610 stations with a 45-foot spacing. The second recording starts a few seconds after the first recording ends. Although the exact location of the NPE array is unknown, the data shows that the first 66 stations were located at an angle with respect to the other stations. Starting from station 67 located at 0 m in Figure 2, 512 stations are analyzed. The first arrivals were muted from the first record. All records were filtered in the frequency-wavenumber domain with a high-cut cosine filter centered around an angle corresponding to a velocity of 770 m/s, determined from Figure 3. This removes noise from the coda and interpolates the missing traces. For velocities smaller and equal to 770 m/s this did not affect the estimation of the surface wave velocity as described below.

The spectrum at each station was estimated using a multitaper spectral-estimation technique. This provides several statistically independent estimates of the spectrum and decreases spectral leakage (Prieto et al., 2007, 2008b). In this procedure, a time record with N samples is first multiplied with a set of K orthogonal Slepian tapers (Thomson, 1982). Second, the discrete Fourier transformation is computed for each tapered trace as follows:

$$x^k(\omega) = \sum_{t=0}^{N-1} x(t)\nu^k(t)\exp\left\{-i\omega t\right\}, \tag{9}$$

where the k^{th} Slepian taper is denoted by $\nu^k(t)$. The cross-spectrum $\rho(\omega)$ between two traces x and y is calculated from the spectral estimations $x^k(\omega)$ and $y^k(\omega)$ according to

$$\rho(\omega) = \sum_{k=1}^{K} x^k(\omega)\left\{y^k(\omega)\right\}^*. \tag{10}$$

Interferometric gathers in the frequency-domain are computed by consecutively selecting each station in the array as a master station and computing the cross-spectra between all other stations and this master station. The collection of interferometric gathers is further analyzed in the midpoint-offset domain (m, h). To enhance the signal-to-noise ratio, the retrieved gathers are smoothed over 50 midpoints, corresponding to a length of 680 m.

We first study the result from processing the records in the time domain, as is common in seismic interferometry practices. A common-midpoint section at $m = 5144$ m is given in

Figure 4, and a common-offset section for $h = 261$ m is given in Figure 5. de Ridder (2008) did not observe any event, besides the surface-wave events intersecting each other at $(h, t) = (0, 0)$ that is coherent across different midpoints. When we study the common-offset section in Figure 5, we can see the arrival time of the surface wave event at ~ 0.5 s slightly varying with offset. If we neglect this and assume the earth is horizontally layered, the recovered gathers can be stacked over common-offsets as shown in Figure 6. Estimating the slope of the event visible in Figures 4 and 6, de Ridder (2008) found a velocity of $c_r = 770$ m/s. The subscript r refers to a Rayleigh wave, which is the dominant surface-wave type recorded in the vertical component of particle velocity in groundroll. It is difficult to extract more information from the time-domain images. Additional analysis of the retrieved gathers can be performed in the frequency-domain, by inverting for phase velocity and attenuation factors. The frequency-domain equivalents of the time-domain gathers in Figures 4 and 6 are shown respectively in Figures 7 and 9.

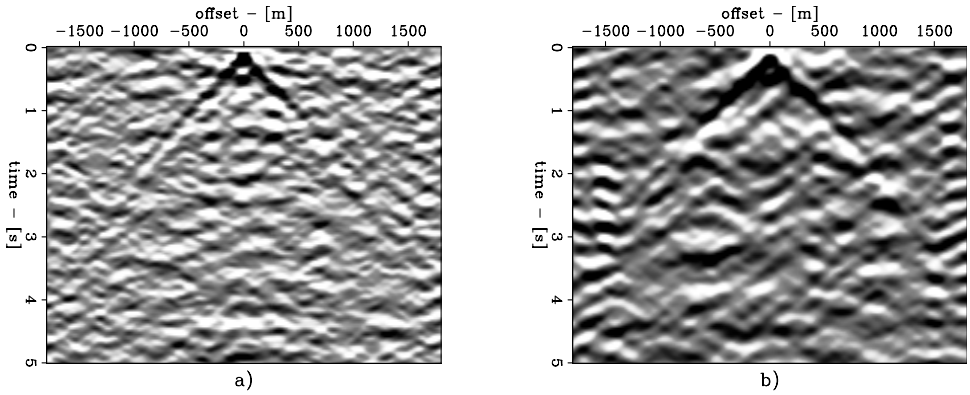


Figure 4: Interferometric common-midpoint gather gather, at $m = 5144$ m; a) retrieved from recording 1, b) retrieved from recording 2. **[ER]** sjoerd1/. si375m

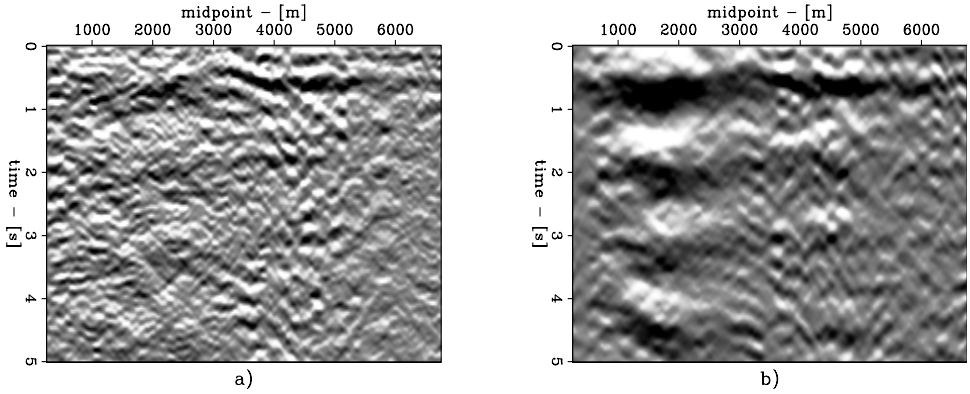


Figure 5: Interferometric common-offset gather, for $h = 261$ m; a) retrieved from recording 1, b) retrieved from recording 2. **[ER]** sjoerd1/. si375o

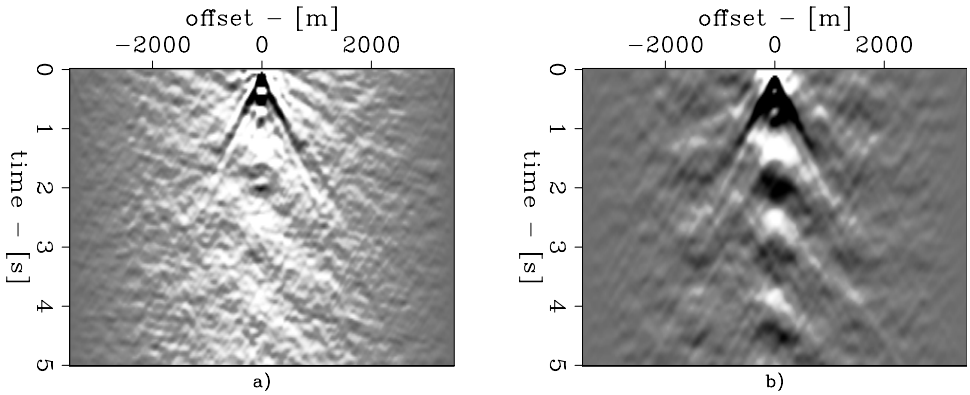


Figure 6: Common-offset stack for the interferometric gathers; a) retrieved from recording 1, b) retrieved from recording 2. **[ER]** sjoerd1/. SIHS

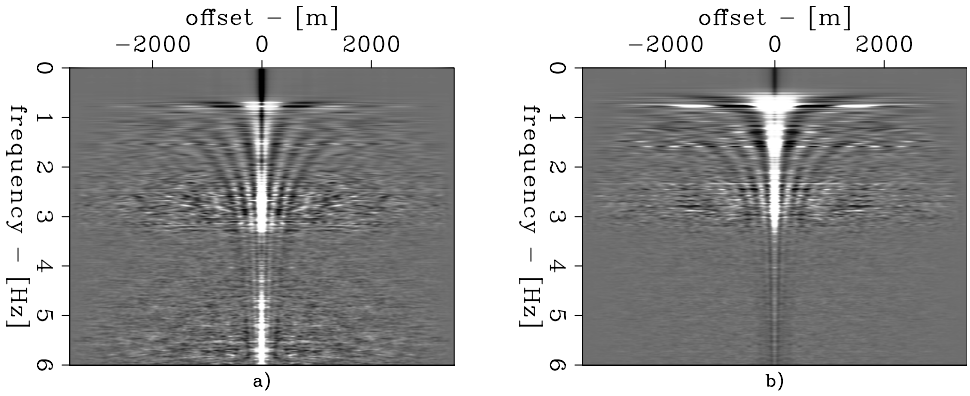


Figure 7: Frequency-domain common-offset stack; a) retrieved from recording 1, b) retrieved from recording 2. **[ER]** sjoerd1/. SPACHS

INVERSION FOR DISPERSION CURVES

In Figures 7 and 9, Bessel functions are clearly observed at different oscillation periods for different frequencies. Moreover, from Figure 5 it can be concluded that the period of oscillation of the Bessel functions should not remain constant for all gathers along the array. We use our model stated in equation 8 to estimate phase velocity and attenuation from retrieved frequency-domain common-midpoint gathers. This analysis can be repeated for all gathers at all midpoints, thereby detecting variation in dispersion curves along the array.

The optimization procedure is done by a grid-search for the two dimensions in the model space for each frequency. The inversion for phase velocity $c(\omega)$ is fairly well posed because of the zero crossings of the Bessel function; in a second step the decay rate observed in the gathers is inverted for an attenuation factor $Q(\omega)$. The division of the cross-spectrum by the auto-correlation of the recording at station A is very unstable, and is avoided by multiplication of equation 8 with the auto-correlation of the recording at station A.

An example in which theory matches the data reasonably well is at midpoint $m =$

5144 m. In Figure 8, the cross-spectrum for this midpoint and at 2.4 Hz calculated from recording 1 and 2 are shown together with the fitted Bessel and damped Bessel functions. The estimated phase velocities and attenuation factors are $c(4.8\pi) = 560$ m/s from recording 1, and $Q(4.8\pi) = 22.3$ and $c(\omega) = 562$ m/s, $Q = 17.8$ from recording 2.

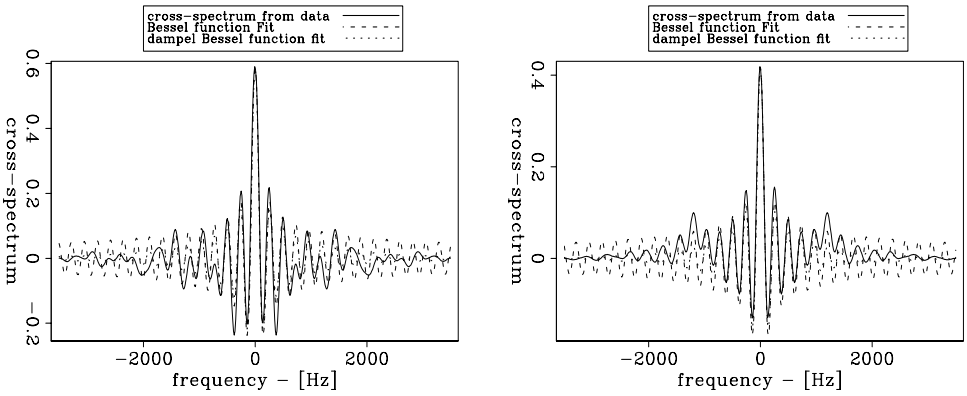


Figure 8: Cross-spectrum at 2.4 Hz at midpoint $m = 5144$ m; a) retrieved from recording 1, b) retrieved from recording 2. Included are the fitted Bessel and damped Bessel functions, from recording 1; $c(4.8\pi) = 560$ m/s, $Q(4.8\pi) = 22.3$, from recording 2; $c(4.8\pi) = 562$ m/s, $Q(4.8\pi) = 17.8$. **[ER]** sjoerd1/. bessels

This analysis is performed for all frequencies and on both recordings. The cross-spectrum calculated at midpoint $m = 5144$ m for all frequencies from recording 1 and 2 are shown in Figure 9. The best fit of Bessel functions to the midpoint-gathers in Figure 9 are shown in Figure 10 and the best fit of damped Bessel functions to the midpoint gathers in Figure 10 are shown in Figure 11. The estimated dispersion curves and attenuation factors for the midpoint gather at $m = 5144$ m are shown in Figure 12. The dispersion curves estimated for all midpoints from recording 1 and 2 are shown together in Figure 13. The dispersion curves range from values twice as large as the estimated group velocity of 770 m/s to velocities slightly smaller than the estimated group velocity. Although there are some spurious estimated phase velocities, a clear trend can be seen towards the right side of the array for generally higher phase velocities. The estimated attenuation coefficients vary much more strongly and become seemingly more incoherent with frequency. For the midpoint gather at $m = 5144$ m and the frequency band of 1 to 3 Hz, the attenuation factor seems to be in the range of $Q = 5$ to $Q = 30$.

DISCUSSION AND CONCLUSIONS

Consistent with the observation of de Ridder (2008), we have observed that the cross-spectra of the NPE coda contain coherent energy that corresponds to the surface-wave. This observation shows that for the coda of the NPE, the SI approach is analogous to the SPAC approach. Both methods provide different information; in addition to an estimate of group velocity for the surface wave from SI, we have also estimated phase velocity from the cross-spectra. Only the energy between 1 and 4 Hz can be fitted unambiguously with J_0 functions, because higher frequencies contain too much spurious energy that is not equipartitioned, and lower frequencies do not contain sufficient energy. We also show that by analysing common-midpoint gathers, we can begin to see spatial variation of phase velocities along the array.

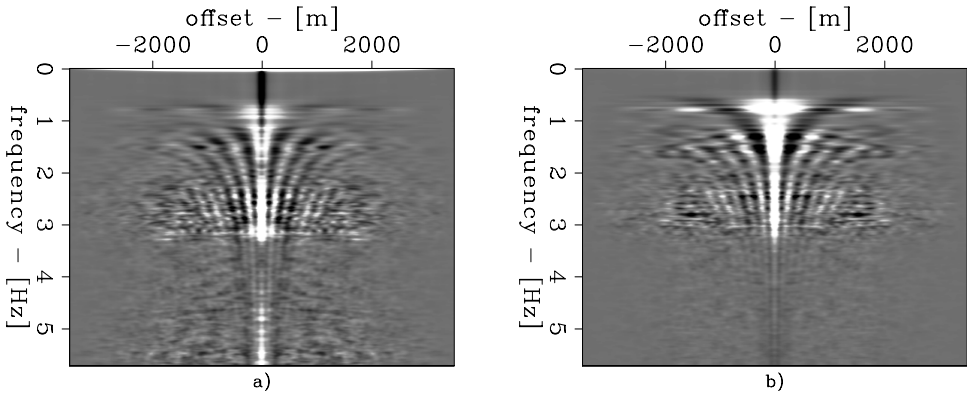


Figure 9: Cross-spectra calculated from recording 1, in a) and 2 in b) at $m = 5144$ m. **[ER]** sjoerd1/. spacm375o

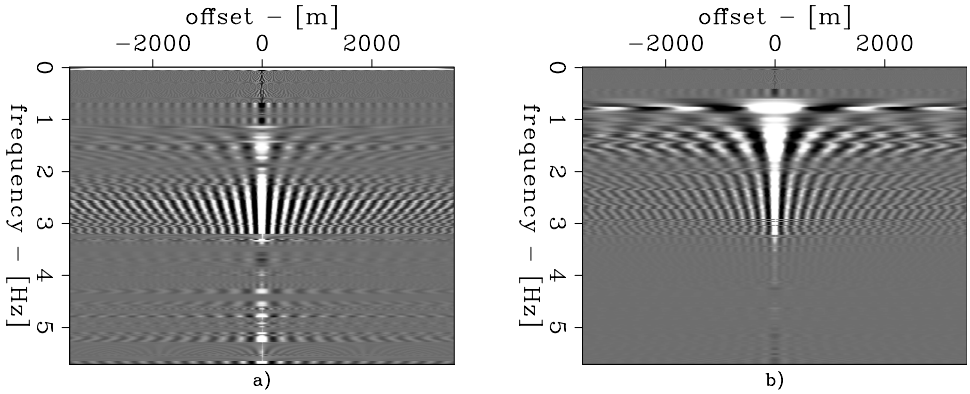


Figure 10: Bessel functions fitted to the cross-spectra calculated from recording 1, in a) and 2 in b), at $m = 5144$ m. The estimated phase velocity curves are shown in Figure 12. **[ER]** sjoerd1/. spacm375f

This suggest the possibility to study 3D structure using cross-correlation techniques, as has also been proposed by Harmon et al. (2008) and Prieto et al. (2008a).

The relatively accurate fit of the damped Bessel function for Figure 8 suggests that most of the energy is part of the fundamental mode of the surface wave. Small discrepancies can be explained by many factors; the energy might not be completely equipartitioned, there could be some energy in higher modes, and very likely there are heterogeneities on a scale smaller than the smoothing length that invalidate our assumptions of a horizontally layered earth. Another effect of scattering is attenuation (Hong et al., 2005). The effects of scattering attenuation and intrinsic attenuation are not easily distinguishable at this point and further research is needed.

The source position of the NPE, transversely oriented to the direction of the array, seems very unfavorable for SI or SPAC. But small-scale heterogeneities in the earth rendered the coda wavefield equipartitioned to first order. Since the recording is sign-bit, this property is fully exploited, because amplitude changes with azimuth are suppressed. Moreover, the energy is relatively well distributed in both directions along the array, since the array is

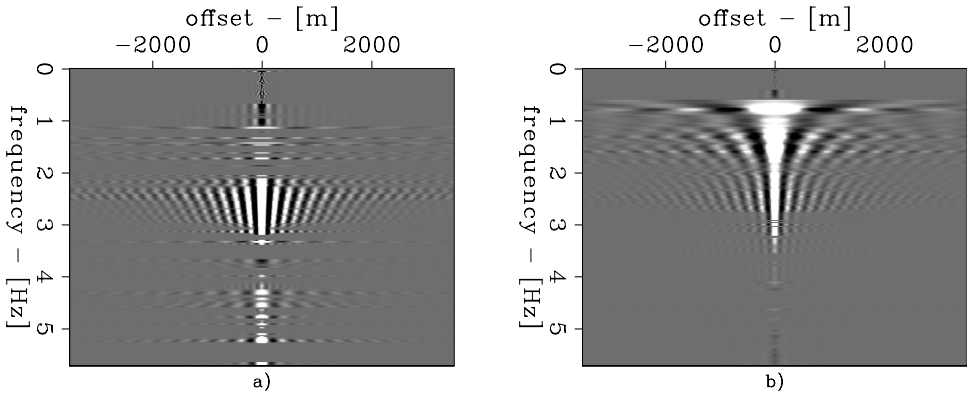


Figure 11: Damped Bessel functions fitted to the cross-spectra calculated from recording 1, in a) and 2 in b), at $m = 5144$ m. The estimated attenuation factors are shown in Figure 12. [**ER**] sjoerd1/. spacm375df

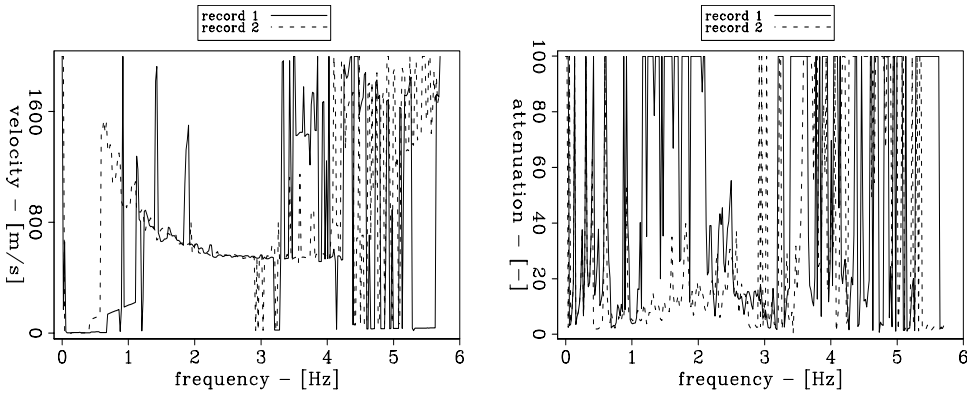


Figure 12: Dispersion curve and attenuation factors estimated from the cross-spectra at $m = 5144$ m calculated from recordings 1 and 2. [**ER**] sjoerd1/. VQcurve

nearly transverse with respect to NTS.

Although the wavefield does not contain sufficient energy for imaging, subsurface information can still be extracted by inverting the dispersion curves for a subsurface velocity profile. The recording of the NPE and geometry of the array in Railroad Valley provided the rare opportunity to study the properties of a small earthquake-like event in the transverse direction, with high spatial resolution.

ACKNOWLEDGMENTS

The author thanks Germán Prieto for many discussions on Green's function retrieval and his multitaper spectrum analysis Fortran library, Jon Claerbout for his continuing encouragement to investigate this dataset and John Vidale for supplying the data tape file.

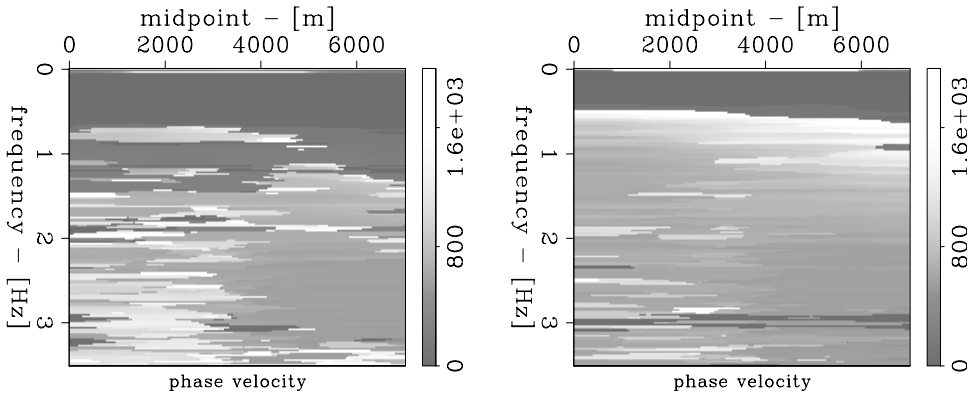


Figure 13: Estimated phase velocity from the cross-spectra calculated along the array. A clear trend of increasing phase velocities to the right-side can be observed. **[CR]** sjoerd1/. SPACV2d

REFERENCES

Aki, K., 1957, Space and time spectra of stationary stochastic waves, with special reference to microtremors: Bulletin of the Earthquake Research Institute, **35**, 415–456.

Aki, K. and P. G. Richards, 2002, Quantitative Seismology - second edition: University Science Books.

Asten, M. W., 2006, On bias and noise in passive seismic data from finite circular array data processed using spac methods: Geophysics, **71**, V153–V162.

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.

———, 1976, Fundamentals of geophysical data processing; with applications to petroleum prospecting: Blackwell Scientific Publications.

CTBTO, 2008, CTBTO Preparatory Commission: http://www.ctbto.org/. (last visited: Oktober 15, 2008).

de Ridder, S., 2008, Spectral analysis of the non-proliferation experiment: Technical report, SEP-134.

Harmon, N., P. Gerstoft, C. A. Rychert, G. A. Abers, M. S. de la Cruz, and K. M. Fischer, 2008, Phase velocities from seismic noise using beamforming and cross correlation in costa rica and nicaragua: Geophys. Res. Let., **35**, L19303–1–L19303–6.

Hennino, R., N. Trégourès, N. M. Shapiro, L. Margerin, M. Campillo, B. A. van Tiggelen, and R. L. Weaver, 2001, Observation of equipartition of seismic waves: Physical Review Letters, **86**, 3447–3450.

Hong, T.-K., R.-S. Wu, and B. L. N. Kennett, 2005, Stochastic features of scattering: Physics of the Earth and Planetary Interiors, **148**, 131–148.

Okada, H., 2003, The microtremor survey method. Geophysical Monograph, No. 12: Society of Exploration Geophysicists.

Prieto, G. A., J. F. Lawrence, and G. C. Beroza, 2008a, Anelastic earth structure from the coherency of the ambient seismic field. submitted to J. Geophys. Res.

Prieto, G. A., R. L. Parker, and F. L. Vernon, 2008b, A Fortran 90 library for multitaper

spectrum analysis. Accepted in 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: Geophys. J. Int., **168**, 1227–1234.

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.

Sánchez-Sesma, F. J. and M. Campillo, 2006, Retrieval of the Green's function from cross correlation: The canonical elastic problem: Bull. Seism. Soc. Am., **96**, 1182–1191.

Sánchez-Sesma, F. J., J. A. Párez-Ruiz, M. Campillo, and F. Luzón, 2006, Elastodynamic 2d Green's function retrieval from cross-correlation: Canonical inclusion problem: Geophys. Res. Let., **33**, L13305–1–L13305–6.

Schuster, G. T., 2001, Theory of daylight/interferometric imaging - tutorial: 63rd Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, Session A32, Expanded Abstracts, A32.

Schuster, G. T., J. Yu, J. Sheng, and J. Rickett, 2004, Interferometric/daylight seismic imaging: Geophys. J. Int., **157**, 838–852.

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: Proceedings of the IEEE, **70**, 1955–1996.

Tinker, M. A. and T. C. Wallace, 1997, Regional phase development of the non-proliferation experiment within the Western United States: Bull. Seism. Soc. Am., **87**, 383–395.

van Manen, D.-J., J. O. A. Robertsson, and A. Curtis, 2005, Modeling of wave propagation in inhomogeneous media: Phys. Rev. Lett., **94**, 164301–1–164301–4.

Wapenaar, K., 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. Let., **87**, 134301–1 – 134301–4.

———, 2004, Diffuse fields in open systems and the emergence of the Green's function (L): J. Acoust. Soc. Am., **116**, 2731–2734.

Yokoi, T. and S. Margaryan, 2008, Consistency of the spatial autocorrelation method with seismic interferometry and its consequence: Geophysical Prospecting, **56**, 435–451.

3D pyramid interpolation

Xukai Shen

ABSTRACT

Seismic data geometries are not always as nice and regular as we want due to various acquisition constraints. In such cases, data interpolation becomes necessary. Usually high-frequency data are aliased, while low-frequency data are not, so information in low frequencies can help us interpolate aliased high-frequency data. In this paper, I present a 3D data interpolation scheme in pyramid domain, in which I use information in low-frequency data to interpolate aliased high-frequency data. This is possible since in pyramid domain, only one prediction error filter (PEF) is needed to represent any stationary event (plane-wave) across all offsets and frequencies. However, if we need to estimate both the missing data and PEF, the problem becomes nonlinear. By alternately estimating the missing data and PEF, we can linearize the problem and solve it using a conventional least-squares solver.

INTRODUCTION

Data interpolation is an important step in seismic data processing that can greatly affect the results of later processing steps, such as multiple removal, migration and inversion. There are many ways to interpolate data, including Fourier-transform-based approaches (e.g., Xu et al., 2005) and PEF-based approaches (e.g., Spitz, 1991; Crawley, 2000). 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 it has the inverse spectrum of the known data, so when it is convolved with known data, it minimizes the convolution result in the least-square sense. PEF estimation can be done in either time-space (t-$\mathbf{x}$) domain or frequency-space (f-$\mathbf{x}$) domain (e.g. Claerbout, 1999; Crawley, 2000; Curry, 2007), however, if PEF estimation is done in the f-$\mathbf{x}$ domain, every frequency needs one distinct PEF.

The pyramid domain was introduced by Ronen (Hung et al., 2005), and is a resampled representation of an ordinary f-$\mathbf{x}$ domain. Although it has frequency and space axes, 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 the 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. So the information in the low frequency data can be better used to interpolate higher frequency data.

In this paper, I present a 3D version of data interpolation in the pyramid domain based on PEF estimation, which is based on Shen (2008). The paper is organized as follows: I first show the 3D pyramid transform and corresponding missing-data interpolation and PEF estimation. I then show synthetic data examples. Finally, I conclude with the advantages and disadvantages of this interpolation method.

METHODOLOGY

There are two important parts of the pyramid-based interpolation algorithm, the first of which is the selection of the pyramid transforms between pyramid domain and f-$\mathbf{x}$ domain. The more accurately these transforms are performed, the better the result we can get for missing data interpolation. The second step combines data interpolation and PEF estimation in the pyramid domain, which are done alternately in an iterative way.

3D pyramid Transform between pyramid domain and f-$\mathbf{x}$ domain

I discussed the 2D pyramid transform in Shen (2008), and the 3D version is almost the same, except that some scalars become vectors. In the 3D pyramid transform, spatial grid spacing is calculated for each frequency f using the equation

$$\Delta\mathbf{x}(f) = \frac{\Delta\mathbf{x_0 v}}{f\mathbf{n_{sf}}}, \tag{1}$$

where $\Delta\mathbf{x}(f)$, $\Delta\mathbf{x_0}$, $\mathbf{v}$ and $\mathbf{n_{sf}}$ are all 2D vectors. $\Delta\mathbf{x_0}$ is the uniform spatial grid spacing in the original f-$\mathbf{x}$ data, $\mathbf{v}$ is the velocity that controls the slope of the pyramid and $\mathbf{n_{sf}}$ is the sampling factor in pyramid domain. By changing this factor we can control how densely the pyramid domain is sampled. In situations where events to be interpolated are not perfectly stationary, dense sampling is preferrable since the information in the low frequencies cannot be represented well by only a few points. In 3D, the inversion scheme that transforms data in f-$\mathbf{x}$ space to the pyramid domain is as follows:

$$\mathbf{Lm} - \mathbf{d} \approx \mathbf{0}, \tag{2}$$

where $\mathbf{m}$ is the data in the pyramid domain, $\mathbf{d}$ is the known data in f-$\mathbf{x}$ space, and $\mathbf{L}$ is the 2D linear interpolation operator in 3D pyramid transform. The 3D pyramid transform from pyramid domain to f-$\mathbf{x}$ domina uses the following equation :

$$\mathbf{d} = \mathbf{Lm}. \tag{3}$$

Where now $\mathbf{m}$ is known and $\mathbf{d}$ is unknown data in f-$\mathbf{x}$ space.

PEF estimation and Missing data estimation

The missing data estimation algorithm presented here is different from what I presented in the previous paper (Shen, 2008). Missing data are fitted in the f-$\mathbf{x}$ domain to ensure better fitting of known data. Also, the 3D version of these algorithms use helical coordinates (Claerbout, 1999) to perform the convolution.

For PEF estimation, I try to solve the following problem assuming known pyramid data $\mathbf{m}$. Denoting convolution with m as operator $\mathbf{M}$, with $\mathbf{W}$ being a diagonal masking matrix that is 1 where pyramid data can be used for PEF estimation and 0 elsewhere, I try to solve for the unknown PEF $\mathbf{a}$ using the following fitting goal(Claerbout, 1999):

$$\mathbf{WMa} \approx \mathbf{0}, \tag{4}$$

For missing-data estimation, I start with a known PEF $\mathbf{A}$, and try to solve the following least-squares problem (Claerbout, 1999):

$$\begin{aligned} \mathbf{K}\left(\mathbf{Lm}-\mathbf{d}\right) &\approx \mathbf{0} \\ \epsilon\mathbf{WAm} &\approx \mathbf{0}, \end{aligned} \tag{5}$$

where $\mathbf{K}$ 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, and $\mathbf{W}$ is the same as explained above.

Linearized nonlinear problem

To estimate both missing data and the PEF, the problem becomes nonlinear. To avoid directly solving nonlinear problem, I linearize it by alternately estimate missing data and PEF, and use them to update each other. Corresponding pseudo code is as follows:

```
for each iteration i{
 estimate missing data m_i from PEF a_(i-1) using equation 5
 estimate PEF a_i from missing data m_i using equation 4
}
```

I start with my guess of the PEF, $\mathbf{a_0}$. First, I make an operator $\mathbf{A_0}$ that is convolution with $\mathbf{a_0}$, and I use it to estimate the missing data $\mathbf{m_0}$. From $\mathbf{m_0}$, I make an operator $\mathbf{M_0}$ that is a convolution with $\mathbf{m_0}$. Then I update $\mathbf{a_0}$ using $\mathbf{M_0}$, calling the updated $\mathbf{a_0}$ as $\mathbf{a_1}$. This process makes one iteration of the linearized problem. Then I repeat this process, making $\mathbf{A_1}$ from $\mathbf{a_1}$, updating $\mathbf{m_0}$ to $\mathbf{m_1}$ using $\mathbf{A_1}$, making $\mathbf{M_1}$ from $\mathbf{m_1}$, updating $\mathbf{a_1}$ to $\mathbf{a_2}$ using $\mathbf{M_1}$, and so on... Finally the algorithm will converge to some $\mathbf{m}$ and $\mathbf{a}$; hopefully, by careful choosing of $\mathbf{a_0}$, I will converge to the correct $\mathbf{m}$ and $\mathbf{a}$.

EXAMPLE

Here I show two examples of missing-data interpolation using the linearized iterations described above. Both are 3D synthetic examples. The first example interpolates 3D plane-waves. The second dataset is a patch from the qdome data set Claerbout (1999).

Synthetic plane-waves

In this example, there are three plane-waves with different frequency components and dips (Figure 1a). I sub-sampled the data cube by a factor of three along both the x and y axes. This causes aliasing in two of the three plane-waves (Figure 1b), which can also be seen from their f-k spectrum(Figure 2). Then with the initial guess of the PEF being a 2D Laplacian operator, the above algorithms converged to a decent result at most places, except for data points close to the edges (Figure 3a). More specifically, the remaining data, after being transformed into the pyramid domain, looks like Figure 4a. There are a lot of holes, which are caused by missing data and big sampling factors ($\mathbf{n_{sf}}$ in equation 1) used in pyramid domain. In this example, a sampling factor of 6 along the inline direction and

a factor of 8 along crossline direction are used to make sure we have enough sample to represent all the frequencies, especially low frequencies. With the initial guess of the PEF being a 2D Laplacian operator, these holes in the pyramid domain will be filled; however, since this is just a guess of the PEF, the filled information is not necessarily correct (Figure 5b). Actually, the t-$\mathbf{x}$ domain data of interpolation with this PEF are step-like functions for aliased plane-waves (Figure 3b). However, for the unaliased low-frequency plane-wave, the missing data is already correctly interpolated. After five iterations of the algorithm, using the information from low-frequency data, the interpolated data finally have the correct dips for all the frequency components (Figure 5a).

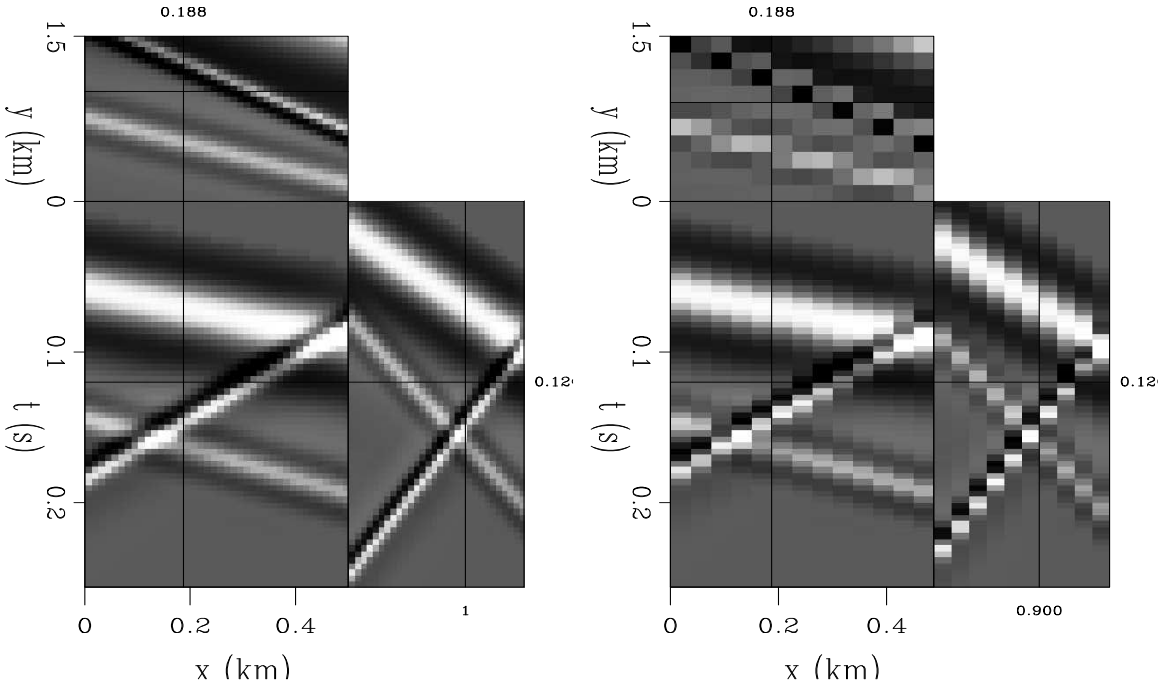


Figure 1: a) Original data, consisting of three plane waves with different dips and frequency contents. b) Sub-sampling by a factor of three along both x and y axes. [**ER**] xukai1/. 3pw

Patch from the qdome data set

In this example, I windowed out one fourth of the qdome data set, in which the reflectors are almost stationary (Figure 6 a). I then sub-sampled the data cube by a factor of four along both the x and y axes (Figure 6 b). This caused aliasing of some reflectors, especially in the cross line direction, as the f-$\mathbf{k}$ spectrum of Figure 7 shows. Since this data set is more complicated then the previous one, and not all the reflectors are ideally stationary, in the pyramid domain, I use ten as the sampling factor along both x and y directions to make sure low frequency data are well represented. Then with the initial guess of the PEF also being 2D Laplacian operator, the above algorithms converged to a satisfactory result for most of the reflectors (Figure 8b). Notice that for reflectors with small amplitude, the interpolation works not so well, and tuning the PEF size may help solving this problem. By looking at the the depth slice in Figure 8b , zeros can be observed for data at small x and y values, this is due to insufficient number of data points for PEF estimation at these

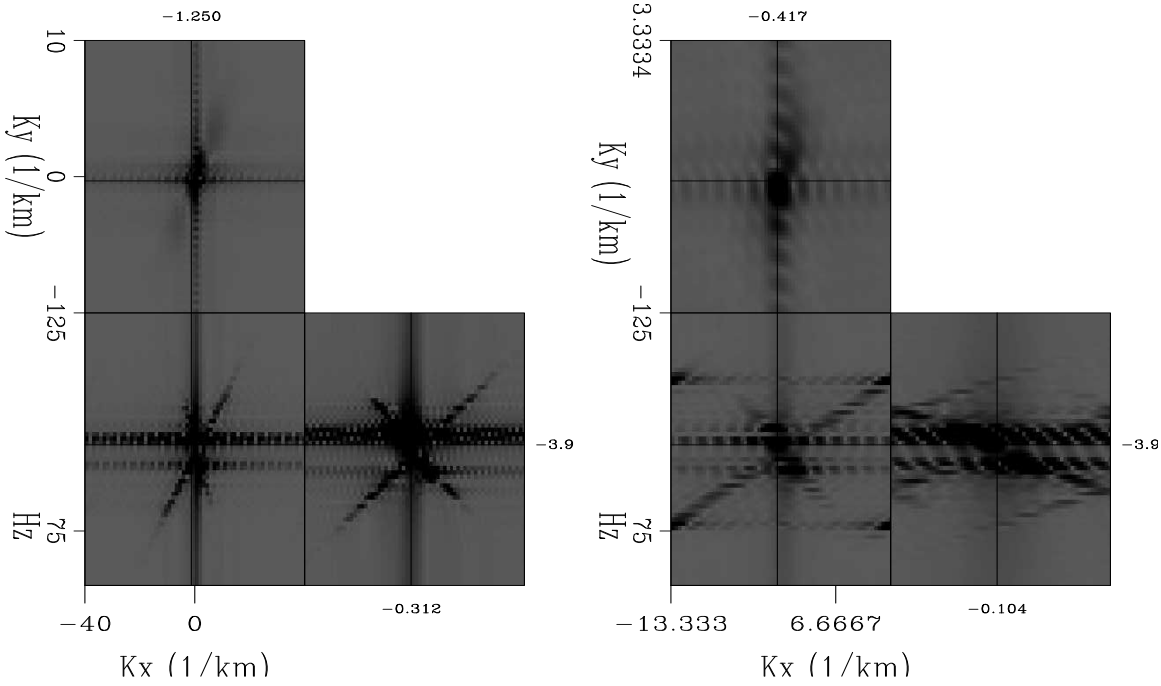


Figure 2: a) The f-k spectrum of the original data. b) The f-k spectrum of the sub-sampled data. [**ER**] xukai1/. wk3pw

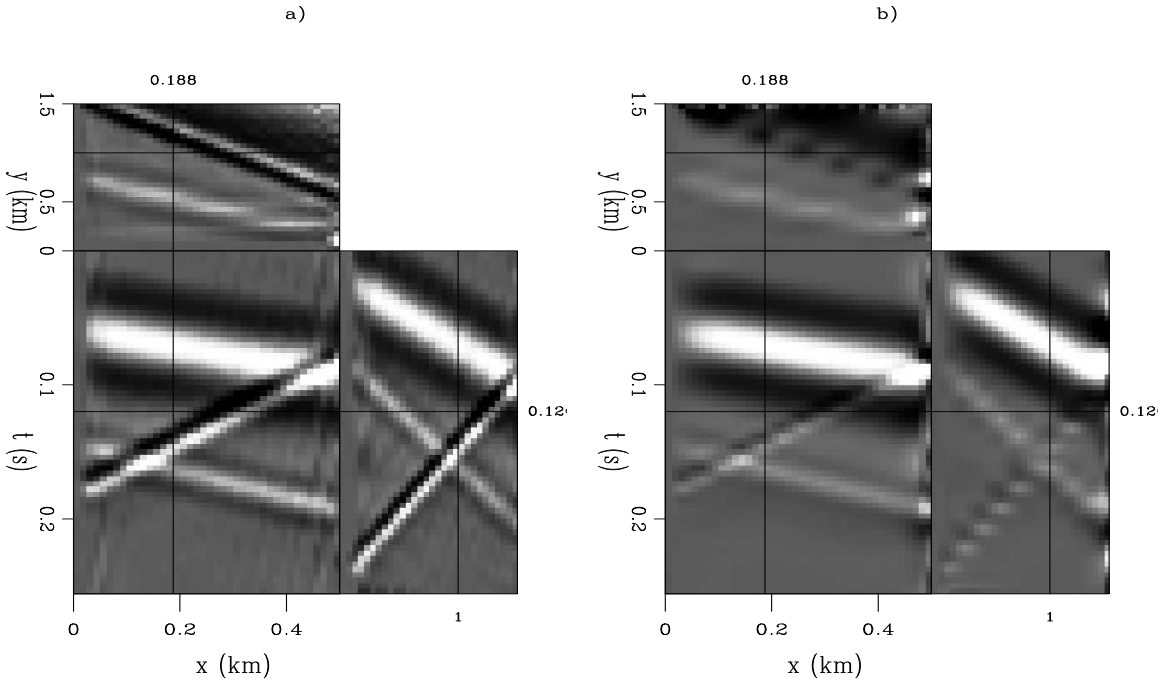


Figure 3: a) Interpolated data. b) Data interpolated with the Laplacian operator in the pyramid domain. [**CR**] xukai1/. 3pwi0

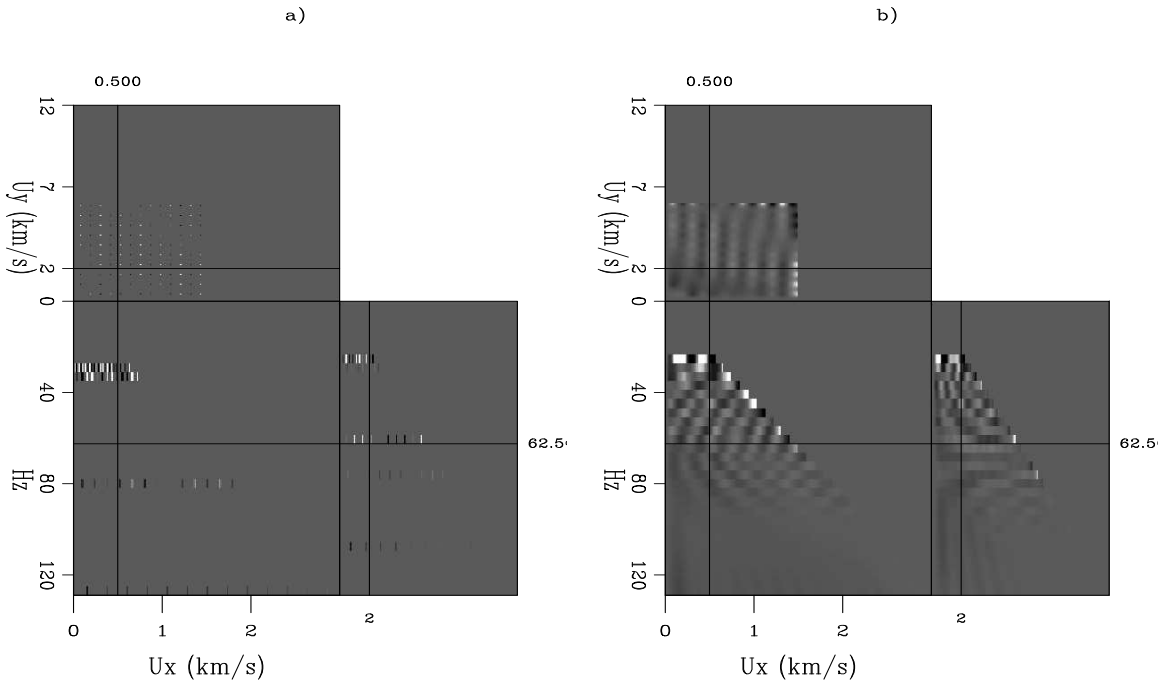


Figure 4: a) Remaining data in the pyramid domain. b) Data interpolated with the Laplacian operator in pyramid domain. [**CR**] xukai1/. py3pw0

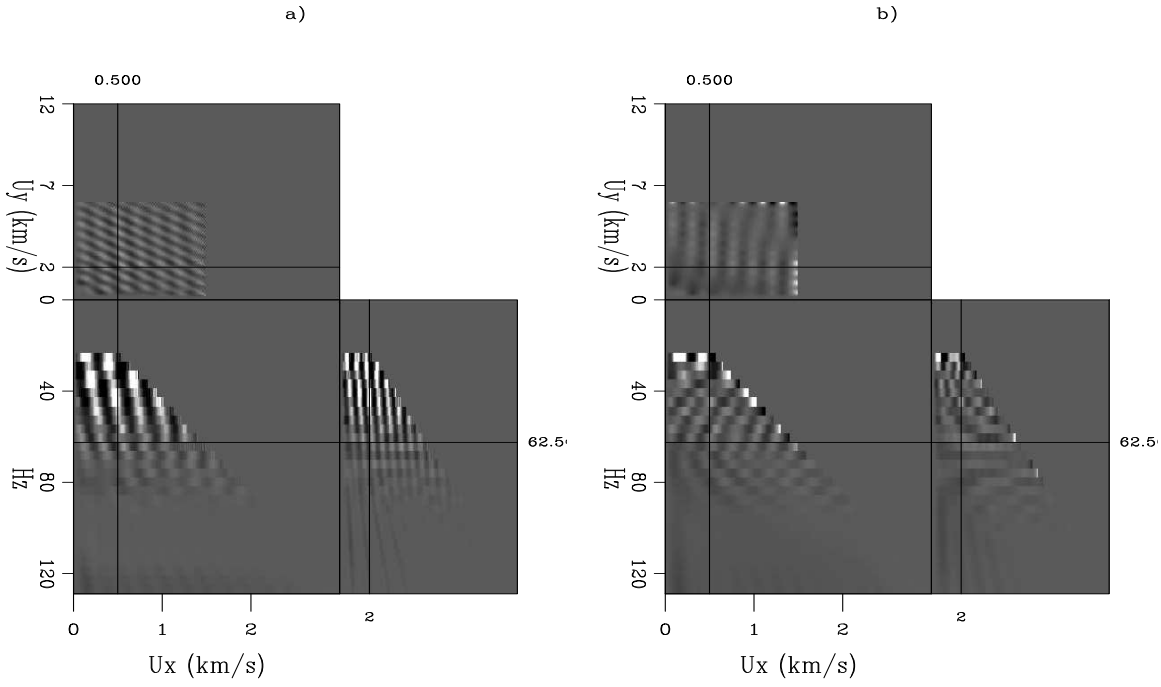


Figure 5: a) Interpolated data (final version) in pyramid domain. b) Data interpolated with the Laplacian operator in pyramid domain. [**CR**] xukai1/. py3pw

locations. On the other hand, the other two edges have much stronger amplitude artifacts, which requires further investigation.

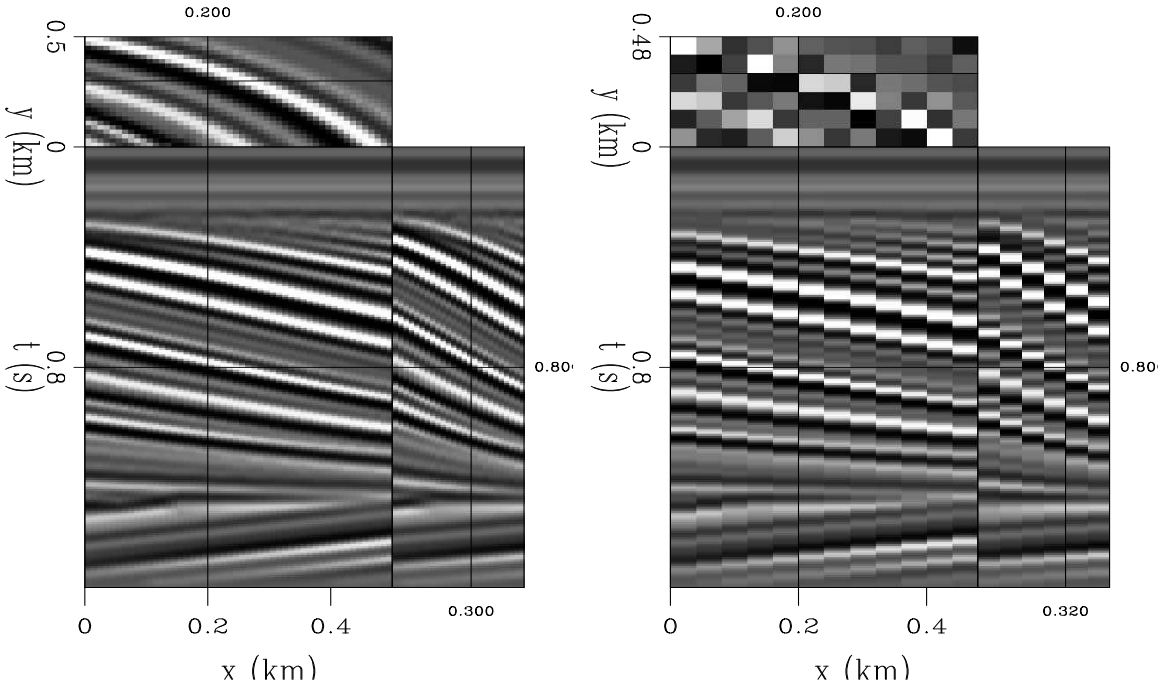


Figure 6: a) Original patch of qdome data, where all the reflectors are almost stationary. b) Sub-sampling by a factor of four along both the x and y axes. [**ER**] xukai1/. qd

FUTURE WORKS

So far, for stationary events, this algorithm works quite well. For real data, the assumption of stationarity holds if we break the data into small patches and look at each individual patch. So the next step will be to looking at more complex data and try to apply this algorithm with patching technique.

CONCLUSIONS

The pyramid domain is a very promising domain for missing-data interpolation. The synthetic examples demonstrate that, with a good initial PEF estimate, we can use the information in the low frequency to interpolate the aliased missing data relatively accurately and interpolate the unaliased missing data fairly well.

One disadvantage of this interpolation scheme is computation cost. First, to get a decent result, the data samples in pyramid domain is an order more than that in f-$\mathbf{x}$ domain along each spatial axis; in 3D, that amounts to a factor of 100 or more. In addition, the linearized nonlinear iteration adds a factor of about five in the synthetic test. In other words, we have to do both PEF estimation and data interpolation five times in total. Altogether, we first increase the data size by a factor of 100, then run about 5 rounds of data estimation. So the overall computational cost is about 500 times greater than a conventional PEF based

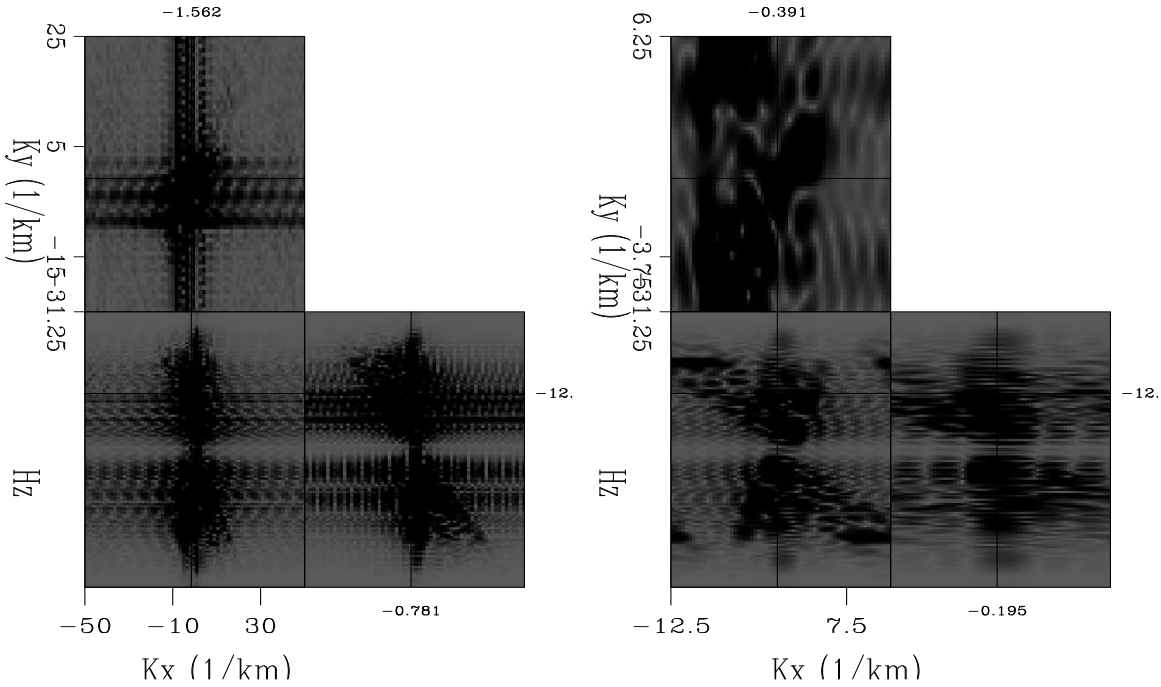


Figure 7: a) The f-k spectrum of the original data. b) The f-k spectrum of the sub-sampled data. [**ER**] xukai1/. wkqd

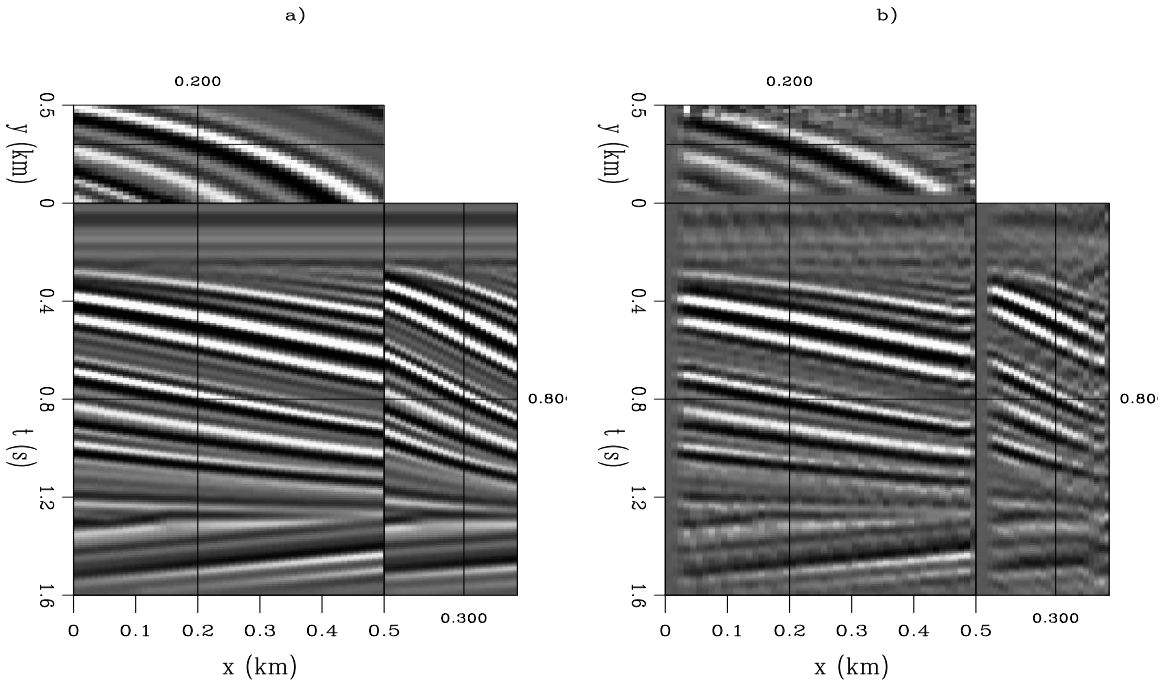


Figure 8: a) Original patch of qdome data. b) Interpolated data. [**CR**] xukai1/. qdi

interpolation scheme (e.g. Spitz, 1991). However, with a patching technique, many patches of data can be interpolated simutaneously using parallelized version of this algorithm.

ACKNOWLEDGMENTS

I would like to thank Jon Claerbout for ideas about data interpolation in the pyramid domain, Bob Clapp for synthetic example suggestions, and Antoine Guitton and William Curry 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**.

Hung, B., C. Notfors, and S. Ronen, 2005, Seismic data reconstruction using the pyramid transform: 67th Meeting, EAGE, Expanded Abstracts, A039.

Shen, X., 2008, Data interpolation in pyramid domain: **SEP-134**, 225–230.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **56**, 785–794.

Xu, S., Y. Zhang, D. Pham, and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, V87–V95.

Anti-crosstalk

Jon Claerbout

ABSTRACT

Inverse theory can never be wrong because it's just theory. Where problems arise and opportunities are overlooked is where practitioners grab and use inverse theory without recognizing that alternate assumptions might be better. Here I formulate *anti-crosstalk* operators to supplement (or replace) familiar regularizations in image estimation. Applications appear abundant.

Theoretical example

First let us consider an extremely simple theoretical example with an issue of "cross talk". Notice that word itself is absent from linear inverse theory. (Anyway, it's absent from the index of Tarantola's book.) Normally we think of "high frequencies" as being orthogonal to "low frequencies". In this example that will be true from a theoretical viewpoint, but no practical person would consider it to be true.

Consider a signal that is an impulse which is to be split into its low and high frequencies. We might write this as $0 = l + h - d$, namely, zero equals low plus high minus data. Let us extract low frequencies from the data with a step function in the Fourier domain or convolving with a sinc function in the time domain. Clearly the high frequency component in the Fourier domain is the constant function minus the step. In the time domain it is a delta function minus a sinc (since high plus low is a delta function). Theoretically everything is fine. Usually a vanishing cross product means a sum of terms vanishes. Here every term vanishes in the Fourier domain when we multiply the step times one minus the step. That is powerful orthogonality! In the time domain the convolution of the two, $sinc * (1 - sinc) = sinc - sinc * sinc = sinc - sinc = 0$ vanishes not just at zero lag, but at all lags. That is powerfully strong orthogonality too. Theoretically, high and low frequency components of the data are orthogonal at every frequency and at every lag. What would the experimentalist have to say? The experimentalist would look in the time domain at the low frequency function and at the high frequency function and say, "Everywhere I look at these two functions they are the same. They can't be orthogonal. They have a massive amount of *crosstalk*." Of course the two functions are not exactly the same. They have opposite polarities, and they are not the same at the origin point. But everywhere else they look the same. We don't like it. These signals should not be coherent but they are.

I first encountered this crosstalk issue in a serious geophysical application of a very simple nature. The lack of a good place for crosstalk in the theoretical framework was blatantly obvious.

Lake example

A depth sounding survey was made of a lake. A boat with a depth sounder sailed a gridwork of passes on the lake. Upon analysis the final image contained obvious evidence of the survey grid. Oops! We should always hide our data acquisition footprint. The footprint is not the geography or geology we wish to show. How did this happen? We guessed the water level changed during the survey. Perhaps it rained or perhaps the water was used for agriculture. Perhaps the wind caused the lake to "pendulum" (seiche). Perhaps the operator sat in the front of the boat, or sometimes its back, or ran it at various speeds giving the depth measurement a different bias. Our data measured the difference between the top and the bottom of the lake; yet we had no idea how to model the top. Eventually we modeled the top as an "arbitrary low frequency function" in data space (a one dimensional function following the boat). This got rid of the tracks in the model space (map space) but led to much embarrassment. It was embarrassing to discover that the geography (as seen in data space) was correlated with the lake surface (rain and drain).

Let us express these ideas mathematically: $\mathbf{d}$ is the data, depth along the survey coordinate $d(s)$. Here s is a parameter like time. It increases steadily whether the boat is sailing north-south or east-west or turning inbetween. The model space is the depth $h(x, y)$. There will be a regularization on the depth, perhaps $0 \approx \nabla h(x, y)$.

For the top of the lake with the ship we need some slowly variable function of location s. It's embarrassing for us to need to specify it because we have no good model for it. So, we specify a slowly variable function $\mathbf{u} = u(s)$ by asking a random noise function $\mathbf{n} = n(s)$ to run thru a low frequency filter, say $\mathbf{L}$. We are not comfortable also about needing to choose $\mathbf{L}$. We call the function $\mathbf{u} = \mathbf{Ln}$ the rain and drain function. We take the regularization for the unknown $\mathbf{n}$ to be $\mathbf{0} \approx \mathbf{n}$. (Least squares will tend to drive components of $\mathbf{n}$ to similar values, and under some conditions likewise the spectrum of $\mathbf{n}$ will tend to white, so we expect (and often find) the spectrum of $\mathbf{u}$ comes out that of $\mathbf{L}$.)

The operator we do understand very clearly is the geography operator $\mathbf{G}$. Given we wish to make a theoretical data point (water depth), the geography operator $\mathbf{G}$ tells us where to go on the map to get it. Of course each of the two regularizations $0 \approx \nabla h(x, y)$ and $\mathbf{0} \approx \mathbf{n}$ has its own epsilon which is annoying because we need to specify those too. With all these definitions our unknowns are the geography $\mathbf{h}$ and the noise $\mathbf{n}$ that builds us a drift function. Our data fitting goal says the data should be the separation of the top and bottom of the lake.

$$\mathbf{0} \quad \approx \quad \mathbf{Gh} + \mathbf{Ln} - \mathbf{d} \tag{1}$$

In my free on-line textbook GEE all this seemed rather conventional and rather fine. Our embarrassment came when we compared the geographically modeled part of the data $\mathbf{Gh}$ to the drift (rain and drain) modeled part of the data $\mathbf{Ln}$. They were visibly correlated. This is crazy! The boat being in deep water should not correlate with rain (or drain). We needed to add an ingredient to the formulation saying $\mathbf{u} = \mathbf{Ln}$ should be orthogonal to $\mathbf{Gh}$ (which is practically the same as $\mathbf{d}$) in some generalized sense. Let us see how this might be done.

A regression to minimize crosstalk

Observing the geographically modeled data **Gh** correlating with the data drift $\mathbf{u} = \mathbf{Ln}$ we wish to articulate a regression that says they should not correlate. Since the drift **u** is a small correction to the data **d**, in other words $\mathbf{Gh} \approx \mathbf{d}$, we can simplify the goal by asking that the dot product of **d** with **u** should vanish, vanish not necessarily over the entire data set; but that it should vanish under many triangular weighed windows.

Let us define **D** as a diagonal matrix with **d** on the diagonal. This may be a little unfamiliar. Often we see positive weighting functions on the diagonal. Here we see data (possibly with both polarities) on the diagonal. Additionally, let us define a matrix **T** of convolution with a triangle. Columns of **T** contain shifted triangle functions, likewise do rows. Take $\mathbf{t}'$ to be any row of **T**. Then $\mathbf{t}'\mathbf{D}$ is a row vector of triangle weighted data. We want the regression $\mathbf{0} \approx \mathbf{t}'\mathbf{Du}$ for all shifts of the triangle function. The way to express this is:

$$\mathbf{0} \approx (\mathbf{TD})\mathbf{u} \tag{2}$$

$$\mathbf{0} \approx (\mathbf{TD})\mathbf{Ln} \tag{3}$$

Hooray! Now we know what coding to do! But first, to better understand the regression (2) imagine instead that **T** is a square matrix of all ones, say **1**. That would be like super wide triangular windows. Then every component of the vector **1Du** contains the same dot product $\mathbf{d} \cdot \mathbf{u}$. Using **T** instead of **1** gives us those dot products under a triangle weight, each final vector component having a shifted triangle.

What is a good name for **TD**? It measures the similarity of **d** and **u**. It might be called the "data similarity" operator. What is a good name for its adjoint **DT**? Assuming whatever comes out of **T** is a smooth positive function, then **DT** is a data gaining operator (its input being a gain function). Do we have any geophysical problems where the unknown is the gain?

Crosstalk in a more general context

We seemed to escape nonlinearity in the lake depth sounding example above, but that was a lucky accident. Since the data there was mostly explained by geography with a small perturbation by rain and drain, the crosstalk while fundamentally nonlinear was practically linear. More generally anti-crosstalk strategies seem little (if ever!) developed because they lead us directly into nonlinear regression. Let us work through the general case, the nonlinear theory.

Consider data **d** a shot gather or CMP gather. We might choose to model it as reflections (hyperbolas) $\mathbf{d}_1$ plus linear events $\mathbf{d}_2$ (noises or head waves). We might thus set up the regression

$$\mathbf{0} \approx \mathbf{d}_1 + \mathbf{d}_2 - \mathbf{d} \tag{4}$$

$$\mathbf{0} \approx \mathbf{F}_1\mathbf{m}_1 + \mathbf{F}_2\mathbf{m}_2 - \mathbf{d} \tag{5}$$

Of course we need some damping regularization on $\mathbf{m}_1$ and $\mathbf{m}_2$ which for simplicity of exposition I will take to be $\mathbf{0} \approx \mathbf{m}_1$ and $\mathbf{0} \approx \mathbf{m}_2$. Is that all there is to this problem? Not

necessarily. We'll be annoyed if we discover a lot of cross talk between $\mathbf{d}_1$ and $\mathbf{d}_2$. Two different physical mechanisms are supposed to have created our data. We'll be annoyed to discover they both make the same contribution or that they make opposite contributions. The regularization should reduce (or prevent) the contributions from coming out opposite. If they are opposites, it could represent our lack of analytic skills in formulating the regularization, or it could represent our need for the anti-crosstalk methodology being proposed here.

We'd like that the modeled data parts $\mathbf{d}_1$ and $\mathbf{d}_2$ do not "look like" each other. It's not enough that the dot product $\mathbf{d}_1 \cdot \mathbf{d}_2$ vanish. That dot product should be small under all shifted (say triangular) weighting windows. Since $\mathbf{d}_1$ is a linear function of the model $\mathbf{m}_1$ and likewise for $\mathbf{d}_2$, the orthogonality we seek involves the product of $\mathbf{m}_1$ with $\mathbf{m}_2$ so our goals are a non-linear function of our unknowns. Never fear. We have done non-linear problems before. They don't turn out badly when we are able to define a good starting location (which we do by solving the linearized non-linear problem first).

The full non-linear derivation

For warm up we linearize in the simplest possible way. Suppose we allow only $\mathbf{m}_1$ to vary keeping $\mathbf{m}_2$ fixed. We put $\mathbf{d}_2$ on the diagonal of a matrix, say $\mathbf{D}_2$. The regression for anti-crosstalk is now

$$\mathbf{0} \approx \mathbf{T}\mathbf{D}_2\mathbf{d}_1 \tag{6}$$

$$\mathbf{0} \approx \mathbf{T}\mathbf{D}_2\mathbf{F}_1\mathbf{m}_1 \tag{7}$$

Define the element-by-element cross product of $\mathbf{d}_2$ times $\mathbf{d}_1$ to be $\mathbf{d}_1 \times \mathbf{d}_2$. Now let us linearize the full non-linear anti-crosstalk regularization. Let a single element of $\mathbf{d}_1 \times \mathbf{d}_2$ be decomposed as a base plus a perturbation $d = \bar{d} + \tilde{d}$. A single component of the vector $\mathbf{d}_1 \times \mathbf{d}_2$ is $(\bar{d}_1 + \tilde{d}_1)(\bar{d}_2 + \tilde{d}_2)$. Linearizing the product (neglecting the product of the perturbations) gives

$$\bar{d}_2\tilde{d}_1 + \bar{d}_1\tilde{d}_2 + \bar{d}_2\bar{d}_1 \tag{8}$$

This is one component. We seek an expression for all. It will be a vector which is a product of a matrix with a vector. We want no unknowns in matrices; we want them all in vectors so we will know how to solve for them.

$$\bar{\mathbf{D}}_2\tilde{\mathbf{d}}_1 + \bar{\mathbf{D}}_1\tilde{\mathbf{d}}_2 + \bar{\mathbf{D}}_1\bar{\mathbf{d}}_2 \tag{9}$$

Express the perturbation parts of the vectors as functions of the model space

$$\bar{\mathbf{D}}_2\mathbf{F}_1\tilde{\mathbf{m}}_1 + \bar{\mathbf{D}}_1\mathbf{F}_2\tilde{\mathbf{m}}_2 + \bar{\mathbf{D}}_1\bar{\mathbf{d}}_2 \tag{10}$$

This vector should be viewed under many windows (triangle shaped, for example). Under each window we hope to see the product have a small value. The desired anti-crosstalk regression is to minimize the length of the vector below by variation of the model parameters $\tilde{\mathbf{m}}_1$ and $\tilde{\mathbf{m}}_2$.

$$\mathbf{0} \approx \mathbf{T}(\bar{\mathbf{D}}_2\mathbf{F}_1\tilde{\mathbf{m}}_1 + \bar{\mathbf{D}}_1\mathbf{F}_2\tilde{\mathbf{m}}_2 + \bar{\mathbf{D}}_1\bar{\mathbf{d}}_2) \tag{11}$$

This regression augments our usual regularizations. Perhaps it partially or significantly supplants them. Unfortunately, it requires yet another epsilon.

Upon finding $\tilde{\mathbf{m}}_1$ and $\tilde{\mathbf{m}}_2$ we update the base model $\bar{\mathbf{m}} \leftarrow \bar{\mathbf{m}} + \tilde{\mathbf{m}}$ and iterate.

Outlook

Many examples suggest themselves.

1. We might model reflection data as a superposition of primaries and multiples. We might model it as a superposition of pressure waves and shear waves.
2. In tomography we might model event flatness as a superposition of shallow and deep slownesses. The shallow and deep slownesses have wholly different causes separated by millions of years. They should not show crosstalk.
3. The problem of segregating signal and noise offers many examples. We'd like to see signal containing no evident noise and vice versa.
4. In time-lapse seismology we would like to see the the image change unpolluted by the original image. Unfortunately, the methodology proposed here does not allow for time-shifted crosstalk.

Getting started

As the concepts here are quite new to us, the first thing we should do is cook up some super simple synthetic examples. With working synthetic codes in hand we should see if we can go ahead and repair the Galilee survey. Hopefully we'll recognize we have built some reusable software to facilitate other projects.

Getting started will not be easy. Most commonly we have a simple synthetic example under control and struggle to find an appropriate real data set. Here we have a suitable beginners' data set (Galilee) but we need to find a synthetic data set to provide examples that give clarity to the whole process. Just one issue is dealing with the relative scaling of the three regularizations. We'd like meaningful examples where only one or two of the regularizations are actually required.

There are many paths to explore with anti-crosstalk technology. Besides the many potential applications one can hope that the anti-crosstalk regularization eliminates (or reduces) the need for the usual regularizations. That would be nice if true. The need to specify a suitable regularization is often what makes it difficult to automate data analysis based on inversion.

I'm worried about the job as I defined it for the first-year SEP students. Given the linearization I suggested to them, did they know how to measure success?

Seismic investigation of natural coal fires: A pre-fieldwork synthetic feasibility study.

Sjoerd de Ridder and Seth S. Haines

ABSTRACT

Natural underground coal fires are a world-wide concern, emitting carbon dioxide and other pollutant gasses into the atmosphere; one such coal fire is located at Durango, Colorado. We carried out elastic modeling in order to investigate the potential of applying P-wave seismic methods to the problem of differentiating between burned and unburned coal in the upper ∼30 m of the subsurface at the site near Durango. This is a challenging problem for any geophysical method, but preliminary modeling results show that the problem is tractable under certain circumstances. Our highly simplified model suggests that imaging the coal layer can potentially be accomplished with adequately high frequencies (source center frequency $>$ 125 Hz); imaging the actual burned zone would be more difficult. The model neglects the major near-surface heterogeneity known to exist at the site; features such as fissures would surely result in diffractions and reflections that could obscure much of the desired signal.

INTRODUCTION

Ongoing research at a site outside Durango, Colorado, is focused on characterizing and modeling a coal bed that has been burning underground for a number of years. The goal is to develop an understanding of how coal fires burn for the purpose of designing optimum solutions for extinguishing coal fires around the world. The unique aspect of this particular fire is that it is small-scale and easily accessible. We were asked whether seismic methods could potentially provide useful information at the site, ideally as a way to map the burned and unburned coal. Any related information would be useful, such as mapping only the unburned coal, and/or locating fissures and other subsidence features that might not reach the surface.

We conducted elastic modeling to simulate a P-wave seismic survey at the site, to assess the likelihood of successfully applying seismic methods to this problem, and to determine optimal parameters for any field data acquisition. The model is highly simplified given that it neglects the fissures and other overburden heterogeneity. The main target is the partially burned coal layer.

The site is roughly 300m by 100m. Researchers of the Global Climate & Energy Project at Stanford University (GCEP, 2008) have been studying the site using a number of different approaches to understand the progression of the fire over time and try to predict its likely future trends. One approach has been the mapping of surface anomalies across the site – mainly fissures resulting from the compaction of the ash as the coal burns. Researchers have also drilled a total of 14 wells across the site, in areas where they know the coal has burned, where they think the burn front is located, and in unburned areas. A 100 ft core

was recovered from one of these wells. Five wells have geophysical logs (caliper, gamma and density), and rock cutting samples were taken at 5-ft intervals in all the wells.

The coal bed is approximately 5 m thick where it is unburned and closer to 2-4 m thick where it has burned. It outcrops northwest of the site and dips to the southeast. In the area of the site itself, the coal bed ranges from approximately 10 m depth to approximately 30 m. Wells showing ash are predominantly in the northwest (up-dip) end of the site, with unburned coal located down-dip. The coal bed is embedded within a sandstone formation that also includes some shale layers. A thin unconsolidated layer composed of rocky soil up to approximately 1.5 m thick overlies parts of the site.

We present specifics of our elastic model, and then present the resulting synthetic data. We examine the data in various domains, and show a brute stack of the data set. Finally, we discuss our findings in relation to the original question, and consider the impacts of neglected aspects of the modeling.

ELASTIC MODELING

We have used the code e3d, described by Larsen and Grieger (1998), to simulate a seismic survey at the coal fire field site. E3d is an explicit 2D/3D elastic finite-difference wave-propagation code that is widely used (e.g., Martin et al., 2006). It is accurate to fourth order in space and second order accurate in time.

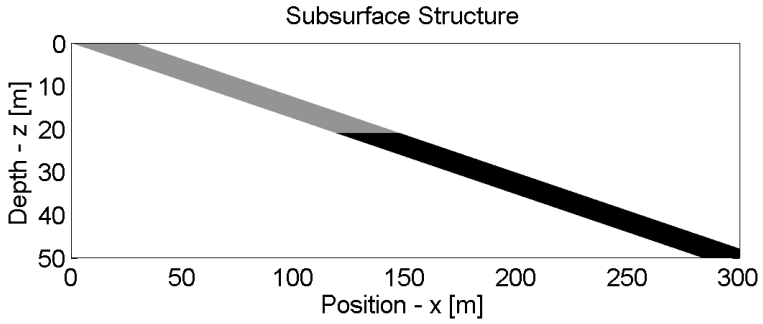


Figure 1: Subsurface structure; a dipping, partially burned (grey) coal (black) layer. P and S wave velocities are given in Table 1. [**NR**] sjoerd2/. Vmodel

Figure 1 shows the layout of our 2D model space, with layer properties shown in Table 1. A 5m thick layer of coal, half burned, dips 10° across a background of sandstone. We synthesized a seismic survey across the model with shots and receivers at 1-m spacings. The source is a vertical force with the waveform of a 125-Hz Ricker wavelet. This is intended to simulate a better-than-average sledgehammer impact. The frequency of a real sledgehammer impact is likely to be considerably lower and the waveform completely dependant on the local conditions. The receivers record both vertical, V_z, and horizontal, V_x, (in-plane) components of particle velocity. In this report we focus on the vertical (PP) component but include the horizontal component for interested readers. The model neglects the major near-surface heterogeneity like soil layer, shale layers and fissures known to exist at the site.

DATA

Figure 2 shows a pair of shot gathers for a shot at position 150 m, directly over the end of the burn front. The reflection arriving at ∼0.03 s at zero offset is the reflection from the top of the coal seam. The reflection from the base of the coal very closely follows the reflection from the top, and the two interfere. Multiples dominate the remainder of the record. In the common-offset domain (Figure 3) we see clear, though faint, diffractions from the end of the coal and note that the reflection from the base of the coal/ash layer is difficult to discern here as well. To the right of 150 m, the reflections from the top and bottom of the coal combine to create the observed pattern of interfering surface and reflected waves. To the left of 150 m, the basal reflection is obscured by the multiples. At 150 m, we can just see the basal reflection diving away due to velocity pull-down. The jumbled patterns at the left end of the shot gather and common-offset gather are observed where the ash layer crops out at the surface.

A quick processing flow (f-k filter with high-cut of $\frac{\omega}{k_x} = 1250$ m/s for surface waves, and with NMO correction but no DMO or migration) yields the stacked section in Figure 4. The reflection from the top of the coal/ash layer is clearly observed. The slight lateral change observed at position 150 m is due to the change from ash to coal, and represents the velocity pull-down of the basal reflection where low-velocity ash exists. Left of 150 m the observed reflection is due only to the top of the coal/ash layer. Right of 150 m the reflection is due to both the top and bottom of the coal.

Figure 5 shows a stacked section low-pass filtered with 100 Hz, to show what we might see for a lower source bandwidth. It is important to note that 100 Hz is still at the high end of what can reliably be expected with a sledgehammer (e.g., Miller et al., 1992; Miller et al., 1994). The reflections for the top and bottom of the coal are much more difficult to resolve from each other, and would be very difficult to interpret.

DISCUSSION

Our modeling suggests that, with a suitably high-frequency source and optimal field conditions, useful information can be gained from seismic surveys at the Durango coal fire site. We could realistically hope to discern changes across the burn front such as diffractions; these would be difficult to interpret because fissures would produce other, possibly stronger, diffractions. Separately imaging the top and bottom of the coal would be much more useful and would require frequencies at least as high as those used in the model. A lower source bandwidth would be inadequate.

Attaining frequencies even as high as those in the model (125 Hz) is difficult in prac-

Table 1: Seismic P and S wave velocities and velocity gradients.

Unit	V_p [m/s]	$\partial_z V_p [1/s]$	V_s [m/s]	$\partial_z V_s [1/s]$
Sandstone	2000	.2	1200	.154
Coal	1200	0.0	800	0.0
Ash	300	0.0	200	0.0

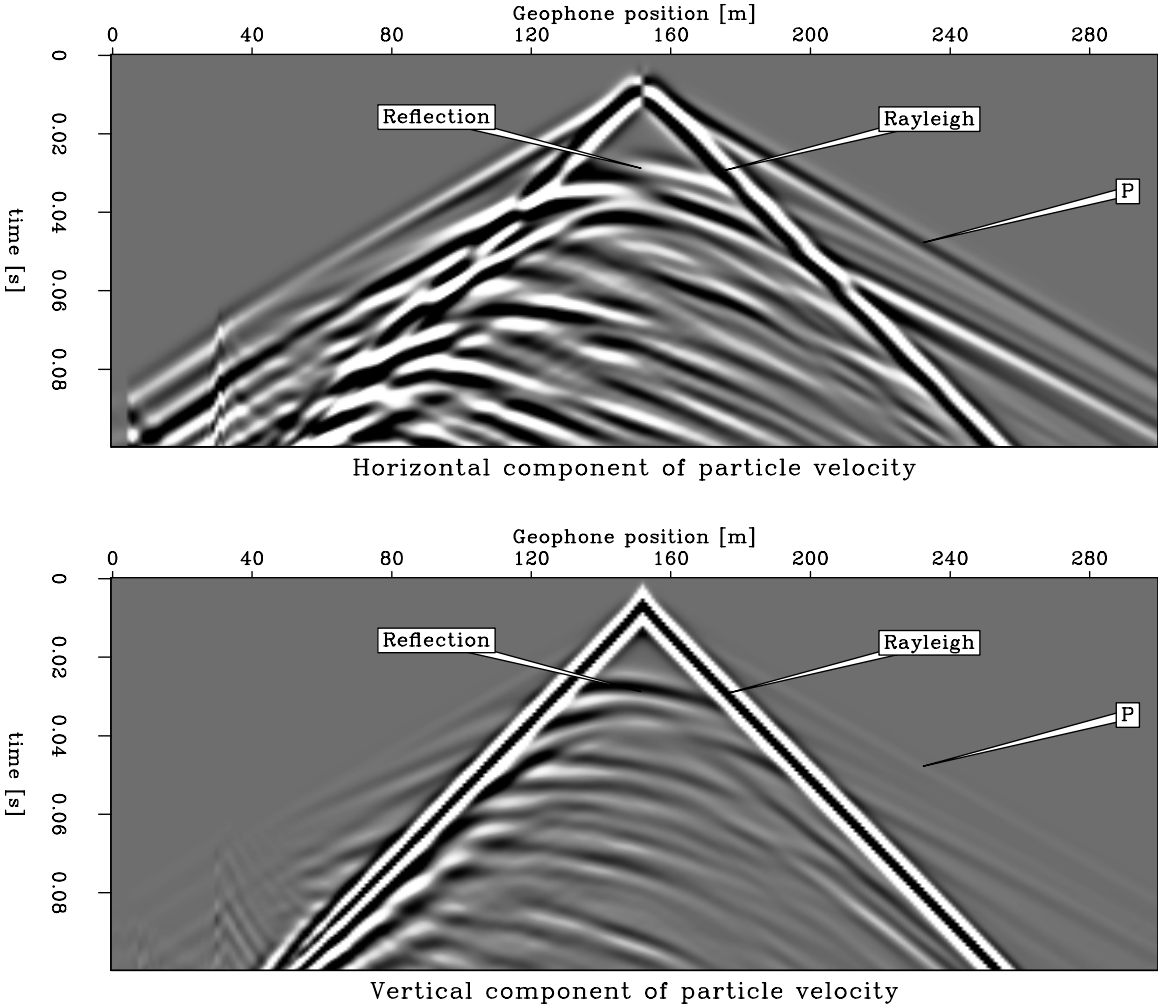


Figure 2: Shot gather for a shot at 150m. In the horizontal component, we can distinguish both the S and P surface wave. In the vertical component, we can only distinguish the Rayleigh wave. At $t = 0.03$ we see the reflections from the top of the coal layer. **[CR]** sjoerd2/. 150sg

tice. With sandstone at the surface, coupling of sources and geophones should be good and attenuation reasonably low. But in areas where dry rocky soil overlies the sandstone, coupling issues and attenuation would almost certainly filter the higher frequencies in addition to significantly impacting the propagating wavefront. A successful survey would require working directly on bare rock (likely gluing geophones into drilled holes) and quite possibly a high frequency source such as a small vibrator. The application of either of those, would significantly increase the field-work effort. Attaining sufficiently high-frequency energy with a hammer is not impossible at this site, but certainly challenging.

The model omits very important known heterogeneity at the site, a key problem that would surely impact the quality of any acquired data. The fissures extending from the burned zones to the surface would undoubtedly have a major impact on propagating wave fields, quite possibly masking the reflections from subsurface features in the burned areas.

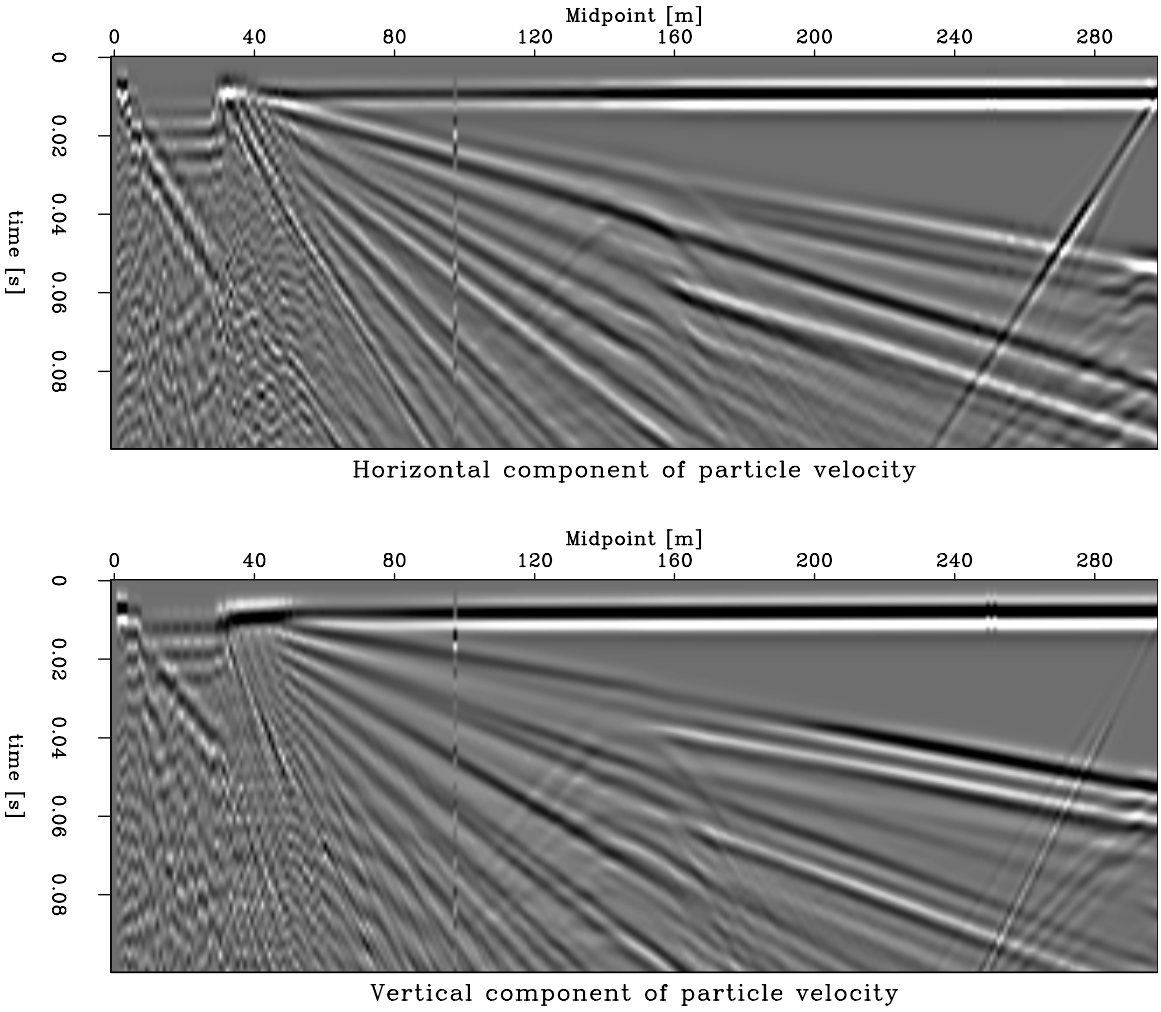


Figure 3: Zero offset gather. Clearly visible is the reflection from the top of the coal/ash layer. Less clearly and almost unseparated is the reflection from the bottom of the coal layer, the bottom of the ash layer is not visible. Weak diffractions from the ash-edge are visible. **[CR]** sjoerd2/. 0co

These fissures are less common (absent?) above the unburned coal, so we are more confident in the validity of our results in that part of the model. But shale layers exist throughout the sandstone in the field area, and could also cause reverberations and spurious reflections in the data. Thus even under ideal circumstances, where source and receiver coupling are optimal and source frequencies exceed 100 or even 200 Hz, the recorded waveforms might be very difficult to interpret.

In addition to the missing fissures, other aspects of the model may be wrong. For instance, we do not really know what parameters are reasonable for the burned coal; we simply divided the coal properties by four. But the ash might be better modeled as a void, which would make a major difference. We could expect stronger diffractions from the edge of the burned zone. But because the reflection from the coal is already quite strong, not much difference is expected between the reflections of the burned and not unburned parts

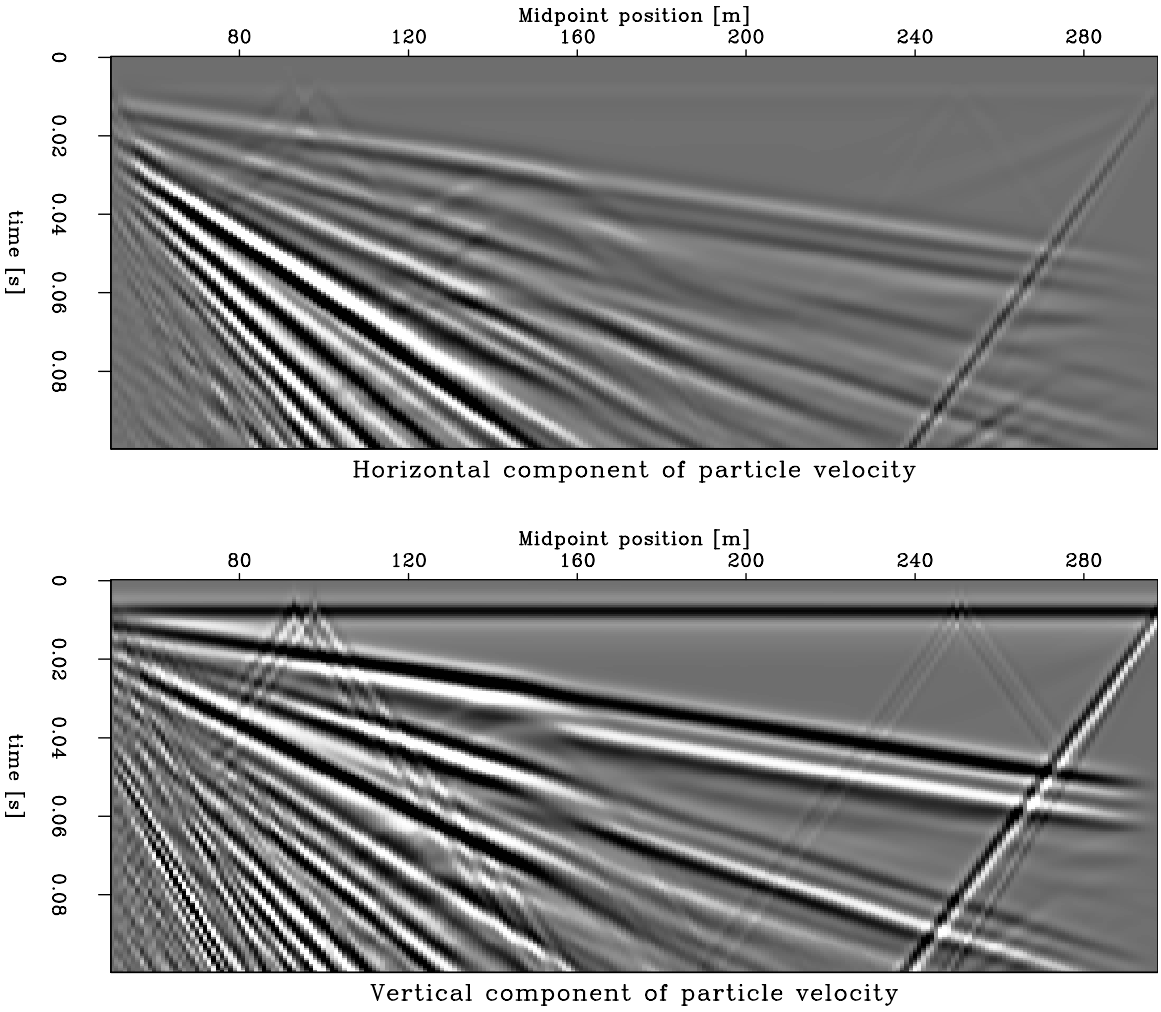


Figure 4: Stacked section, after surface wave removal and NMO correction. The bottom of the coal layer is visible, but clearly overlaps with the reflection from the top of the coal layer. **[CR]** sjoerd2/. ZOsect

of the coal seam. In any case, the known subsidence and rock changes that exist above the burned zone at this site and observed at other such sites (e.g., Wolf, 2006) would surely impact both wave propagation above, and within, the burned zone. The absence of this heterogeneity in the model is a major simplification of the real case.

We have not simulated an SH-wave survey at the site, an option that should be considered in any near-surface application. Being strongly sensitive to voids, shear waves might prove useful in this case. Their lower propagation speed would lead to shorter wavelength and better resolution for a given frequency, but generation of high frequency sources is even harder for SH-wave sources than it is for P-wave sources. In addition, the jumbled nature of the fissured sandstone would likely lead to many spurious S-P conversions that would complicate interpretation of the observed wave field. Accurately simulating the complexities of an SH-wave survey at the site would be difficult.

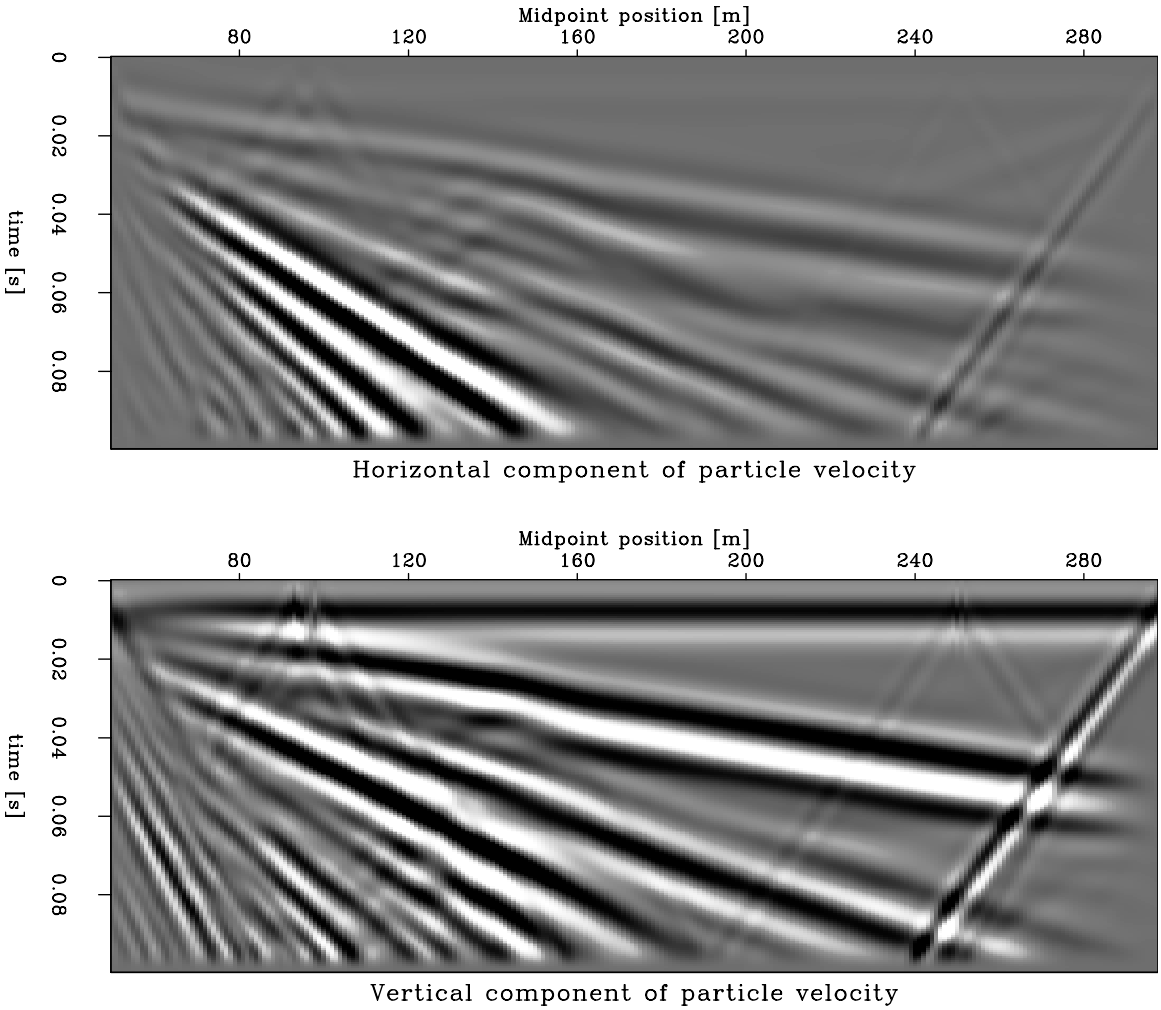


Figure 5: Stacked section, after surface wave removal and NMO correction. Data are bandpass-filtered to 100 Hz to more accurately represent possible data. The bottom of the coal layer, and the difference between coal and ash, is no longer visible. **[CR]** sjoerd2/. ZOsectF

CONCLUSIONS

Our modeling results provide important insights into the feasibility of conducting seismic studies at the Durango coal fire site. The modeling results suggest that under ideal circumstances, P-wave reflection could successfully image the unburned coal and could help to constrain the burned zone. Successful data acquisition would likley require thorough planning and seismic source selection, very careful field procedures including working directly on the rock surface (not on any soil), and great care to ensure high frequencies. A key unknown is the impact that fissures and other near-surface uncertainty would have on the data; these features might completely corrupt the wavefield and render useless data collected even under the most careful procedures.

ACKNOWLEDGMENTS

We are grateful to Bob Clapp for helpful discussions on this topic. We are also grateful to Lynn Orr, Taku Ide, and Nigel Crook for bringing this project to our attention and helping us to understand the key aspects of the problem. And finally, we thank Shawn Larsen for his help with e3d.

REFERENCES

GCEP, 2008, Global energy and climate project, Stanford University: http://www.gcep.org/. (last visited: October 20, 2008).

Larsen, S. C. and J. C. Grieger, 1998, Elastic modeling initiative, Part III: 3-D computational modeling: 68th Ann. Internat. Mtg, 1803–1806, Soc. of Expl. Geophys.

Martin, G. S., 2006, Marmousi2: An upgrade for marmousi: The Leading Edge, **25**, 156–166.

Miller, R. D., S. E. Pullan, D. W. Steeples, and J. A. Hunter, 1994, Field comparison of shallow P-wave seismic sources near Houston, Texas: Geophysics, **59**, 1713–1728.

Miller, R. D., S. F. Pullan, D. W. Steeples, and J. A. Hunter, 1992, Field comparison of shallow seismic sources near Chino, California: Geophysics, **57**, 693–709.

Wolf, K. H. A. A., 2006, The interaction between underground coal fires and their roof rocks.: PhD thesis, Delft University of Technology.

Hubbert math

Jon Claerbout and Francis Muir

ABSTRACT

Hubbert fits growth and decay of petroleum production to the logistic function. Hubbert's relationship is commonly encountered in four different forms. They are all stated here, then derived from one of them, thus showing they are equivalent.

PREFACE

This article was rejected by Wikipedia. They say an encyclopedia should not contain the derivation of important equations so we prepared this article for that purpose. The many references to this paper (missing here) are web links which may be found at http://sep.stanford.edu/sep/jon/tarsand/

THE FOUR FORMS OF HUBBERT'S EQUATION

Over the long haul populations grow and decay. To describe the growth and decay of civilization's dependence on nuclear and fossil fuels, M. King Hubbert chose an equation that describes many natural processes. Introduce bacteria to food and their population will grow exponentially until there no longer is food. As we catch all the fish in the lake our daily catch will be proportional to the number of remaining fish. Hubbert's equation models both exponential growth and decay with a single equation of three parameters to be chosen from the data. He predicted 52 years ago that worldwide oil production would be peaking about now (2008). It is.

Hubbert's math has four different forms which we examine before showing they are mathematically equivalent.

Basic definitions

We define:

- t is time in years
- $Q(t)$ is cumulative production in billion barrels at year t.
- Q_∞ is the ultimate recoverable resource.
- $P(t) = dQ/dt$ is production in billion barrels/year at year t.
- τ is the year at which production peaks.

- ω is an inverse decay time (imaginary frequency).

The Hubbert's equation can be expressed in four forms. First, the differential form

$$\frac{dQ}{dt} = P = \omega\ Q \left(1 - \frac{Q}{Q_\infty}\right) \tag{1}$$

This equation is non-linear in Q but it reduces to familiar linear equations near the beginning and $Q \approx 0$ and near the end at $Q \approx Q_\infty$. As production begins and Q/Q_∞ is small, equation (1) reduces to $dQ/dt = \omega Q$ which displays exponential growth at a rate ω. As production ends near $Q \approx Q_\infty$ the non-linear equation reduces to exponential decay. To prove this fact change variables from Q to q by inserting $Q = Q_\infty - q$. Then evaluate the result at small q. The form (1) exhibits growth and decay as a dynamic process.

The second form of the Hubbert equation is found by dividing equation (1) by Q. It is sometimes called the *Hubbert Linearization*.

$$\frac{P}{Q} = \omega\ \left(1 - \frac{Q}{Q_\infty}\right) \tag{2}$$

The important thing about this equation is that it is linear in the two variables Q and P/Q. If you have historical measurements of P_i and Q_i, you can plot these points in the $(Q, P/Q)$-plane and hope for them to reasonably fit a straight line. Fitting the best line to the scattered points we can read the axis intercepts. At $Q = 0$ with equation (2) we can read off the value of the growth/decay parameter $\omega = (P/Q)_{\text{intercept}}$. For world oil, according to Deffeyes (Remember! The references are links on the web version of this paper.) it is 5.3 percent/year. At the other intercept, $P/Q = 0$ we must have $Q = Q_\infty$. According to Deffeyes, Q_∞ is two trillion barrels.

The third form of Hubbert's equation is the one best known. It looks like a Gaussian, but it isn't. (A Gaussian decays much faster.) The current production $P = dQ/dt$ is

$$P(t) = Q_\infty\ \omega\ \frac{1}{(e^{-(\omega/2)(\tau - t)} + e^{(\omega/2)(\tau - t)})^2} \tag{3}$$

This is the equation of a blob, also known as "Hubbert's pimple", symmetric about the point $t = \tau$. Asymptotically it decreases (or increases) exponentially towards its maximum value at the center at $t = \tau$. The function resembles a Gaussian but exponential decay is much weaker than Gaussian decay. Exponential growth is common in ecological systems which may also decay exponentially as resources are depleted or predator numbers grow exponentially.

All that remains is to figure out τ. The Hubbert curve is symmetrical and reaches its maximum when half the oil is gone. That happens when $Q = Q_\infty/2$. In the case of USA production which has passed its peak we can find the year that Q reached that value (about 1973). There is some debate about what year world production peaks, but general agreement is that it is about now (2008). Under Hubbert assumptions the decline curve is a mirror of the rise curve. That means we start down gently over the next decade, but

about 25 years from now we hit the inflection point and see a 5 percent/year decline every year thereafter.

In real life there is no reason for the decay rate to match the growth rate. The decay could be faster because of horizontal drilling. The decay could be slower because we tax to conserve or successfully invest in technologies. As liquid oil depletes, society is switching to mining tar sands.

The Hubbert equation, in all its forms, follows as a consequence of the definition of the "logistic" function $Q(t)$. It ranges from 0 in the past to Q_∞ in the future.

$$Q(t) = \frac{Q_\infty}{1 + e^{\omega(\tau - t)}} \tag{4}$$

VERIFICATION THE FOUR FORMS ARE EQUIVALENT

If you buy the idea that your data scatter in $(Q_i, P_i/Q_i)$-space is a straight line, then you have bought equation (2). If you buy any one of equations (1),(2),(3), or (4), then you have bought them all because they are mathematically equivalent. Starting from the definition (4) using the rule from calculus that $d(1/v)/dt = -(dv/dt)/v^2$ yields equation (3).

$$\frac{dQ}{dt} = P(t) = Q_\infty\ \omega\ \frac{e^{\omega(\tau - t)}}{(1 + e^{\omega(\tau - t)})^2} \tag{5}$$

$$P(t) = Q_\infty\ \omega\ \frac{1}{(e^{-(\omega/2)(\tau - t)} + e^{(\omega/2)(\tau - t)})^2} \tag{6}$$

which is equation (3).

Equation (4) allows us to eliminate the denominator in equation (5) getting equation (2)

$$P/Q = (Q/Q_\infty)\ \omega\ e^{\omega(\tau - t)} \tag{7}$$

$$P/Q = (Q/Q_\infty)\ \omega\ ((1 + e^{\omega(\tau - t)}) - 1) \tag{8}$$

$$P/Q = (Q/Q_\infty)\ \omega\ (Q_\infty/Q - 1) \tag{9}$$

$$P/Q = \omega\ (1 - Q/Q_\infty) \tag{10}$$

which is equation (2). Multiplying both sides by Q gives equation (1).

REFERENCE

1. http://www.hubbertpeak.com/hubbert/1956/1956.pdf contains Hubbert(1956, M.King Hubbert, Nuclear energy and fossil fuels, Publication 95, Shell Oil Company.

2. Kenneth S. Deffeyes, 2006, Beyond Oil: The view from Hubbert's Peak: Hill and Wang ISBN 0-8090-2957-X.

POSTFACE

One day I learned that Firefox had a much better way of zooming web pages, zooming the pictures too. Knowing that equations are pictures I went to Wikipedia, and looked up "Fourier Analysis". I was delighted. A table of equations looked beautiful and could be zoomed up to a size suitable for public lectures! It was as if html had finally incorporated math. In reality the math had been done via LaTeX and inserted as photos. Wanting to have on-line lectures drawn exactly from my books I learned to contribute to Wikipedia including equations.

At the same time I was reading Deffeyes book "Beyond Oil" (a play on the slogan "Beyond Petroleum"). I wanted to play with Hubbert's curve fitting of worldwide oil production. Francis Muir gave me the algebraic tips I needed. I prepared my contribution in my "sandbox" and then moved it to the main encyclopedia. One of their volunteer managers soon found it and didn't like it. Rather than quote his opinions, I paraphrase saying "equation derivations do not belong in an encyclopedia."

So, I gave up and prepared this PDF file instead. It's not as seamlessly web viewable as html, but I'm much happier with it – and I am able to include it in this report!

A goal I cannot meet today is to write a single LaTeX file that becomes two things: (1) print media that is attractive, readable, and contains its web references (perhaps in footnotes), and (2) a PDF file with references as web links that work. I would like to see my books and all SEP report articles in this form. I'm not going to develop this myself. Someone else will do it; and I'll try to be an early adopter.

Tar sands: Reprieve or apocalypse?

Jon Claerbout

ABSTRACT

Based on a Hubbert-type analysis two projections are made of tar sands production. With tar sand production growing at 5%/year total petroleum production declines at an annual rate of 1-2%. With tar sand production growing at 10%/year total petroleum production continues rising at almost the historic rate until 2040 followed by a catastrophic rate of 50%/decade.

PREFACE

This paper with references as embedded web links may be found at http://sep.stanford.edu/sep/jon/tarsand/

TAR SANDS

Tar Sands are an amazing resource, mostly in Canada. Already in 2007, 44% of Canadian petroleum production came from tar sands. The total Canadian resource (not all producable) is estimated to be 2.2 trillion barrels. That's a lot! For comparison producible world petroleum liquids are estimated at 2 trillion barrels, half of which have already been consumed. Below are two scenarios for future tar-sands production. The plot on the left (labeled Malthusian) shows tar sands growing and decaying at the same 5% annual rate that liquid petroleum has done. The decay of the sum is not severe, roughly 15% per decade, 1-2% per year. The plot on the right (labeled Cornucopian) shows what might happen in case of laissez-faire or perceived grave urgency, a 10% tar-sands growth rate for 30-35 years followed by precipitous collapse about 50% per decade.

Unfortunately, of the Canadian 2.2 trillion barrels only 20% is estimated to be recoverable by strip mining. Having myself seen that the residual output of a tar sand extraction plant is clean white sand, we may take that resource to be 100% fully recoverable, 440 billion barrels. Hooray! Unfortunately, what is not reachable by strip mining is more difficult to recover. I took this to be 25%, so that 80% adds 420 billion barrels.

You may hear, "Tar sands are profitably produced at $28-$44/barrel". Yes, and there may be oil in the Middle East that can be produced at $1-$2/barrel. That oil will be sold for the "market clearing price". Even if someone could rapidly produce great volumes from tar sands it would only accelerate the subsequent decline (as the Cornucopian plot shows). Tar sand production requires monsterous industrial machinery. "Obscene profits" are building it now.

We neglect here many things. Natural gas is an important resource about the size of oil with a peak trailing oil by a few decades. Tar sands will need to replace the

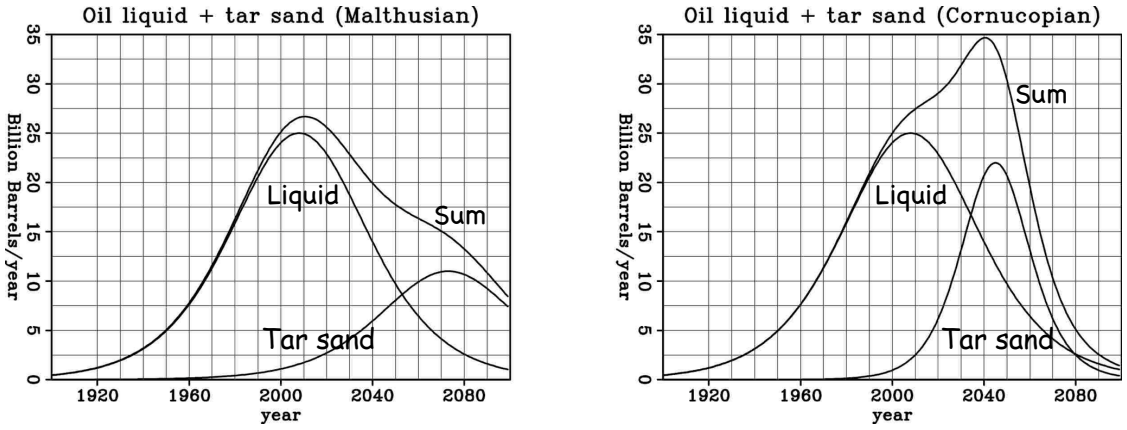


Figure 1: Two scenarios for tar sands exploitation.[NR] jontarsand/. tarannotated

natural gas too! Venezuela has a reserve of heavy oil possibly comparable to Canada's. Past nationalization of investments has effectively stopped current investment, but the oil in place has been quoted as 1.3 trillion barrels with a 20% recovery, netting 260 billion barrels – a bump a quarter the size of Canada occuring later in time. Shale oil can be burned like coal. Unfortunately liquifaction of shale oil and world-wide gas hydrates are nice to dream about, but there are no functioning demonstration production facilities anywhere. Recently USGS forecast 1/10 trillion barrels in the arctic, a pimple on the graphs above. Nuclear is not a transportation fuel. Coal (with its own Hubbert curve) can be converted to the gaseous fuel dimethyl ether (DMD) particularly suitable as a replacement for diesel fuel. Ironically, nuclear, including the fast breeder reactor, including thorium as well as uranium is more fully described by the old-time geologist Hubbert in 1956 than by today's Cal Tech physicist and vice provost David Goodstein in his 2004 book "Out of gas". http://www.usgs.gov/newsroom/article.asp?ID=1980 http://www.hubbertpeak.com/hubbert/1956/1956.pdf

WILL THIS AFFECT YOU?

You may ask, "What makes you believe these predictions?" The classic paper of M.King Hubbert, written in 1956 predicted a USA liquid oil peak in 1973. It did peak then. That paper also predicted a world peak about the year 2000, a bit early, but not much. Most recent predictions [1] [2] calculate twice as much oil as Hubbert predicted, but the peak has moved only about a decade later, about now. Why does double the oil move the peak so little? Because so much is depleted in the middle decades. http://www.amazon.com/Beyond-Oil-View-Hubberts-Peak/dp/080902957X/

Didn't you assume the total oil in place? Not for the liquids. When the production peaks, half is gone; the other half is left. There is widespread agreement the production will soon or has already peaked. For the tar sands resource size I refer to The Petroleum Society of Canada. http://www.energy.gov.ab.ca/OilSands/pdfs/RPT_Chops_app3.pdf

The next question is, "Does it matter to me?" I suggest you print this page. Upon the printed page draw a horizontal arrow from the day you were born to then plus 84 years.

Then draw the arrows for your parents, your children, and grandchildren.

Poor people will be the first to conserve. After that people with long commutes. After that comes you. You have some time prepare yourself before the crowds. Don't waste it! How much time do we have? If the choice is political between the Malthusian and the Cornucopian which choice do you think will be made? I fear the disaster. (Sorry.) Would that the Canadians learned the lesson of the East Texas Oil Field and the Texas Railroad Commission.

CREDITS

Helpful comments were received from Roland Horne, David Lumley, Amos Nur, George Sibbald, and Kevin Wolf.

CONFLICT OF INTEREST STATEMENT

The author is an emeritus (retired) professor specialized in the geophysical prospecting industry, an industry dependent for customers on the oil industry (but little on its tar sands sector).

SEP PHONE DIRECTORY

Name	Phone	Login Name
Al Theyab, Abdullah	723-0463	altheyab
Ayeni, Gboyega	723-6006	gayeni
Berryman, James	–	berryman
Biondi, Biondo	723-1319	biondo
Cardoso, Claudio	723-1250	claudio
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
de Ridder, Sjoerd	723-1250	sjoerd
Guitton, Antoine	–	antoine
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Li, Elita	723-9282	myfusan
Maysami, Mohammad	723-9282	mohammad
Shen, Xukai	723-0463	xukai
Shragge, Jeff	723-0463	jeff
Tang, Yaxun	723-1250	tang
Wong, Mandy	723-9282	mandyman

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*".

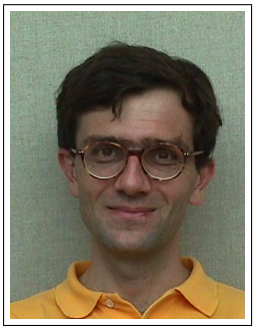
WORLD-WIDE WEB SERVER INFORMATION

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.

Research Personnel

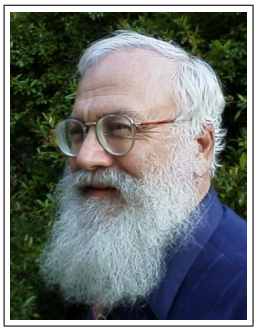
Abdullah Al Theyab graduated from the University of Tulsa with a double major in Geosciences and Computer Science. In 2005, he joined the Exploration Organization in Saudi Aramco working in seismic data processing and interpretation. Currently, he is pursuing a M.S. degree in Stanford Exploration Project.

Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell sholar 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.



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.

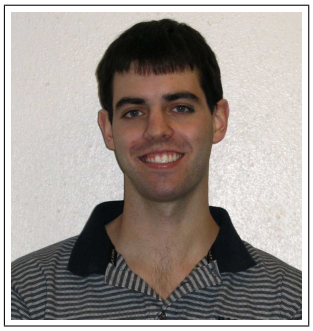


Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.

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.

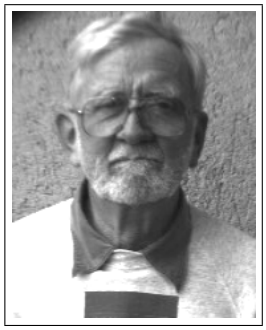
Seth Haines received a B.A. in Geology and Physics (double major) from Middlebury College in 1997, and began graduate work in the Stanford Geophysics Department in 1998 studying the Tibetan crust with seismic methods. He received his M.Sc. in 2001, and a Ph.D. in 2004 (SEP119) with a thesis titled "Seismoelectric imaging of shallow targets". He is currently a Research Geophysicist with the USGS Central Energy Resources Team in Denver, working on near-surface seismic methods for various applications including gas hydrates.

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.

Francis Muir has a degree in Mathematics from Balliol College, Oxford. In 1954 he joined Seismograph Service, where he worked in research and on reflection and refraction crews in England, North Africa, West Africa, Papua and Western Australia. He joined West Australian Petroleum in 1962, where he supervised seismic crews around Exmouth Gulf, the Great Sandy Desert and Barrow Island, and in 1967 transferred to Chevron Oil Field Research Company. He left Chevron as a Senior Research Associate in 1983. Francis has an appointment as Consulting Professor at Stanford, and also consults with industry on various aspects of exploration. He is a Fellow of the Royal Astronomical Society, and is a member of the SEG Research Committee.



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 persuing a Ph.D. in Geophysics at Stanford University.

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.

SEP ARTICLES PUBLISHED OR IN PRESS

Ayeni, G., A. Huck, and P. de Groot, 2008, Extending reservoir property prediction with psedo-wells: First Break. **26**, 11–15.

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.

de Ridder, S., Prieto, G.A., 2008, Seismic Interferometry and the Spatial Auto-Correlation Method on the Regional Coda of the Non-Proliferation Experiment: AGU Fall Meeting, S31A-1885.

Gray, S., D. Trad, B. Biondi, and L. Lines, 2006, Towards wave-equation imaging and velocity estimation: Recoder, **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.

Halpert, A., Clapp, R.G., Lomask, J. and B. Biondi, 2008, Image segmentation for velocity model construction and updating: SEG Technical Program Expanded Abstracts **27**, 3088-3092.

Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3D salt boundaries: Geophysics, in press.

Maysami, M., and F. J. Herrmann, 2008, Lithology constraints from seismic waveforms: application to opal-A to opal-CT transition: SEG Technical Program Expanded Abstracts **27**, 2011-2015.

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.

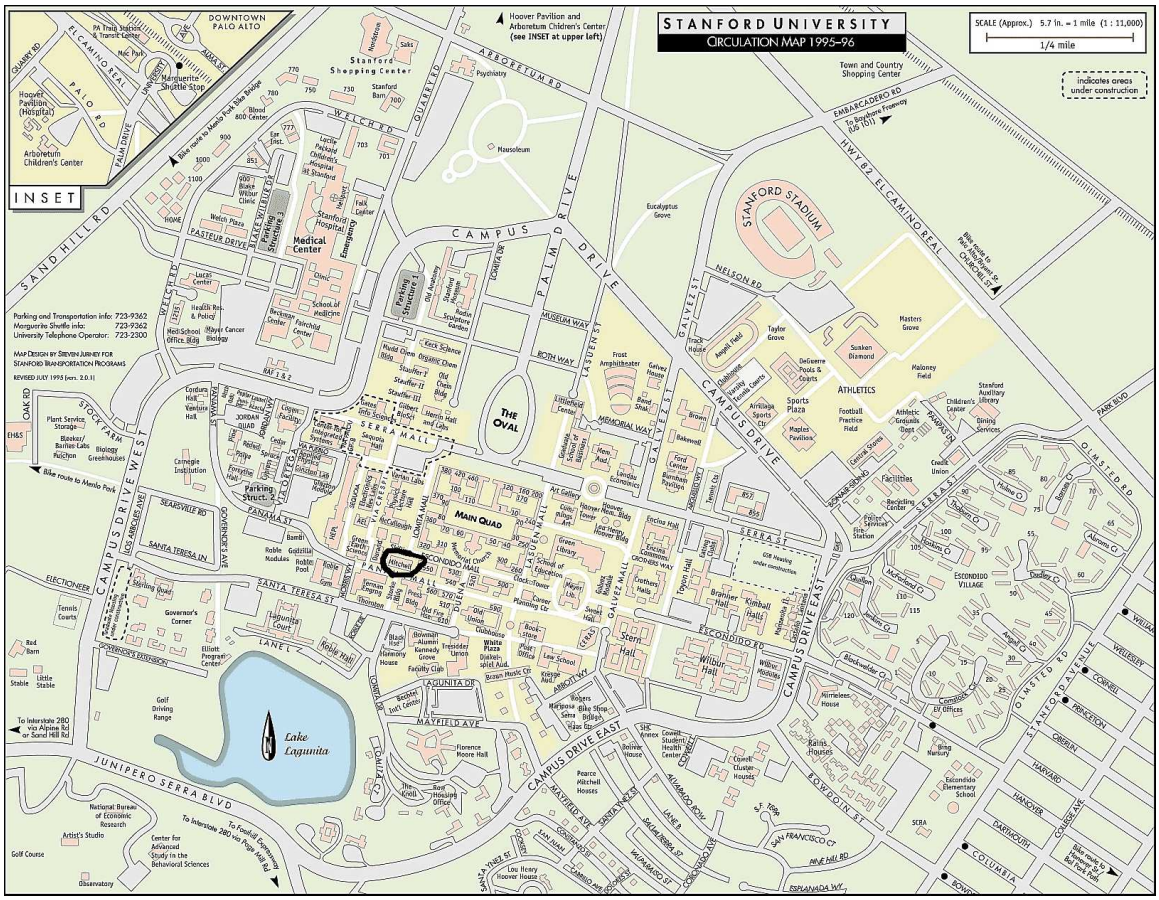
Shragge, J. and G. Shan, 2007, Prestack depth migration in elliptic coordinates: SEG Technical Program Expanded Abstracts, **26**, 2185–2189.

Tang, Y. and R. G. Clapp, 2006, Selection of reference-anisotropy parameters for wavefield extrapolation by Lloyd's algorithm: SEG Technical Program Expanded Abstracts **25**, 189-193.

Tang, Y., 2007, Selective stacking in the reflection-angle and azimuth domain: SEG Technical Program Expanded Abstracts **26**, 2320-2324.

———, 2008, Wave-equation Hessian by phase encoding: SEG Technical Program Expanded

Abstracts **27**, 2201-2205.
Valenciano, A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35.


STANFORD UNIVERSITY
CIRCULATION MAP 1995–96
SCALE (Approx.) 5.7 in. = 1 mile (1 : 11,000)
1/4 mile
indicates areas under construction
Hoover Pavilion and Arboretum Children's Center (see INSET at upper left)
INSET
DOWNTOWN PALO ALTO
Marguerite Shuttle Stop
Stanford Shopping Center
Town and Country Shopping Center
Medical Center
Stanford Hospital
Emergency
Mausoleum
Eucalyptus Grove
STANFORD STADIUM
ATHLETICS
THE OVAL
MAIN QUAD
Lake Lagunita
Golf Driving Range
Governor's Corner
Tennis Courts
Stern Hall
Wilbur Hall
Kimball Hall
Branner Hall
Toyon Hall
Lagunita Court
Florence Moore Hall
Escondido Village
Escondido Elementary School
Kairos Houses
Synergy Houses
Cowell Cluster Houses
Cowell Student Health Center
Pearce Mitchell Houses
Driver House
Lou Henry Hoover House
Sloan Management Bldg
Graduate School of Business
Law School
Crothers Halls
Parking Structure 1
Parking Structure 2
Parking Structure 3
Parking Structure 4
Parking Structure 5
Parking Structure 6
Parking Structure 7
Parking Structure 8
Parking Structure 9
Parking Structure 10
Parking Structure 11
Parking Structure 12
Parking Structure 13
Parking Structure