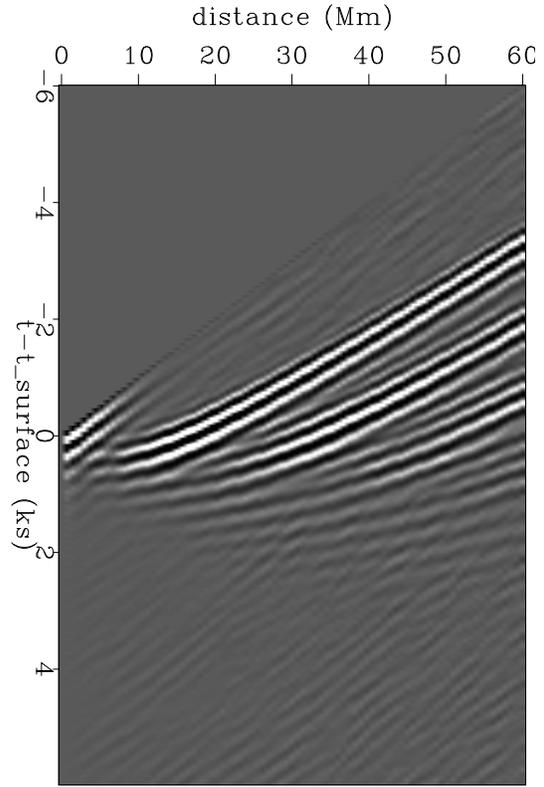
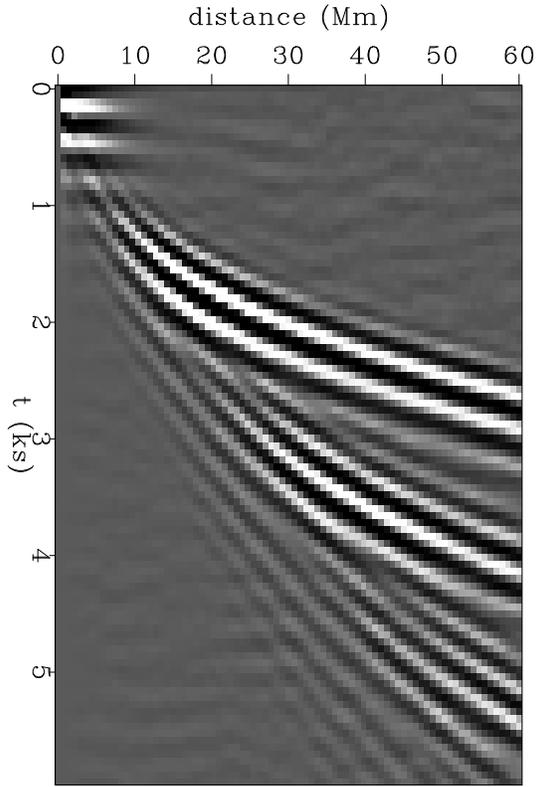


STANFORD EXPLORATION PROJECT

Tariq Alkhalifah, James Berryman, Biondo Biondi, Morgan Brown, Jon Claerbout, Robert Clapp, Sean Crawley, Joe Dellinger, Sergey Fomel, Antoine Guitton, Diane Levin, Stewart Levin, Francis Muir, Marie Prucha, Jianliang Qian, James Rickett, Paul Sava, Mamta Sinha, Yalei Sun, William Symes, Louis Vaillant, Herbert Wang, and Yi Zhao

Report Number 100, May 1999



Copyright © 1999

by the Board of Trustees of the Leland Stanford Junior University

Stanford, California 94305

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

Preface

The electronic version of this report¹ makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR], are promises by the author about the reproducibility of each figure result.

Reproducibility is a way of organizing computational research, which allows both the author and the reader of a publication to verify the reported results at a later time. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

ER denotes Easily Reproducible and are the results of a 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. 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 SEP latex package). 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.

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. SEP staff have only attempted to make sure that the makefile rules exist and the source codes referenced are provided. Some of the reason for the CR designation are: a proprietary data set, a parallel computer, a large amount of processing time (20 minutes or more), or commercial packages such as Matlab or Mathematica.

NR denotes Non-Reproducible. This class of figure is considered non-reproducible. SEP discourages authors from flagging their figures as NR except for artist drawings, scanings, etc.

Our testing is currently limited to IRIX 6.5 and LINUX 2.1, but the code should be portable to other architectures. Reader's suggestions are welcome. For more information on reproducing SEP's electronic documents, please visit

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

Jon Claerbout, Biondo Biondi, Robert Clapp, and Sergey Fomel

¹<http://sepwww.stanford.edu/public/docs/sep100>

SEP-100 — TABLE OF CONTENTS

Velocity

<i>Claerbout J.</i> , Everything depends on $V(x,y,z)$	1
<i>Biondi B. and Sava P.</i> , Wave-equation migration velocity analysis	11
<i>Clapp R. G. and Biondi B.</i> , Preconditioning tau tomography with geologic constraints	35
<i>Clapp R. G. and Biondi B.</i> , Why tau tomography is better than depth tomography	51
<i>Fomel S.</i> , Velocity continuation in migration velocity analysis	59
<i>Symes W. W.</i> , All stationary points of differential semblance are asymptotic global minimizers: layered acoustics	71
<i>Symes W. W.</i> , Coherent noise suppression in velocity inversion	93

Imaging

<i>Prucha M. L., Biondi B. L., and Symes W. W.</i> , Angle-domain common image gathers by wave-equation migration	101
<i>Biondi B.</i> , Subsalt imaging by common-azimuth migration	113
<i>Vaillant L. and Biondi B.</i> , Extending common-azimuth migration	125
<i>Sinha M. and Biondi B.</i> , Comparing Kirchhoff with wave equation migration in a hydrate region	135
<i>Fomel S. and Prucha M.</i> , Angle-gather time migration	141
<i>Sava P.</i> , On Stolt prestack residual migration	151
<i>Sun Y.</i> , Anti-aliasing multiple prediction beyond two dimensions	159

Helix filtering

<i>Rickett J. and Claerbout J.</i> , Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring	171
<i>Crawley S.</i> , Interpolation with smoothly nonstationary prediction-error filters	181
<i>Clapp R. G., Fomel S., Crawley S., and Claerbout J. F.</i> , Directional smoothing of nonstationary filters	197

<i>Brown M.</i> , Texture synthesis and prediction error filtering	211
<i>Claerbout J.</i> , Polarity and PEF regularization	223
<i>Sava P. and Fomel S.</i> , Spectral factorization revisited	227
<i>Zhao Y.</i> , Helix derivative and low-cut filters' spectral feature and application	235
<i>Claerbout J.</i> , Helical meshes on spheres and cones	251

Traveltimes

<i>Sava P.</i> , 3-D traveltime computation by Huygens wavefront tracing	257
<i>Qian J. and Symes W. W.</i> , An adaptive finite difference method for traveltime and amplitude	269
<i>Rickett J. and Fomel S.</i> , A second-order fast marching eikonal solver	287

Optimization

<i>Guitton A. and Symes W. W.</i> , Robust and stable velocity analysis using the Huber function	293
<i>Symes W. W.</i> , Extremal regularization	315

Reservoir Characterization

<i>Muir F.</i> , Backus revisited: Just in time	331
<i>Berryman J. G. and Wang H. F.</i> , Elastic wave propagation and attenuation in a double-porosity dual-permeability medium	335

Anisotropy

<i>Alkhalifah T.</i> , A short tour through the Stanford Exploration Project contributions to anisotropy	361
--	-----

Reunion

<i>Levin S. A., Levin D. P., and Dellinger J.</i> , 25 for 100: Remembrances of the SEP 25th Anniversary Reunion	365
--	-----

SEP article published or in press, 1998-99	379
SEP phone directory	381
Research personnel	383
SEP sponsors for 1998-99	391

Everything depends on $V(x,y,z)$

*Jon Claerbout*¹

ABSTRACT

Estimating 3-D velocity $V(x,y,z)$ is the most important problem in exploration geophysics. It is a very difficult problem. In order to really solve it, SEP has turned to fundamentals of estimation theory with topographic data, regridding, interpolation, truncation, erratic noise, instrument drift, etc. This return to fundamentals has proven rewarding, leading us to the helix discovery. This discovery is revitalizing wave equation migration in 3-D, preconditioning many estimations (big speed up), and regularizing velocity estimation (blends measured and prior information). To enable young people to become productive with 3-D seismic data, Biondo Biondi and Bob Clapp have built a 3-D seismic software infrastructure that is able to address real 3-D problems, such as $V(x,y,z)$ and aliasing in 3-D. This infrastructure is unique in the academic world. None of the other academic organizations have enough computing power and infrastructure to allow routine research activities with 3-D field data.

INTRODUCTION

Reflection seismology gives us magnificent volumetric images of the earth's interior. Image construction is a two-stage process. The first stage estimates the three-dimensional velocity. The second stage uses it to make the reflectivity cube. Both reflectivity estimation and velocity estimation are inverse problems. For most datasets, reflectivity would be better determined than velocity, if somebody gave us the velocity function. But nobody does. Velocity estimation remains an art, a mixture of every skill we have and more. It requires ideas we haven't coded yet and it requires ideas we haven't yet concocted. We have no chance to estimate velocity reliably if we do not honor the data complexity when migrating. We need better migration methods when the data are poorly sampled and the velocity function is rapidly varying. I claim that SEP needs to avoid being distracted by the ultimate goals of the petroleum community. We need to solve particular problems in estimating the earth's 3-D velocity (and reflectivity). Everyone else depends on us to get it right.

¹email: jon@sep.stanford.edu

AVOIDING CHIMERAS

Some would say that seismic structural imaging is finished. They would say the future lies in linking the seismic image with porosity or permeability. They would interpret the AVO (Amplitude Verses Offset) to estimate it.

I was “present at the creation.” I was a consultant to Chevron Overseas Inc. when Bill Ostrander did his founding work on AVO. People asked me then whether I didn’t feel a conflict, whether I didn’t want to pursue this exciting new area at the university. After all, I was in on the ground floor. My answer was no and remains no. Even at the beginning we quickly learned that **AVO often fails**. How often does it fail? At a university we cannot know. I suppose some companies keep good records of their AVO failures while others forget their failures as quickly as they can. The reason I didn’t want to jump on the AVO bandwagon is that universities don’t drill holes. We would never know if we were right or wrong.

Since then we have gotten a pretty clear idea why AVO often fails. We learned a lot from the PhD thesis of Einar Kjartansson. I took his key ideas and results for my textbook, *Imaging the Earth’s Interior*. Both Einar’s thesis and my book are freely available at the SEP web site. Maybe I’ll attach a few of Einar’s results here to remind you. The bottom line is this: After the seismic waves encounter some $V(x, y, z)$ which isn’t merely $v(z)$, the AVO analysis becomes extremely tricky.

Simply stated, the goal of reservoir characterization is to depict the extent of the reservoir, mathematically this is the lateral heterogeneity of the reservoir. But to do this from the earth’s surface, we need to be able to compensate for the lateral heterogeneity everywhere above the reservoir.

A bright young man recently showed us his innovative statistical analysis leading to a 3-D color display of clay probability $P_{\text{clay}}(x, y, z)$. I was impressed. (I’m from the old school where the weather man predicts whether it will rain. He doesn’t tell us the *probability* of rain.) We asked the young man, “how did you find the impedance?” He replied that he had used a commercial package from a very respected firm. I do respect that firm so they must have a disclaimer somewhere that says, “garbage in, garbage out”. What happened to the young man is that he had a little carbonate above his reservoir. His image at small offset was different from that at larger offset. Not merely the Amplitude was different. The whole blessed image was different. The impedance program failed to tell the young man that his data was unsuitable for impedance estimation. Who’d want to purchase a program that failed to produce its product? Better to sell it with some fine print.

Some companies envision us as the ones who should solve the problem and pass the solution along to the contractors. An alternative would be for us to set our sights lower and try to establish standards for estimating the quality of other people’s impedance estimates.

As I said, an academic like me cannot really give a fair and balanced picture of AVO. I want to give you a fair picture that is not wholly my own opinion. The biggest bloodbath in corporate research this year might tell us something. I was told that everyone in the imaging group got offers (though most spurned them), the rock physics group survived (though its lab

is in moth balls), but everyone in the AVO group got fired (“got the package”). I was surprised by this result. “Doesn’t the AVO group produce the much sought-after porosity?” I asked. The answer given was this: “AVO is easily over sold”. So there you have it. In industry there is accountability, which is absent at universities. We play poker for pennies, but they gamble their careers. I’m agreeable if some of my students want to work on AVO, but only if they are among the strongest students.

Two-dimensional marine data often has regular data coverage. Land data and three-dimensional data never really does. To be honest about our day-to-day activities with seismic amplitude, we do not spend our time making small and subtle measurements of amplitude. We spend our time preventing massive amplitude fluctuations that are associated with irregular data coverage. We often spend our time trying to hide the data-acquisition footprint.

Never-the-less, AVO remains an active area of industrial activity and research. There must be successes and those successes may be more secret than the failures. Today AVO may give us rock properties in simple situations (e.g. $v(z)$). One of our challenges is to make those estimates reliable in more complex areas. We know that those estimates are a *very sensitive* function of $V(x, y, z)$. That’s my claim. Everything depends on $V(x, y, z)$.

VELOCITY ESTIMATION IS NOT EASY

Where is SEP going? Alumni have been known to comment that SEP looks unfocused. I think it was my good friend Rick Ottolini who said we look like an oil-company research lab from the 1980’s. They see me playing with radar images of volcanos, bottom-sounding surveys from deep ocean ridges, and some badly recorded depth soundings of the Sea of Galilee. Is Claerbout oblivious to the industry and the changes in the world? Here is why we are going off in these seemingly irrelevant directions.

Oil men think porosity and permeability are everything. Academics believe they can solve all problems with “inversion”. We academics are way ahead of the oil men in our high count of failures, both reported and unreported. I have had many more failures than successes. Successes are rare. Likewise, all the people I most respect have reported many failures. Still we persist, despite grave frustrations. “Inversion” (I prefer to call it “estimation”) is a systematic approach to problems. It is not merely a black art. There are theories, claims, counter claims, etc. And many of these claims can be tested. Because of my experience with a high percentage of failures I decided to start young people off with simple problems where the meaning of failure is self evident (almost). I chose topographic analysis. Putting data on a regular mesh. Filling in gaps. Segregating signal and noise. That’s why we have Galilee, Vesuvius, Madagascar, Fernandina, and other nonpetroliferous data everywhere on display around here. We are learning. And the tools we are building carry straight over to billion dollar industrial problems.

The first billion dollar problem I have in mind is insufficiently dense 3-D seismic data acquisition. Spatially aliased multiples stack into primaries in marine data. This problem is getting worse as more streamers are added to multi-streamer acquisitions. Sean Crawley

and I are successfully interpolating multi-streamer data by estimating and inverting prediction error filters. Poor sampling lowers resolution and creates imaging artifacts in land data. Nizar Chemingui and Biondo Biondi are generating high-resolution images from under-sampled land data by inverting an imaging operator called Azimuth Moveout. These two applications to a billion dollar seismic problems arose from estimation methodologies and tools developed using non-seismic data!

The past two years have brought me the magical helix. This discovery (brought to us by our topographic toys) has spun off projects in three different areas: (1) revitalizing wave-equation migration in 3-D, (2) preconditioning estimations (big speed up), and (3) regularizing velocity estimation (blending measured with prior information). The amazing thing about the helix (as embodied in Sergey Fomel's F90 library) is that any program that does something nifty with two-dimensional data, immediately with no change does the same nifty thing on three- or four-dimensional data.

SEP got started thanks to wave-equation migration. When 3-D came along, for a while SEP lost its competitive edge in migration, as Kirchhoff methods seemed the only way to go. But now the advantages of wave-equation imaging are luring the industry again. We made important progress in that direction; both in the basic technology for wavefield extrapolation and in methods to migrate 3-D marine data. James Rickett and Sergey Fomel worked with me on implicit solutions of the 3-D one-way wave equation that exploit the power of the helix. Biondo Biondi developed common-azimuth migration that drastically reduces the cost of 3-D wave-equation prestack migration and yields encouraging subsalt images when compared with Kirchhoff images.

The dream of estimating velocity by wave-equation methods is getting closer thanks to the recent progress in migration-velocity analysis achieved by Biondo Biondi in collaboration with Paul Sava. Bob Clapp cleared the path by sharpening our tools to handle the nonlinear and undetermined nature of velocity estimation. My title, "Everything depends on $V(x, y, z)$ " says the world is three dimensional. This is where SEP is making its mark, uniquely among university consortia. This is our path.

INFRASTRUCTURE BUILDING

We have been criticized for spending a lot of energy building our infrastructure. By this I mean building software packages, SEPlib, vplot, SEP3D, ratfor90, our reproducible research makefile rules, and Sergey's optimization library. Unlike SU, which was developed at CSM, our software has hardly defined industry standards, nor is it widely used outside SEP (except for Joe Dellinger's vplot which seems to show up in nearly every issue of *Geophysics*). We are not here to produce widely used software, though that is gratifying when it occurs (Rick Ottolini's movie program is another example). We build our infrastructure to get our work done, and to pass along reusable results of our older generation to the next. We do use some commercial packages, and despite hefty discounts, I don't think we get our money's worth from them. I wish we were paying on a "per use" basis. Although some powerful past personages here at SEP have done heroic 3-D work with little more than our SEPlib, vplot, and a

bare-bones Fortran90 compiler, propagating those skills to our broader population of students requires some infrastructure. I'm very happy (absolutely thrilled) that Biondo Biondi and Bob Clapp have built SEP3D for us.

COOPERATION ACROSS DISCIPLINES

We at SEP are not going to give up our relationship with rock physics and geostatistics. Over the years several PhD theses have sprung up from collaborations between SEP and SRB. I expect more of such joint efforts to bear fruit in the future. I recently made a long presentation to the Petroleum Engineering faculty on my work and my perception of its relation to theirs. Andre Journel organized a well-attended reciprocal venture. They find they do not have enough information to produce a unique model, so they produce large numbers of models with the desired statistical properties. Perhaps we should do the same. Instead of providing other geoscientists with a volumetric earth image, we might produce a zillion image cubes (each differing by members of the null space in our best estimate of $V(x, y, z)$). I'm joking of course. They don't want a zillion image cubes. They want us to get a single "right" value for the velocity and give them the "right" cube. Everything depends on it.

KJARTANSSON'S RESULTS

The primary reference for Kjartanson's results is his doctoral dissertation available as SEP report 23. Electronically, this thesis is at

<http://sepwww.stanford.edu/theses/sep23/>

A "digested" version of the thesis is in chapter 4 of my textbook IEI at

<http://sepwww.stanford.edu/oldreports/sep40/>

The particular section is available in html at

http://sepwww.stanford.edu/sep/prof/iei/ofs/paper_html/node6.html

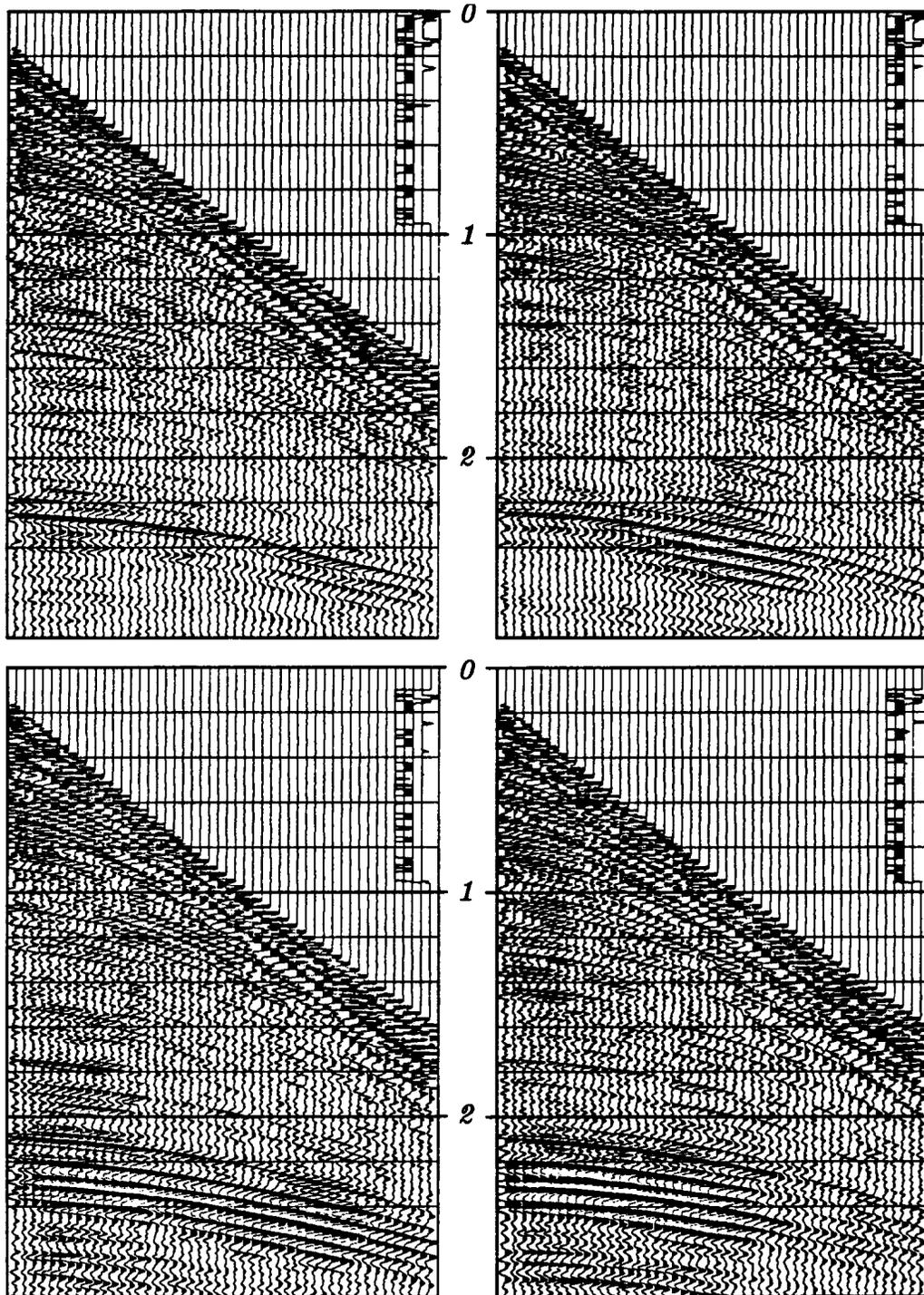


Figure 1: Top left is shot point 210; top right is shot point 220. No processing has been applied to the data except for a display gain proportional to time. Bottom shows shot points 305 and 315. The AVO is too large for realistic layered media models and it changes rapidly from one midpoint to another one nearby. [jonl-kjcmg] [NR]

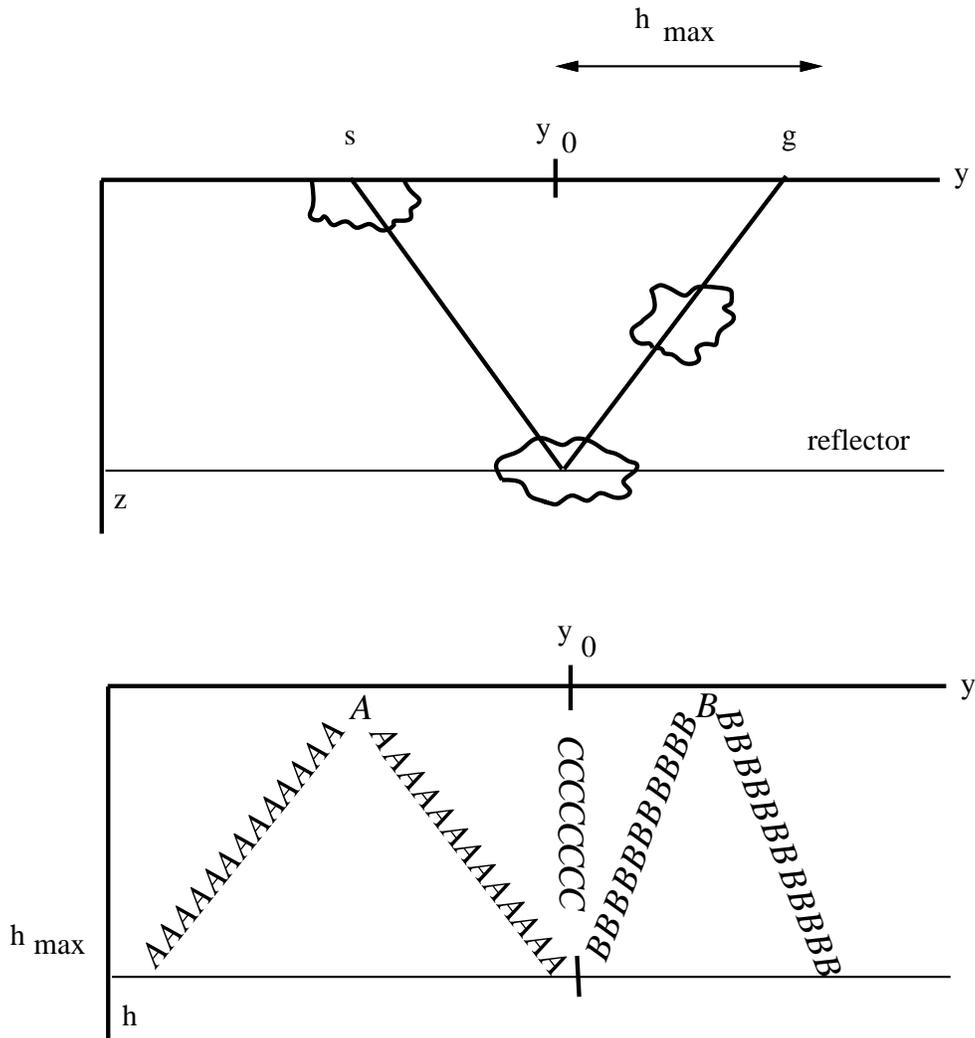


Figure 2: Kjartansson's model. The model on the top produces the disturbed data space sketched below it. Anomalous material in pods A, B, and C may be detected by its effect on reflections from a deeper layer. Midpoint is y . Offset is h . jon1-kjidea [NR]

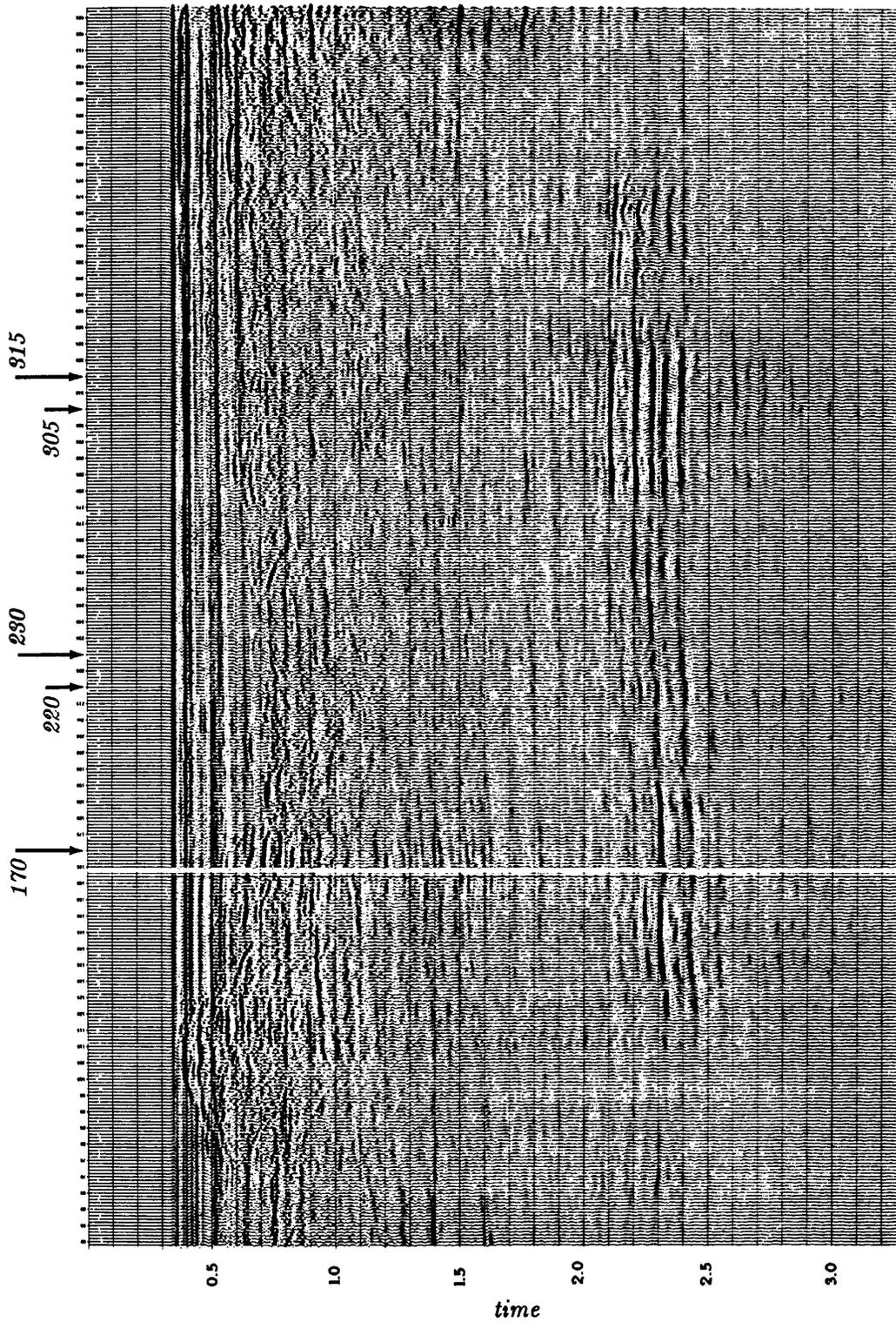


Figure 3: A constant-offset section across the Grand Isle gas field. The offset shown is the fifth from the near trace. (Kjartansson, Gulf) Notice that anomalous amplitudes are not limited to thin “reservoir thickness” zones. They tend to expand over the time axis. Notice a missing trace. [jon1-kjcos] [NR]

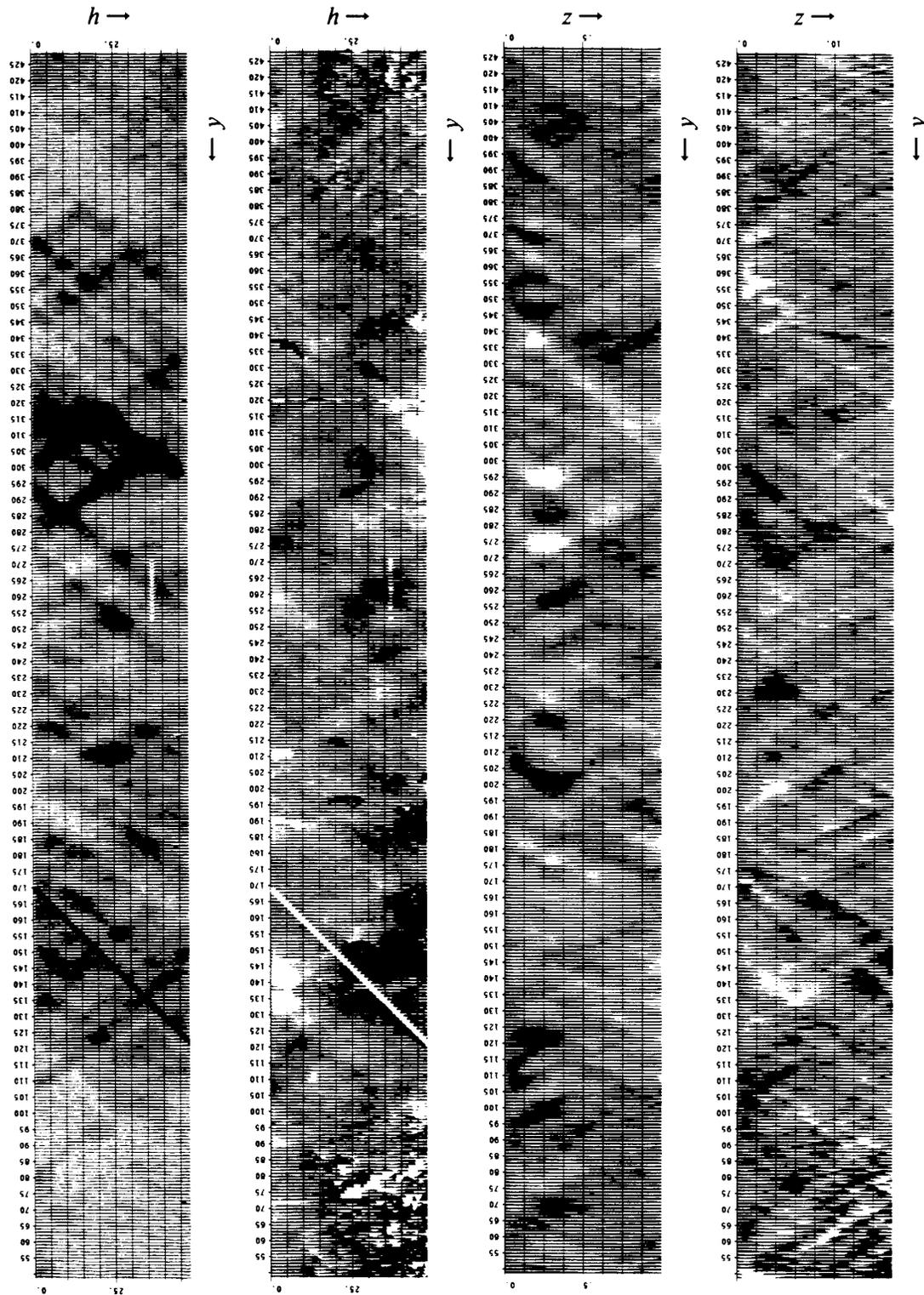


Figure 4: (a) amplitude (h,y), (b) timing (h,y) (c) amplitude (z,y), (d) timing (d,y) The missing trace anomaly shows up exactly at a 45 degree angle. Other anomalies are at lesser angles indicating their nonzero depth. jon1-kja [NR]

Wave-equation migration velocity analysis

*Biondo Biondi and Paul Sava*¹

ABSTRACT

In this report, we introduce a new wave-equation method of migration velocity analysis (MVA). The method is based on the linear relation that can be established between a perturbation in the migrated image and the generating perturbation in the slowness function. Our method consists of two steps: we first improve the focusing of the migrated image and then iteratively update the velocity model to explain the improvement in the focusing of the image. As a wave-equation method, our version of MVA is robust and generates smooth slowness functions without model regularization. We also show that our method has the potential to exploit the power of residual prestack migration to MVA.

INTRODUCTION

Seismic imaging is a two-step process: velocity estimation and migration. As the velocity function becomes more complex, the two steps become more and more dependent on each other. In complex depth-imaging problems, velocity estimation and migration are applied iteratively in a loop. To assure that this iterative imaging process converges to a satisfactory model, it is crucial that the migration and the velocity estimation are consistent with each other.

Kirchhoff migration often fails in complex areas, such as sub-salt, because the wavefield is severely distorted by lateral velocity variations, and thus complex multipathing occurs. As the shortcomings of Kirchhoff migration have become apparent (O'Brien and Etgen, 1998), there has been a renewal of interest in wave-equation migration and the development of computationally efficient 3-D prestack depth-migration methods based on the wave equation (Biondi and Palacharla, 1996; Biondi, 1997; Mosher et al., 1997). However, there has been no corresponding progress in the development of migration velocity analysis (MVA) methods that can be used in conjunction with wave-equation migration.

In this paper, we propose a method that aims to fill this gap and that, at least in principle, can be used in conjunction with any downward-continuation migration method. In particular, we have been applying our new methodology to downward continuation based on the Double Square Root equation in two dimensions (Yilmaz, 1979; Claerbout, 1985; Popovici, 1996) and on common-azimuth continuation in three dimensions (Biondi and Palacharla, 1996).

¹email: biondo@sep.stanford.edu, paul@sep.stanford.edu

As for migration, wave-equation MVA is intrinsically more robust than ray-based MVA, because it avoids the well-known problems that rays encounter when the velocity model is complex and has sharp boundaries. The transmission kinematic component of the finite-frequency wave propagation is mostly sensitive to smooth variations in the velocity model. Consequently, wave-equation MVA produces smooth velocity updates and is therefore stable. In most cases, no smoothing constraints are needed to assure stability in the inversion. In contrast, ray-based methods require strong smoothing constraints to avoid quick divergence.

Our method is closer to conventional MVA than other wave-equation methods that have been proposed to estimate the background velocity model (Noble et al., 1991; Bunks et al., 1995; Fogues et al., 1998), because it tries to maximize the quality of the migrated image rather than to match the recorded data. In this respect, our method is related to differential semblance optimization (DSO) (Symes and Carazzone, 1991) and multiple migration fitting (Chavent and Jacewitz, 1995). However, in contrast to these two methods, our method has the advantage of exploiting the power of residual prestack migration to speed up the convergence.

AN ALGORITHM FOR ESTIMATING VELOCITY

We estimate velocity by iteratively migrating the prestack data and looping through the following steps:

1. Downward continuation with current velocity,
2. Extraction of common-image gathers from prestack wavefields (Prucha et al., 1999),
3. Residual prestack migration of common-image gathers,
4. Estimation of image perturbation from the results of residual prestack migration, and
5. Estimation of velocity perturbation from image perturbations.

The core technical element of the method is the estimation of velocity perturbations from image perturbations. The next section presents the linear theory that enables us to achieve this goal.

LINEAR THEORY

In migration by downward continuation, illustrated in Figure 1, data measured at the surface (D) are recursively propagated down in depth to generate the complete wavefield (U). Downward continuation requires us to make an assumption about the magnitude of the slowness field (S). Once the wavefield is known, we can apply the imaging condition, which gives us the wavefield at time zero, or in the other words, the image or reflectivity map at the moment the reflectors explode (R).

In the presence of the background wavefield (U), a perturbation in slowness (ΔS) will generate a scattered wavefield (ΔW), which can, by the same method as the background field, be downward continued (ΔU) and imaged (ΔR), as shown in Figure 1.

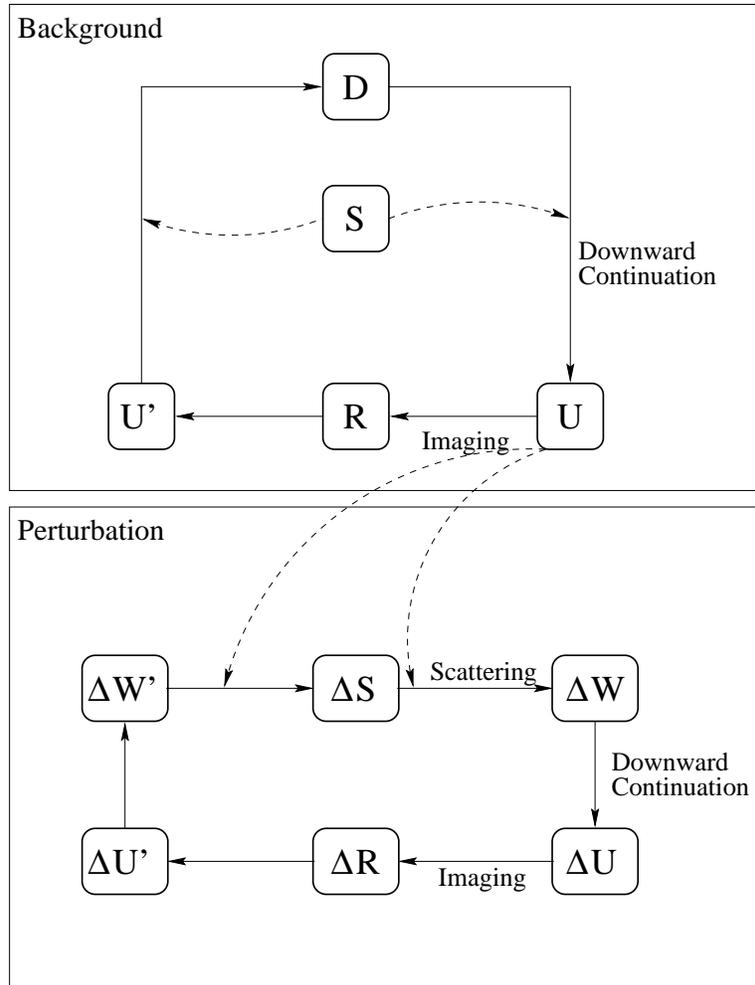


Figure 1: A summary-chart of our MVA method. The upper panel describes the computations done with respect to the background field, while the bottom panel refers to the computations done with respect to the perturbation field. `biondo2-chart` [NR]

We can take the perturbation in image (ΔR) and apply to it the adjoint operation. Doing so creates an adjoint perturbation in wavefield ($\Delta U'$), an adjoint scattered field ($\Delta W'$), and eventually an adjoint perturbation in slowness ($\Delta S'$), as the bottom panel of Figure 1 shows. Considering a first-order Born relation between the perturbation in slowness and the scattered wavefield, we can establish a direct linear relation between the perturbation in image (ΔR) and the perturbation in slowness (ΔS). It follows that if we can obtain a better focused image, we can iteratively invert for the perturbation in slowness that generated the improvement in focusing. This is the foundation of our wave-equation MVA method.

In the next two sections, we briefly present the mathematical relations that form the basis

of our method. A more detailed mathematical description appears in Appendices A and B.

Background field: Forward operator

Migration by downward continuation, in post-stack or prestack, is done in two steps: the first step is to downward continue the data (D) measured at the surface, and the second is to apply the imaging condition, that is, to extract the wavefield at time $t = 0$, or the image (R) at the moment the reflectors explode (Claerbout, 1985).

1. Downward continuation

The first step of migration consists of downward continuation of the wavefield measured at the surface (a.k.a. the data), which is done by the recursive application of the equation

$$u_0^{z+1} = T_0^z u_0^z \quad (1)$$

initialized by the wavefield at the surface

$$u_0^1 = f \cdot d \quad (2)$$

where

- $u_0^z(\omega)$ is the wavefield $u_0(\omega)$ at depth z ,
- $u_0^1(\omega)$ is the wavefield $u_0(\omega)$ at the surface ($z = 0$),
- $T_0^z(\omega, s_0)$ is the downward continuation operator at depth z ,
- $d(\omega)$ is the data, i.e., the wavefield at the surface, and
- $f(\omega)$ is a frequency-dependent scale factor for the data.

2. Imaging

The second step of the migration by downward continuation is imaging. According to the exploding reflector concept, the image is found by selecting the wavefield at time $t = 0$, or equivalently, by summing over the frequencies ω :

$$r_0^z = \sum_1^{N_\omega} u_0^z(\omega) \quad (3)$$

where

- r_0^z is the image (reflectivity) corresponding to a given depth level z .

Perturbation field: Forward operator

If we perturb the velocity model, we also introduce a perturbation in the wavefield. In other words, the perturbation in slowness generates a secondary scattered wavefield.

1. Scattering and downward continuation

If we consider the perturbation in the wavefield at the surface, we can recursively downward continue it, adding at every depth step the scattered wavefield:

$$\Delta u^{z+1} = T_0^z \Delta u^z + \Delta v^{z+1} \quad (4)$$

where

- $\Delta u^z(\omega)$ is the perturbation in the wavefield generated by the perturbation in velocity and downward continued from the surface, and
- $\Delta v^{z+1}(\omega)$ represents the scattered wavefield caused at depth level $z + 1$ by the perturbation in velocity from depth level z .

In the first-order Born approximation, the scattered wavefield can be written as

$$\Delta v^{z+1} = T_0^z G_0^z u_0^z \Delta s^z \quad (5)$$

where

- $G_0^z(\omega, s_0)$ is the scattering operator at depth z ,
- $\Delta s^z(\omega)$ is the perturbation in slowness at depth z , and
- $u_0^z(\omega)$ is the background wavefield at depth z .

If we introduce equation (5) into (4) we find that

$$\Delta u^{z+1} = T_0^z [\Delta u^z + G_0^z u_0^z \Delta s^z] \quad (6)$$

2. Imaging

As for the background image, the perturbation in image (Δr^z), caused by the perturbation in slowness, is obtained by a summation over all the frequencies ω :

$$\Delta r^z = \sum_1^{N_\omega} \Delta u^z(\omega) \quad (7)$$

Equations (6) and (7) establish a linear relation between the perturbation in slowness (Δs^z) and the perturbation in image (Δr^z). We can use this linear relation in an iterative algorithm to invert for the perturbation in slowness based on the perturbation in the image.

Perturbation field: Adjoint operator

In the adjoint operation, we begin by upward propagating the perturbation in wavefield at depth z :

$$\Delta u^{z-1} = T_0^{z'} \Delta u^z + \Delta r^{z-1} \quad (8)$$

where

- $T_0^{z'}$ is the upward continuation operator at depth z .

We can then obtain the perturbation in slowness from the perturbation in wavefield by applying the adjoint of the scattering operator:

$$\Delta s^z = u_0^{z'} G_0^{z'} [T_0^{z'} \Delta u^{z+1} - \Delta u^z] \quad (9)$$

Equations (6) and (7) for the forward operator and equations (8) and (9) for the adjoint operator express the linear relation established between the perturbation in slowness (ΔS) and the perturbation in image (ΔR).

AN EXAMPLE WITH SIMPLE REFLECTIVITY MODELS

This section offers a pictorial description of the theory presented in the preceding section. We use two simple examples to highlight the main features of the method. In both cases, the reflectivity model consists of two flat interfaces, one shallow and one deep. The velocity models are different as follows:

- In the first model, we started with a constant velocity background of 2 km/s, on which we superimposed a positive Gaussian perturbation with a magnitude of 0.25 km/s, as shown in the right panel of Figure 2.
- In the second model, we superimposed a negative Gaussian anomaly with the magnitude of 0.5 km/s on the background velocity (Figure 2), while the perturbation remained the same as in the first model (the right panel of Figure 2). The main purpose of selecting this second velocity model was to demonstrate the robustness of the forward and adjoint operators to triplications in the wavefield.

We started by creating the synthetic data (D) that correspond to each of the individual models (Figure 3). In the second case, the reflection from the deeper interface creates a triplication caused by the Gaussian anomaly in the background velocity. Then, we migrated the synthetic data using the correct velocity models in each case, and obtained the background images (R) shown in Figure 4.

We then repeated the same succession of operations, considering the background velocity models on which we superimposed the perturbation anomaly. We then created the data and

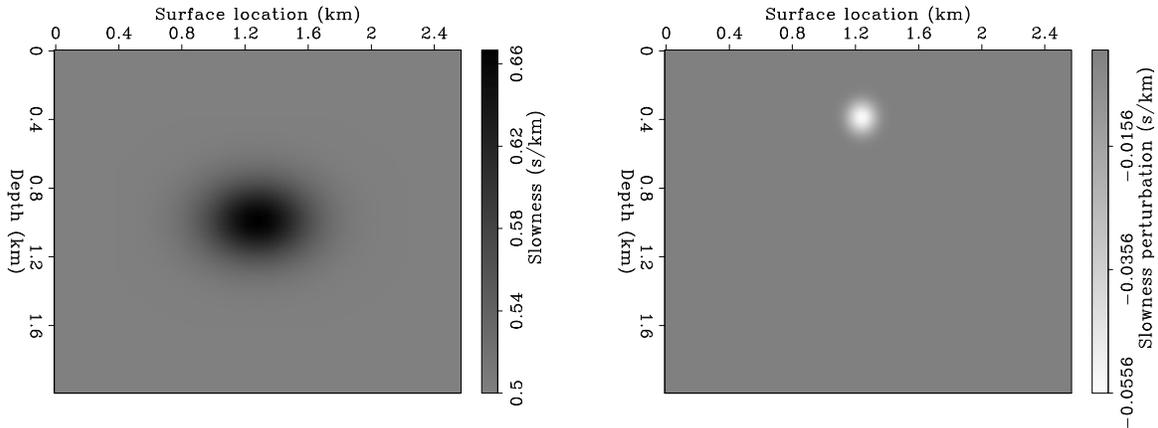


Figure 2: Background slowness (labeled S in Figure 1) – left; Perturbation in slowness (labeled ΔS in Figure 1) – right. `biondo2-slow` [NR]

migrated it with the perturbation in slowness. What we obtained is the perturbation in image depicted in Figure 5. The shape of the image at a given depth is known in the literature as *Kjartansson's V* (Kjartansson, 1979). In the case of the nonconstant background, the triplications of the wavefield created a more complex perturbation in the image, which is especially visible at the level of the deeper reflector. For this case, Kjartansson's V becomes a W (Figure 5).

Finally, we back-projected into the velocity model the perturbations we obtained in the images (Figure 6). To clarify how the back-projection operator works, we have isolated in each panel a single event of the perturbation in image, for a fixed reflection ray parameter. As expected, we have obtained “fat rays” showing which regions of the velocity model are influenced by the perturbation in image. The top panel of Figure 6 displays the straight fat rays corresponding to the constant velocity background. The bottom left panel, shows the rays for a similar perturbation in the image as in the first case, while the bottom right panel displays the rays for the perturbation in image in a region where the wavefield has triplicated when propagating through the anomaly in the background.

AN EXAMPLE OF INVERSION

This section presents an example inversion for the perturbation in slowness using the linear operators derived in the section on linear theory. For the inversion, we have created a set of synthetic data that was inspired by a real dataset, part of a gas-hydrate study, which was recorded at the Blake Outer Ridge, offshore from Florida and Georgia (Ecker, 1998). We have divided this example in two parts: in the first, we show how the focusing of the image can be improved, with application to the real data, and, in the second part, how the inversion works, with application to the smaller synthetic data set.

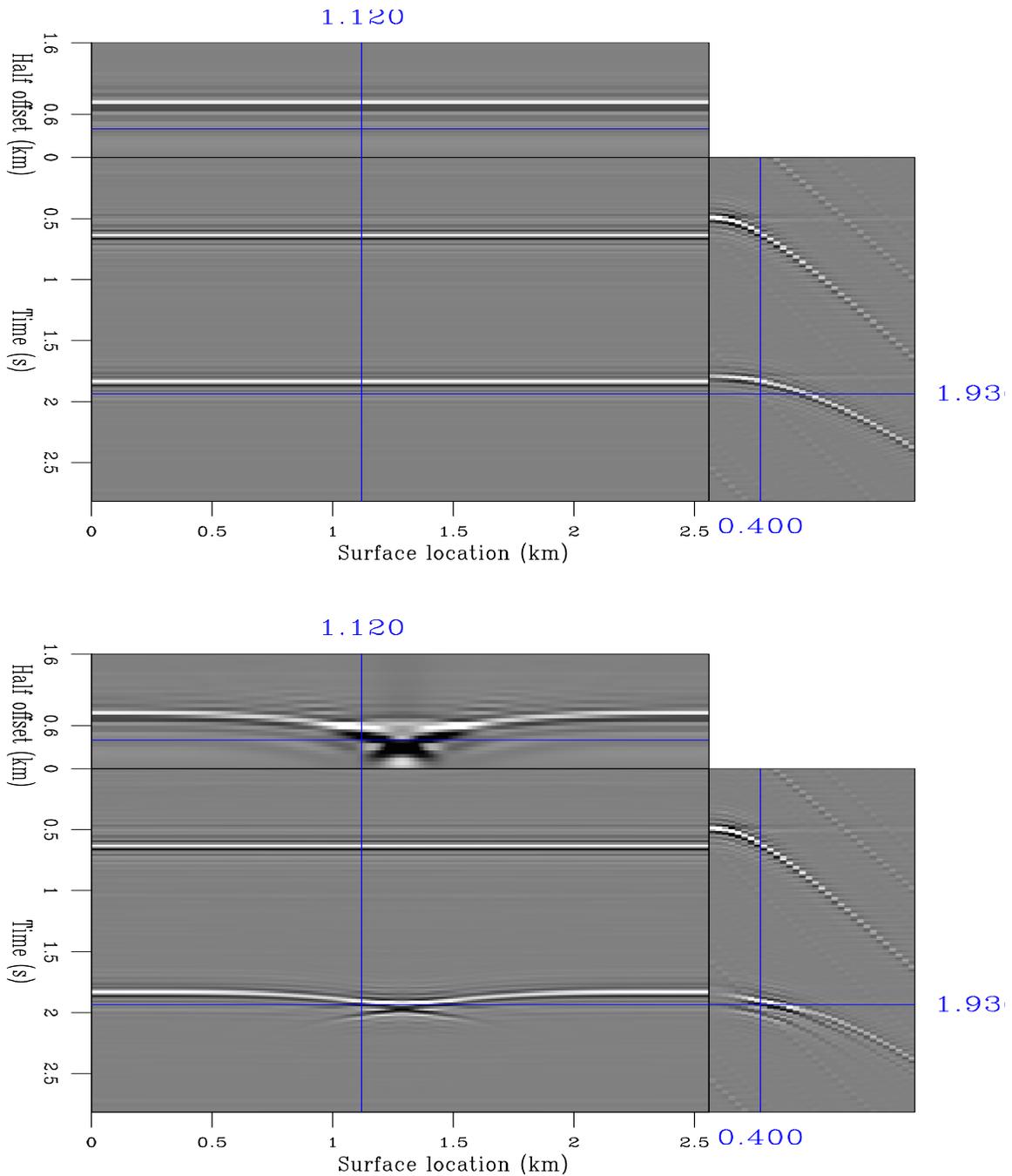


Figure 3: The data (labeled D in Figure 1) measured at the surface for each of the background velocity models. The top panel corresponds to the case of the constant background velocity, while the bottom panel corresponds to the case of the Gaussian anomaly in the background velocity. `biondo2-data` [CR]

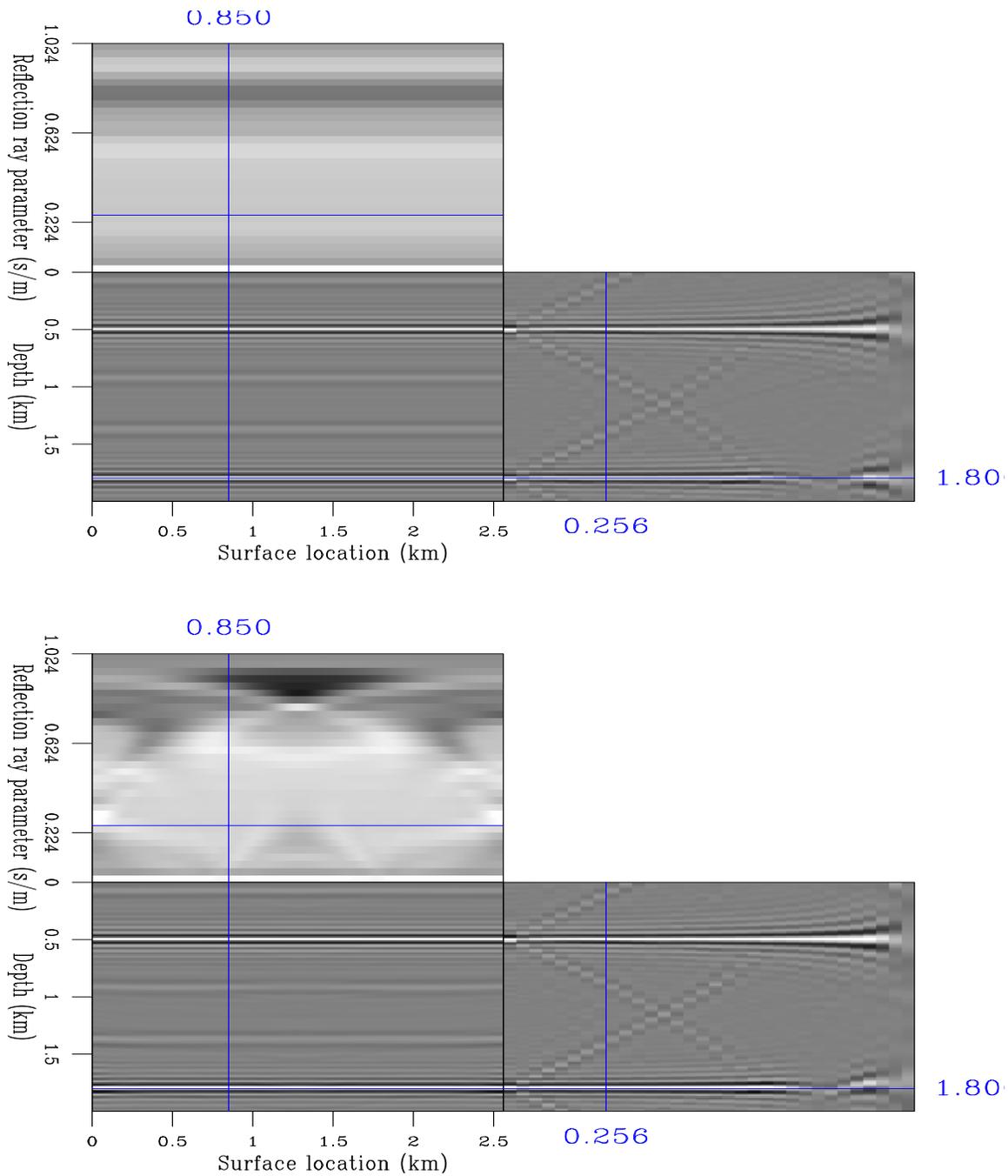


Figure 4: The image (labeled R in Figure 1) obtained after migrating by downward continuation the data in Figure 3. The top panel corresponds to the case of the constant background velocity, while the bottom panel corresponds to the case of the Gaussian anomaly in the background velocity. `biondo2-image` [CR]

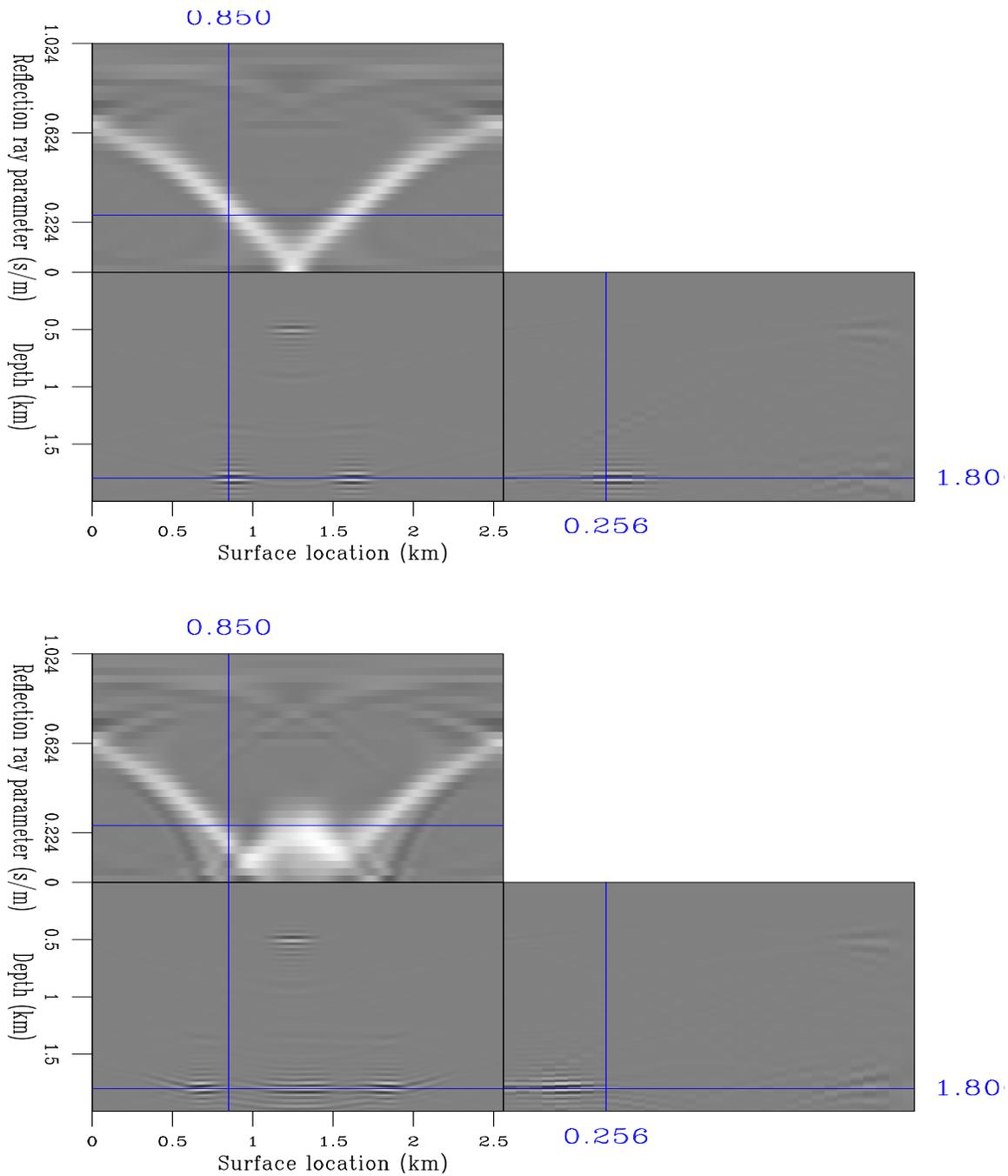


Figure 5: The perturbation in image (labeled ΔR in Figure 1), that is, the data migrated with the perturbation in slowness. The top panel corresponds to the case of the constant background velocity (Kjartansson's V), while the bottom panel corresponds to the case of the Gaussian anomaly in the background velocity. `biondo2-dimage` [CR]

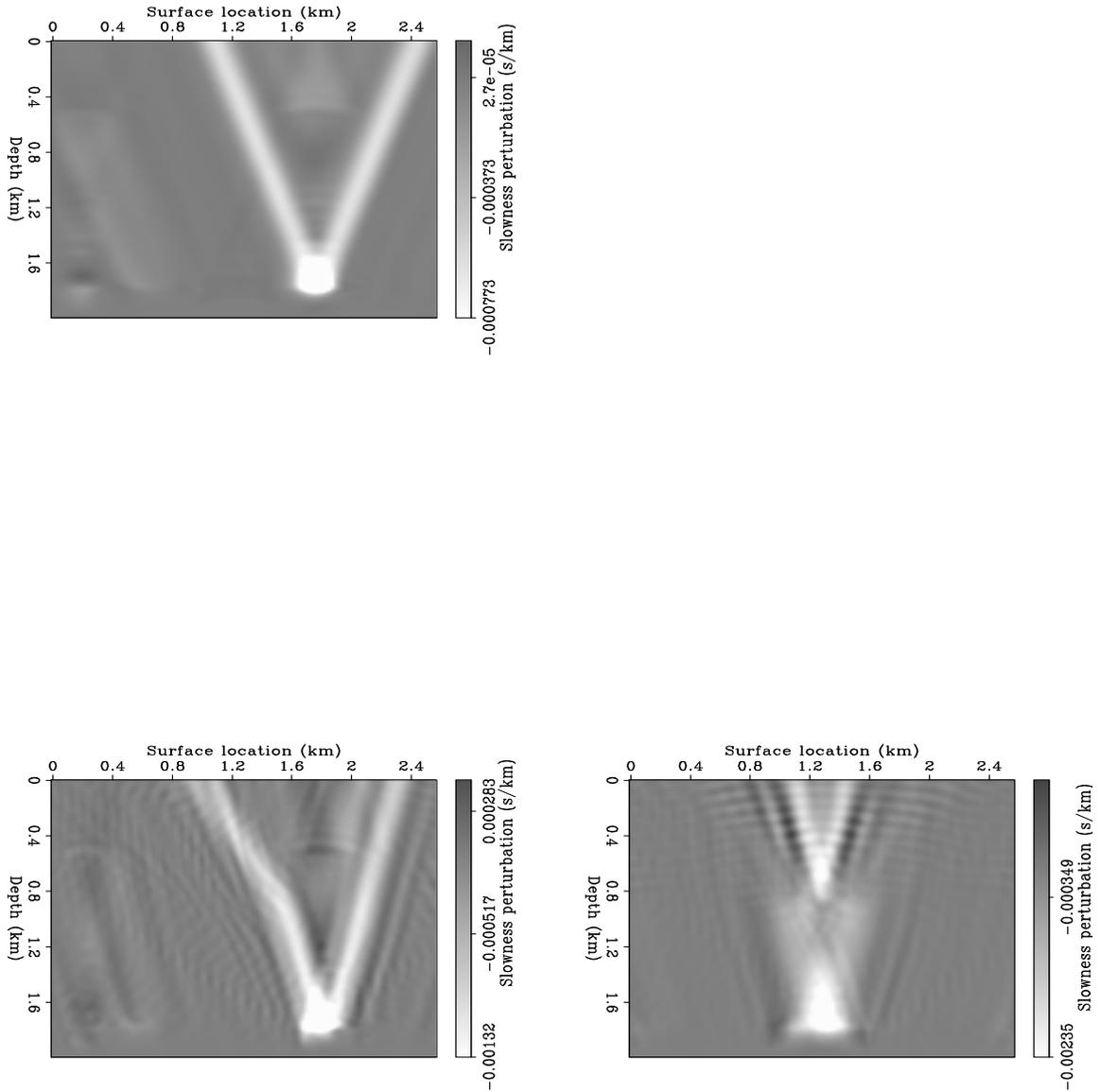


Figure 6: The perturbation in slowness (labeled ΔS in Figure 1) corresponding to part of the perturbation in image (Figure 5). The “fat rays” are the result of the back-projection of the perturbation in slowness. The top panel corresponds to the case of the constant background velocity, while the bottom panel corresponds to the case of the Gaussian anomaly in the background velocity. biondo2-rays [CR]

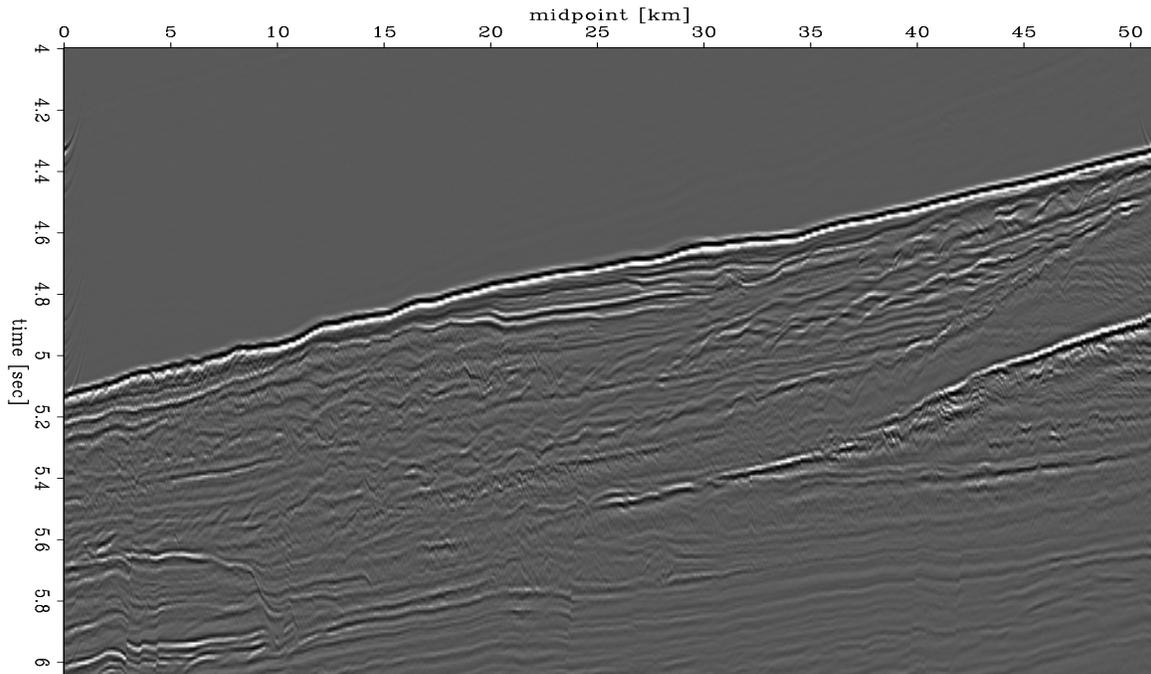


Figure 7: The original image. `biondo2-hydrates1.00` [NR]

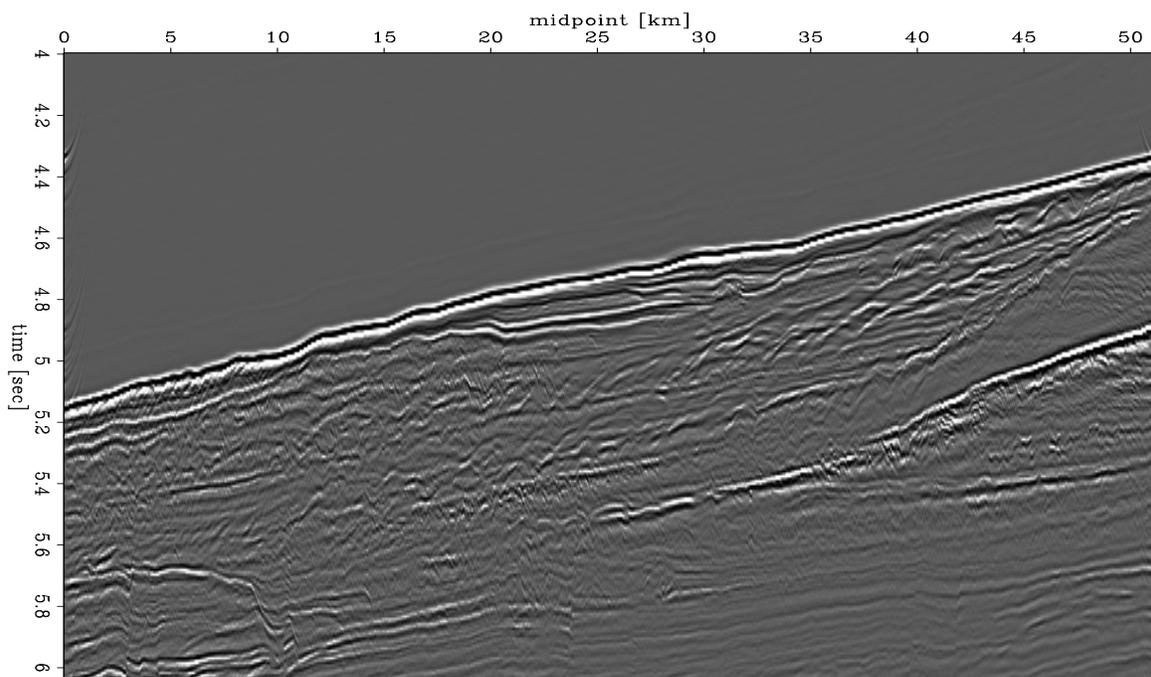


Figure 8: A better focused image after residual migration. `biondo2-hydrates0.98` [NR]

Image enhancement

In the first part of our example, we have concentrated on improving the focusing of the image. We achieved this goal by using Stolt residual migration in the prestack domain (Stolt, 1996; Sava, 1999). Of course, Stolt residual migration is not the only possibility, another alternative being velocity continuation (Fomel, 1997).

Figure 7 shows the image we obtained with a starting velocity model, while Figure 8 displays the image we obtained by applying residual migration to the original. Both images have been created with the same level of clipping. The second image is clearly better focused. We can take the difference between the images in Figures 7 and 8 to be the perturbation in image (ΔR), and use it to invert for the slowness model that generated the improvement in focusing.

Inversion

As a first experiment, we have constructed a synthetic model similar to the sections in Figures 7 and 8. As we said, our goal is to convert the differences in focusing between the two images, or the perturbation in the image, into a better slowness model, that is, to find the perturbation in the slowness.

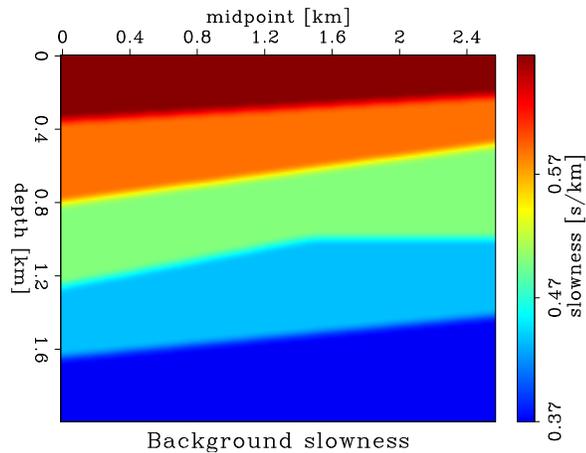
Figure 9 represents the background slowness model (S). We used this model to generate the synthetic data at the surface (S), and then to compute the background wavefield (U) and the background image (R).

The top panel of Figure 10 shows the perturbation in slowness (ΔS). We used this model to generate the scattered wavefield (ΔW), the perturbation wavefield (ΔU), and the perturbation in image (ΔR).

We start the inversion by assuming zero perturbation in slowness. The middle panel of Figure 10 represents the perturbation in slowness obtained at the first iteration. At this stage, we have obtained only a small perturbation in slowness, which is not totally concentrated at the right location. An important part of the energy of the section is spread, for example in the region around the midpoint 2.2 – 2.4km and around the depth 1.3 – 1.4km. This artifact is the result of the still imperfect definition of the slowness anomaly, possibly caused by the proximity of the edge.

By the 20th iteration, shown in the bottom panel of Figure 10, the perturbation in slowness is much better shaped, and the artifact at depth is much weaker. Also, the absolute magnitude of the anomaly is getting very close to the correct value: $s_{max} = 0.088$ s/km for the original, and $s_{max} = 0.084$ s/km for the inversion at the 20th iteration.

Figure 9: The background slowness (S). `biondo2-backslo` [NR]



CONCLUSION

We have presented a recursive wave-equation method of migration velocity analysis operating in the image domain. Our method is based on linearization of the downward continuation operator that relates perturbations in slowness to perturbations in the image. The fundamental idea is to improve the quality of the slowness function by optimizing the focusing of the migrated image.

This iterative method is stable and accurate when applied to a synthetic dataset. It also converges to the solution without the need for any regularization of the slowness model. We are currently in the process of applying the method to the real seismic dataset used as an example in this paper.

REFERENCES

- Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: *Geophysics*, **61**, 1822–1832.
- Biondi, B., 1997, Azimuth moveout + common-azimuth migration: Cost-effective prestack depth imaging of marine data: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1375–1378.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G., 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, no. 5, 1457–1473.
- Chavent, G., and Jacewitz, C. A., 1995, Determination of background velocities by multiple migration fitting: *Geophysics*, **60**, no. 2, 476–490.
- Claerbout, J. F., 1985, *Imaging the Earth's Interior*: Blackwell Scientific Publications.
- Ecker, C., 1998, *Seismic characterization of gas hydrates structures*: Ph.D. thesis, Stanford University.

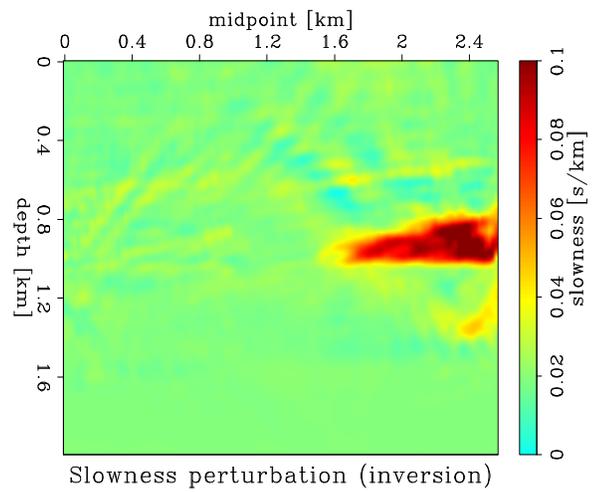
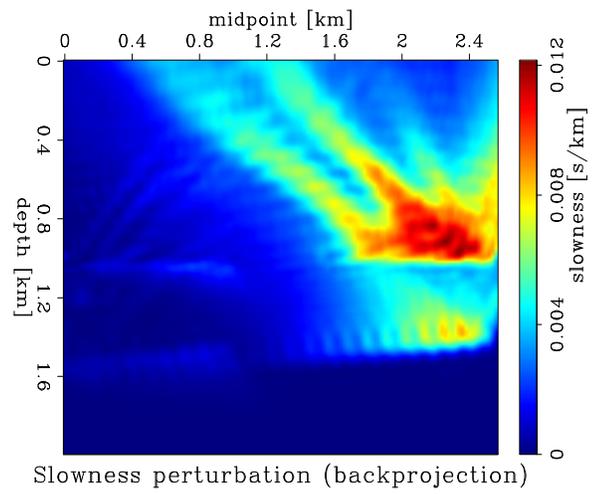
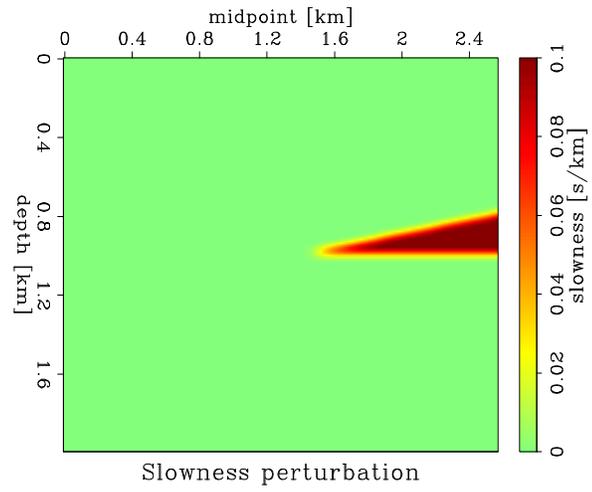


Figure 10: Top: the perturbation in slowness (ΔS). Middle: the perturbation recovered at the first iteration. Bottom: the perturbation recovered at the 20th iteration. `biondo2-inversion` [CR]

- Fogues, E., Scala, E., and Pratt, R. G., 1998, High-resolution velocity estimation from re-fraction and reflection data: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1211–1214.
- Fomel, S., 1997, Velocity continuation and the anatomy of residual prestack migration: 67th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1762–1765.
- Huang, L.-J., Fehler, M. C., and Wu, R.-S., 1999, Extended local Born Fourier migration method: submitted for publication to Geophysics.
- Kjartansson, E., 1979, Attenuation of seismic waves in rocks and applications in energy exploration: Ph.D. thesis, Stanford University.
- Mosher, C. C., Foster, D. J., and Hassanzadeh, S., 1997, Common angle imaging with offset plane waves: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1379–1382.
- Noble, M., Lindgren, J., and Tarantola, A., 1991, Large-sized, nonlinear inversion of a marine data set: Retrieving the source, the background velocity and the impedance contrasts: 61th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 893–896.
- O'Brien, M. J., and Etgen, J. T., 1998, Wavefield imaging of complex structures with sparse point-receiver data: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1365–1368.
- Popovici, A. M., 1996, Short note—prestack migration by split-step DSR: *Geophysics*, **61**, no. 05, 1412–1416.
- Prucha, M., Biondi, B., and Symes, W., 1999, Angle-domain common image gathers by wave-equation migration: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Sava, P., 1999, Short note—on Stolt prestack residual migration: *SEP-100*, 151–158.
- Stolt, R. H., 1996, Short note—a prestack residual time migration operator: *Geophysics*, **61**, no. 02, 605–607.
- Symes, W. W., and Carazzone, J. J., 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, no. 5, 654–663.
- Yilmaz, O., 1979, Prestack partial migration: Ph.D. thesis, Stanford University.

APPENDIX A: DOWNWARD CONTINUATION–MIGRATION AND MODELING

Forward operator: Migration

Migration by downward continuation, either post-stack or prestack, is done in two steps: the first step is to downward continue the data measured at the surface, and the second is to apply the imaging condition, that is, to extract the wavefield at time $t = 0$, the moment the reflectors explode.

1. Downward continuation

The first step of migration consists of downward continuation of the wavefield measured at the surface (a.k.a. the data), which is done by the recursive application of the equation:

$$u_0^{z+1}(\omega) = T_0^z(\omega, s_0)u_0^z(\omega) \quad (\text{A-1})$$

initialized by the wavefield at the surface, as follows:

$$u_0^1(\omega) = f(\omega)d(\omega) \quad (\text{A-2})$$

where

- $u_0^z(\omega)$ is the wavefield $u_0(\omega)$ at depth z ,
- $u_0^1(\omega)$ is the wavefield $u_0(\omega)$ at the surface $z = 0$,
- $T_0^z(\omega, s_0)$ is the downward continuation operator at depth z ,
- $d(\omega)$ is the data, i.e.. the wavefield at the surface, and
- $f(\omega)$ is a frequency-dependent scale factor for the data.

The recursion in Equations (A-1) and (A-2) can be also rewritten in matrix form as

$$[\mathbf{I} - \mathbf{T}_0]\mathbf{U}_0 = \mathbf{D} \quad (\text{A-3})$$

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -T_0^1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -T_0^2 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -T_0^{\ddot{N}_z-1} & 1 \end{bmatrix} \begin{bmatrix} u_0^1 \\ u_0^2 \\ u_0^3 \\ \dots \\ u_0^{\ddot{N}_z} \end{bmatrix} = \begin{bmatrix} f d \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

where

- \mathbf{T}_0 is a square matrix containing the downward continuation operator for all depth levels,
- \mathbf{U}_0 is a column vector containing the wavefield at all depth levels, and

- \mathbf{D} is a column vector containing the scaled data.²

Equation (A-3) represents the downward continuation recursion written for a given frequency. We can write a similar relationship for each of the frequencies in the analyzed data, and group them all in the matrix relationship

$$\boxed{(\mathcal{I} - \mathcal{T}_0)\mathcal{U}_0 = \mathcal{D}} \quad (\text{A-4})$$

$$\begin{pmatrix} \mathbf{I} - \mathbf{T}_0(\omega_1, \mathbf{s}_0) & 0 & \dots & 0 \\ 0 & \mathbf{I} - \mathbf{T}_0(\omega_2, \mathbf{s}_0) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{I} - \mathbf{T}_0(\omega_{N_\omega}, \mathbf{s}_0) \end{pmatrix} \begin{pmatrix} \mathbf{U}_0(\omega_1) \\ \mathbf{U}_0(\omega_2) \\ \dots \\ \mathbf{U}_0(\omega_{N_\omega}) \end{pmatrix} = \begin{pmatrix} \mathbf{D}(\omega_1) \\ \mathbf{D}(\omega_2) \\ \dots \\ \mathbf{D}(\omega_{N_\omega}) \end{pmatrix}$$

where

- \mathcal{T}_0 is a diagonal matrix containing the downward continuation operators for all the frequencies in the data,
- \mathcal{U}_0 is a column vector containing the wavefield data for all the frequencies, and
- \mathcal{D} is a column vector containing the scaled data at all frequencies.³

It follows from Equation (A-4) that the background wavefield (\mathcal{U}_0) can be computed as a function of the measured data (\mathcal{D}), as follows:

$$\mathcal{U}_0 = (\mathcal{I} - \mathcal{T}_0)^{-1} \mathcal{D} \quad (\text{A-5})$$

2. Imaging

The second step of the migration by downward continuation is imaging. In the exploding reflector concept, the image is found by selecting the wavefield at time $t = 0$ or, equivalently, by summing over the frequencies ω :

$$\boxed{r_0^z = \sum_1^{N_\omega} u_0^z(\omega)} \quad (\text{A-6})$$

where

- r_0^z is the image (reflectivity) corresponding to a given depth level z .

2

$$\mathbf{T}_0(\omega, \mathbf{s}_0) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ T_0^1 & 0 & 0 & \dots & 0 & 0 \\ 0 & T_0^2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T_0^{N_z-1} & 0 \end{bmatrix}; \mathbf{U}_0(\omega) = \begin{bmatrix} u_0^1 \\ u_0^2 \\ u_0^3 \\ \dots \\ u_0^{N_z} \end{bmatrix}; \mathbf{D}(\omega) = \begin{bmatrix} f & d \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

3

$$\tilde{\mathcal{T}}_0 = \begin{pmatrix} \mathbf{T}_0(\omega_1, \mathbf{s}_0) & 0 & \dots & 0 \\ 0 & \mathbf{T}_0(\omega_2, \mathbf{s}_0) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{T}_0(\omega_{N_\omega}, \mathbf{s}_0) \end{pmatrix}; \mathcal{U}_0 = \begin{pmatrix} \mathbf{U}_0(\omega_1) \\ \mathbf{U}_0(\omega_2) \\ \dots \\ \mathbf{U}_0(\omega_{N_\omega}) \end{pmatrix}; \mathcal{D} = \begin{pmatrix} \mathbf{D}(\omega_1) \\ \mathbf{D}(\omega_2) \\ \dots \\ \mathbf{D}(\omega_{N_\omega}) \end{pmatrix}$$

We can write the Equation (A-6) in matrix form as

$$\mathcal{R}_0 = \mathcal{H} \mathcal{U}_0 \quad (\text{A-7})$$

$$\begin{pmatrix} r_0^1 \\ r_0^2 \\ \vdots \\ r_0^{N_z} \end{pmatrix} = \left(\begin{array}{cccc|cccc|cccc} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 & 1 & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 0 & 0 & \dots & 1 \end{array} \right) \begin{pmatrix} \mathbf{U}_0(\omega_1) \\ \mathbf{U}_0(\omega_2) \\ \dots \\ \mathbf{U}_0(\omega_{N_\omega}) \end{pmatrix}$$

where

- \mathcal{H} is an operator performing the summation over frequency for every depth level z , and
- \mathcal{R}_0 is a column vector containing the image at every depth level.⁴

Therefore, the image (\mathcal{R}_0), corresponding to the background velocity field, can be computed from the measured data (\mathcal{D}) using the summation (\mathcal{H}) and the downward continuation operators (\mathcal{T}_0) as

$$\mathcal{R}_0 = \mathcal{H}(\mathcal{I} - \mathcal{T}_0)^{-1} \mathcal{D} \quad (\text{A-8})$$

Adjoint operator: Modeling

Equation (A-8) enables us to compute the image corresponding to a given velocity field from data measured at the surface of the earth, that is, to migrate the data. The operation adjoint to migration, modeling, can be derived from the same equation using the adjoint state system

$$\overline{\mathcal{D}} = [\mathcal{H}(\mathcal{I} - \mathcal{T}_0)^{-1}]' \mathcal{R}_0 \quad (\text{A-9})$$

where

- $\overline{\mathcal{D}}$ is the modeled data, computed for a given velocity field.

Therefore, we can obtain the modeled data ($\overline{\mathcal{D}}$) from the reflectivity image (\mathcal{R}_0) by writing

$$\overline{\mathcal{D}} = [(\mathcal{I} - \mathcal{T}_0)^{-1}]' \mathcal{H}' \mathcal{R}_0 \quad (\text{A-10})$$

4

$$\mathcal{H} = \left(\begin{array}{cccc|cccc|cccc} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 & 1 & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 0 & 0 & \dots & 1 \end{array} \right); \mathcal{R}_0 = \begin{pmatrix} r_0^1 \\ r_0^2 \\ \vdots \\ r_0^{N_z} \end{pmatrix}$$

APPENDIX B: FIRST-ORDER BORN LINEARIZATION OF MIGRATION

Forward operator: Perturbation migration

1. Scattering and downward continuation

If we perturb the velocity model we introduce a perturbation in the wavefield. In other words, the perturbation in slowness generates a secondary wavefield, the scattered wavefield. We can downward continue the scattered field as we did with the background wavefield by writing

$$\Delta u^{z+1}(\omega) = T_0^z(\omega, s_0) \Delta u^z(\omega) + \Delta v^{z+1}(\omega) \quad (\text{B-1})$$

where

- $\Delta u^z(\omega)$ is the perturbation in the wavefield generated by the perturbation in velocity, and
- $\Delta v^{z+1}(\omega)$ represents the scattered wavefield caused at depth level $z + 1$ by the perturbation in velocity from the depth level z .

The scattered wavefield can be written as

$$\Delta v^{z+1}(\omega) = T_0^z(\omega, s_0) G_0^z(\omega, s_0) u_0^z(\omega) \Delta s^z(\omega) \quad (\text{B-2})$$

where

- $G_0^z(\omega, s_0)$ is the scattering operator at depth z , and
- $\Delta s^z(\omega)$ is the perturbation in slowness at depth z .

Huang et al. (1999) show that the scattering operator is

$$G_0^z(\omega, s_0) = i\omega \frac{1}{\sqrt{1 - \frac{|\vec{k}|^2}{\omega^2 s_0^2}}} \quad (\text{B-3})$$

and that it can be approximated by

$$G_0^z(\omega, s_0) \approx i\omega \left(1 + \frac{1}{2} \frac{|\vec{k}|^2}{\omega^2 s_0^2} \right) \quad (\text{B-4})$$

which represents the first-order Born approximation. In this equation, \vec{k} represents the horizontal component of the wavenumber.

If we introduce Equation (B-2) into (B-1) we obtain

$$\Delta u^{z+1} = T_0^z \left[\Delta u^z + G_0^z u_0^z \Delta s^z \right] \quad (\text{B-5})$$

which, after rearrangements, becomes the recursion

$$\Delta u^{z+1} - T_0^z \Delta u^z = T_0^z G_0^z u_0^z \Delta s^z \quad (\text{B-6})$$

We can express the recursive relationship between the perturbation in velocity and the perturbation in the wavefield (B-6) as

$$\boxed{[\mathbf{I} - \mathbf{T}_0] \Delta \mathbf{U} = \mathbf{T}_0 \mathbf{G}_0 \widehat{\mathbf{U}}_0 \Delta \mathbf{S}} \quad (\text{B-7})$$

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -T_0^1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -T_0^2 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -T_0^{\ddot{N}_z-1} & 1 \end{bmatrix} \begin{bmatrix} \Delta u^1 \\ \Delta u^2 \\ \Delta u^3 \\ \dots \\ \Delta u^{\ddot{N}_z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ T_0^1 & 0 & 0 & \dots & 0 & 0 \\ 0 & T_0^2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T_0^{\ddot{N}_z-1} & 0 \end{bmatrix} \begin{bmatrix} G_0^1 & 0 & 0 & \dots & 0 & 0 \\ 0 & G_0^2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & G_0^{\ddot{N}_z-1} & 0 \end{bmatrix} \begin{bmatrix} u_0^1 & 0 & 0 & \dots & 0 \\ 0 & u_0^2 & 0 & \dots & 0 \\ 0 & 0 & u_0^3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_0^{\ddot{N}_z} \end{bmatrix} \begin{bmatrix} \Delta s^1 \\ \Delta s^2 \\ \Delta s^3 \\ \dots \\ \Delta s^{\ddot{N}_z} \end{bmatrix}$$

where

- $\Delta \mathbf{U}$ is a column vector containing the perturbation in the wavefield at all depths,
- \mathbf{G}_0 is a diagonal matrix containing the scattering term for all the depth levels,
- $\widehat{\mathbf{U}}_0$ is a diagonal matrix containing the background wavefield data for all the depth levels, and
- $\Delta \mathbf{S}$ is a column vector containing the perturbation in the velocity for all the depth levels.⁵

Note the different arrangement of the background wavefield data at all depths (\mathbf{U}_0 and $\widehat{\mathbf{U}}_0$).⁶

Similarly to the case of the background wavefield, the relationship between the perturbation in the wavefield and the perturbation in slowness can be written for all the frequencies in the data as

$$\boxed{(\mathcal{I} - \mathcal{T}_0) \Delta \mathcal{U} = \mathcal{T}_0 \mathcal{G}_0 \widehat{\mathcal{U}}_0 \Delta \mathcal{S}} \quad (\text{B-8})$$

5

$$\Delta \mathbf{U} = \begin{bmatrix} \Delta u^1 \\ \Delta u^2 \\ \Delta u^3 \\ \dots \\ \Delta u^{\ddot{N}_z} \end{bmatrix}; \mathbf{G}_0 = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ G_0^1 & 0 & 0 & \dots & 0 & 0 \\ 0 & G_0^2 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & G_0^{\ddot{N}_z-1} & 0 \end{bmatrix}; \widehat{\mathbf{U}}_0 = \begin{bmatrix} u_0^1 & 0 & 0 & \dots & 0 \\ 0 & u_0^2 & 0 & \dots & 0 \\ 0 & 0 & u_0^3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_0^{\ddot{N}_z} \end{bmatrix}; \Delta \mathbf{S} = \begin{bmatrix} \Delta s^1 \\ \Delta s^2 \\ \Delta s^3 \\ \dots \\ \Delta s^{\ddot{N}_z} \end{bmatrix}$$

6

$$\mathbf{U}_0 = \begin{bmatrix} u_0^1 \\ u_0^2 \\ u_0^3 \\ \dots \\ u_0^{\ddot{N}_z} \end{bmatrix}; \widehat{\mathbf{U}}_0 = \begin{bmatrix} u_0^1 & 0 & 0 & \dots & 0 \\ 0 & u_0^2 & 0 & \dots & 0 \\ 0 & 0 & u_0^3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_0^{\ddot{N}_z} \end{bmatrix}$$

$$\begin{pmatrix} \mathbf{I} - \mathbf{T}_0(\omega_1, \mathbf{s}_0) & 0 & \dots & 0 \\ 0 & \mathbf{I} - \mathbf{T}_0(\omega_2, \mathbf{s}_0) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{I} - \mathbf{T}_0(\omega_{N_\omega}, \mathbf{s}_0) \end{pmatrix} \begin{pmatrix} \Delta \mathbf{U}(\omega_1) \\ \Delta \mathbf{U}(\omega_2) \\ \dots \\ \Delta \mathbf{U}(\omega_{N_\omega}) \end{pmatrix} = \begin{pmatrix} \mathbf{T}_0 \mathbf{G}_0 \widehat{\mathbf{U}}_0 \Delta \mathbf{S} \\ \mathbf{T}_0 \mathbf{G}_0 \widehat{\mathbf{U}}_0 \Delta \mathbf{S} \\ \dots \\ \mathbf{T}_0 \mathbf{G}_0 \widehat{\mathbf{U}}_0 \Delta \mathbf{S} \end{pmatrix} \begin{matrix} \omega_1 \\ \omega_2 \\ \dots \\ \omega_{N_\omega} \end{matrix}$$

where

- $\Delta \mathcal{U}$ is a column vector containing the perturbation in the wavefield for all the frequencies,
- \mathcal{G}_0 is a diagonal matrix containing the scattering operator for all the frequencies,
- $\widehat{\mathcal{U}}_0$ is a diagonal matrix containing the background wavefield for all the frequencies, and
- $\Delta \mathcal{S}$ is a column vector containing the perturbation in slowness, same for all the frequencies if we disregard dispersion.⁷

Again, it is important to note the different arrangement of the background wavefield data at all frequencies (\mathcal{U}_0 and $\widehat{\mathcal{U}}_0$).⁸

Therefore, we can compute the perturbation in the wavefield ($\Delta \mathcal{U}$) as a function of the perturbation in slowness ($\Delta \mathcal{S}$) like this:

$$\Delta \mathcal{U} = (\mathcal{I} - \mathcal{T}_0)^{-1} \mathcal{T}_0 \mathcal{G}_0 \widehat{\mathcal{U}}_0 \Delta \mathcal{S} \quad (\text{B-9})$$

2. Imaging

As for the background image, the perturbation in image (Δr^z), caused by the perturbation in slowness, is obtained by a summation over all the frequencies (ω):

$$\Delta r^z = \sum_1^{N_\omega} \Delta u^z(\omega) \quad (\text{B-10})$$

We can write Equation (B-10) in matrix form as

$$\Delta \mathcal{R} = \mathcal{H} \Delta \mathcal{U} \quad (\text{B-11})$$

$$\begin{pmatrix} \Delta r^1 \\ \Delta r^2 \\ \dots \\ \Delta r^{N_z} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 & | & 1 & 0 & \dots & 0 & | & \dots & | & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & | & 0 & 1 & \dots & 0 & | & \dots & | & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & | & \dots & \dots & \dots & \dots & | & \dots & | & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & | & 0 & 0 & \dots & 1 & | & \dots & | & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{U}(\omega_1) \\ \Delta \mathbf{U}(\omega_2) \\ \dots \\ \Delta \mathbf{U}(\omega_{N_\omega}) \end{pmatrix}$$

where

$$\Delta \mathcal{U} = \begin{pmatrix} \Delta \mathbf{U}(\omega_1) \\ \Delta \mathbf{U}(\omega_2) \\ \dots \\ \Delta \mathbf{U}(\omega_{N_\omega}) \end{pmatrix}; \mathcal{G}_0 = \begin{pmatrix} \mathbf{G}_0(\omega_1, \mathbf{s}_0) & 0 & \dots & 0 \\ 0 & \mathbf{G}_0(\omega_2, \mathbf{s}_0) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{G}_0(\omega_{N_\omega}, \mathbf{s}_0) \end{pmatrix}; \widehat{\mathcal{U}}_0 = \begin{pmatrix} \widehat{\mathbf{U}}_0(\omega_1) & 0 & \dots & 0 \\ 0 & \widehat{\mathbf{U}}_0(\omega_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \widehat{\mathbf{U}}_0(\omega_{N_\omega}) \end{pmatrix}; \Delta \mathcal{S} = \begin{pmatrix} \Delta \mathbf{S} \\ \Delta \mathbf{S} \\ \dots \\ \Delta \mathbf{S} \end{pmatrix}$$

8

$$\mathcal{U}_0 = \begin{pmatrix} \mathbf{U}_0(\omega_1) \\ \mathbf{U}_0(\omega_2) \\ \dots \\ \mathbf{U}_0(\omega_{N_\omega}) \end{pmatrix}; \widehat{\mathcal{U}}_0 = \begin{pmatrix} \widehat{\mathbf{U}}_0(\omega_1) & 0 & \dots & 0 \\ 0 & \widehat{\mathbf{U}}_0(\omega_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \widehat{\mathbf{U}}_0(\omega_{N_\omega}) \end{pmatrix}$$

- $\Delta \mathcal{R}$ is a column vector containing the perturbation in image at every depth level z .⁹

Therefore, the perturbation in image ($\Delta \mathcal{R}$), corresponding to the perturbation velocity field ($\Delta \mathcal{S}$), can be computed as follows:

$$\Delta \mathcal{R} = \mathcal{H}(\mathcal{I} - \mathcal{T}_0)^{-1} \mathcal{T}_0 \mathcal{G}_0 \widehat{\mathcal{U}}_0 \Delta \mathcal{S} \quad (\text{B-12})$$

Adjoint operator: Back-projection

Equation (B-12) enables us to compute the perturbation in image corresponding to a given perturbation in the velocity field, that is, to migrate the scattered field. The operation adjoint to migration, back-projection, can be derived from the same equation using the adjoint state system:

$$\overline{\Delta \mathcal{S}} = [\mathcal{H}(\mathcal{I} - \mathcal{T}_0)^{-1} \mathcal{T}_0 \mathcal{G}_0 \widehat{\mathcal{U}}_0]' \Delta \mathcal{R} \quad (\text{B-13})$$

where

- $\overline{\Delta \mathcal{S}}$ is the back-projected perturbation in the velocity field obtained from the perturbation in the image.

Therefore, the back-projected perturbation in velocity ($\overline{\Delta \mathcal{S}}$) is derived from the perturbation in the reflectivity image ($\Delta \mathcal{R}$) using the following expression:

$$\overline{\Delta \mathcal{S}} = \widehat{\mathcal{U}}_0' \mathcal{G}_0' \mathcal{T}_0' [(\mathcal{I} - \mathcal{T}_0)^{-1}]' \mathcal{H}' \Delta \mathcal{R} \quad (\text{B-14})$$

Equations (B-12) and (B-14) comprise a pair of adjoint operators that relate the perturbation in velocity to the perturbation in reflectivity image. We can use these two operators to invert for velocity from measurable perturbations in the image.

9

$$\Delta \mathcal{R} = \begin{pmatrix} \Delta r^1 \\ \Delta r^2 \\ \dots \\ \Delta r^{N_z} \end{pmatrix}$$

Preconditioning tau tomography with geologic constraints

Robert G. Clapp and Biondo Biondi¹

ABSTRACT

Seismic tomography is a non-linear problem with a significant null-space. Our estimation problem often converges slowly, to a geologically unreasonable model, or not at all. One reason for slow or non-convergence is that we are attempting to simultaneously estimate reflector position (mapping velocity) and image our data (focusing velocity). By performing tomography in vertical travel-time space, we avoid estimating mapping velocity, instead concentrating on focusing velocity. By introducing anisotropic preconditioning oriented along bedding planes, we can quickly guide the inversion towards a geologically reasonable model. We illustrate the benefits of our tomography method by comparing it to more traditional methods on a synthetic anticline model. In addition, we demonstrate the method's ability to improve the velocity estimate, and the resulting migrated image of a real 2-D dataset.

INTRODUCTION

Tomography is inherently non-linear, therefore a standard technique is to linearize the problem by assuming a stationary ray field (Stork and Clayton, 1991). Unfortunately, we must still deal with the coupled relationship between reflector position and velocity (Al-Chalabi, 1997; Tieman, 1995). As a result, the back projection operator must attempt to handle both repositioning of the reflector *and* updating the velocity model (van Trier, 1990). The resulting back projection operator is sensitive to our current guess at velocity and reflector position.

In addition to non-linearity, tomography problems are often under-determined. To create more geologically feasible velocity models and to speed up convergence, Michelena (1991) suggested using varying sized grid cells. Unfortunately, such a parameterization is prone to error when the wrong size blocks are chosen (Delprat-Jannaud and Lailly, 1992). Other authors have suggested locally clustering grid cells (Carrion, 1991) or characterizing the velocity model as a series of layers (Kosloff et al., 1996). These methods are also susceptible to errors when the wrong parameterization is chosen. An attractive alternative approach is to add an additional model regularization term to our objective function (Toldi, 1985). In theory, this regularization term should be the inverse model covariance matrix (Tarantola, 1987). The question is how to obtain an estimate of the model covariance matrix. The obvious answer is through *a priori* information sources such a geologist's structural model of the area, well log information, or preliminary stack or migration results. Incorporating these varied information

¹email: bob@sep.stanford.edu, biondo@sep.Stanford.EDU

sources into our objective function has always been problematic. For years, geostatisticians have successfully combined these mixed types of information to produce variograms (Issaks and Srivastava, 1989). Unfortunately, the geostatistical approach does not easily fit within a standard global tomography problem. In this paper we follow the course outlined in Clapp and Biondi (1998) to address both the velocity-depth ambiguity and the problem of adding geologic constraints. We formulate our tomography problem in vertical travel-time (τ) coordinates rather than depth. In this coordinate system, reflectors are significantly less sensitive to velocity (Biondi et al., 1997) and the resulting back projection operator is less sensitive to the background velocity model (Clapp and Biondi, 1999). We make the assumption that velocity follows geologic dip or some other known trend. We then approximate the model covariance matrix by creating small, plane-wave annihilation filters (Claerbout, 1992), or *steering filters* oriented along geologic dip (Clapp et al., 1997, 1999). To speed up convergence, we reformulate our regularization problem to a preconditioned problem (Claerbout, 1998a) using the helix transform and polynomial division (Claerbout, 1998c).

We create a synthetic anticline velocity model and compare the inversion result using a symmetric regularization operator in depth, steering filter in depth, and finally steering filter in vertical travel time space. We study the speed and quality of our tomographic estimate using two different synthetic models. We conclude with some preliminary tests on a 2-D marine dataset with gas hydrates. Preliminary migration results are encouraging.

THEORY

Following the method described in Clapp (1998), we began by linearizing the tomography problem around an initial guess at our slowness model \mathbf{s}_0 . We assumed ray stationarity and described the change in travel time ($\Delta\mathbf{t}$) as being linearly related to our change in slowness ($\Delta\mathbf{s}$):

$$\Delta\mathbf{t} \approx \mathbf{T}_z \Delta\mathbf{s}. \quad (1)$$

\mathbf{T}_z is composed of two portions. The first, $\mathbf{T}_{z,\text{ray}}$ simply applies

$$\delta t = \tilde{l} \delta s, \quad (2)$$

or that the change in the travel time is (δt) is equal to change in slowness (δs) times length of the ray of the ray segment (\tilde{l}) of the ray connecting the source, reflector, and the receiver. The second component, $\mathbf{T}_{z,\text{ref}}$, can be thought of as a chain of two operators: the first maps our change in slowness ($\Delta\mathbf{s}$) into reflector movement, the second maps the reflector movement into our change in travel times ($\Delta\mathbf{t}$) (van Trier, 1990). This second term amounts to performing residual migration and can be done by back projecting a ray located at the reflection point perpendicular to the reflector (Stork, 1994), Figure 1. Taking both components into account our tomography fitting goal becomes

$$\Delta\mathbf{t} \approx (\mathbf{T}_{z,\text{ray}} + \mathbf{T}_{z,\text{ref}}) \Delta\mathbf{s}. \quad (3)$$

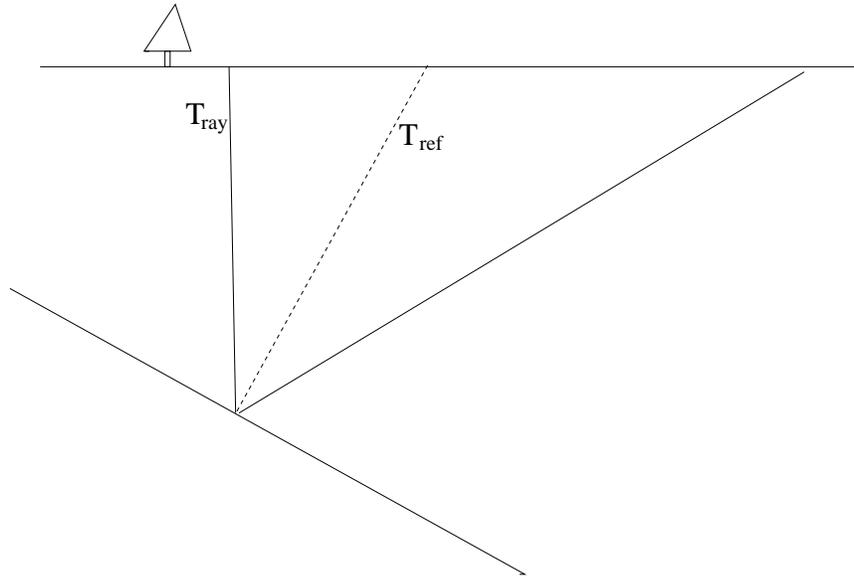


Figure 1: The two portions of the back projection operator. $\mathbf{T}_{z,\text{ray}}$ is the pair of rays through \mathbf{s}_0 from the source to the receiver that obey Snell's law at the reflector. $\mathbf{T}_{z,\text{ref}}$ is the raypath from this reflection point to the surface. `bob1-schematic` [NR]

Smoothing slowness rather than change of velocity

In general, the tomography problem is under-determined and requires some type of regularization. Ideally, this regularization should be the inverse model covariance (Tarantola, 1987) but that is not readily available. In many cases we do have well logs, initial migration surfaces, or a geologist's model of the region that can at least indicate the trend that velocity should follow. Following the method described in Clapp, et al. (1999) we can build a space-varying operator composed of small plane wave annihilation filters that can smooth our velocity along this pre-determined trend. The problem is that our model is not slowness, but change in slowness. To a degree, we can get around this problem by following the method similar to the one described by Bevc (1994). We start by stating our goal to smooth the slowness field:

$$\mathbf{0} \approx \mathbf{A}\mathbf{s} \quad (4)$$

where \mathbf{A} is our steering filter operator. But \mathbf{s} is actually $\mathbf{s}_0 + \Delta\mathbf{s}$, so we can write a new regularization goal as

$$\begin{aligned} \mathbf{0} &\approx \mathbf{A}(\mathbf{s}_0 + \Delta\mathbf{s}) \\ -\mathbf{A}\mathbf{s}_0 &\approx \mathbf{A}\Delta\mathbf{s}. \end{aligned} \quad (5)$$

A problem with this method is where the adjoint of our modeling operator (\mathbf{T}'_z) does not contribute at all to the model we can introduce artifacts. Our best solution to date for this problem is to introduce a smooth masking operator that tapers off to zero in locations unaffected by \mathbf{T}'_z .

Preconditioning

The proposed regularized tomography problem still has the problem of slow convergence. By reformulating the problem in helix space (Claerbout, 1998c), we can take advantage of 1-D theory to change our regularized problem into a preconditioned one. We start by defining a new variable \mathbf{p} :

$$\mathbf{p} = \mathbf{A} \Delta \mathbf{s}. \quad (6)$$

By applying polynomial division to our steering filters, we can create \mathbf{A}^{-1} which becomes a smoothing operator. We can then rewrite our fitting goals as

$$\begin{aligned} \Delta \mathbf{t} &\approx (\mathbf{T}_{\mathbf{z},\text{ray}} + \mathbf{T}_{\mathbf{z},\text{ref}}) \mathbf{A}^{-1} \mathbf{p} \\ -\mathbf{A} \mathbf{s}_0 &\approx \epsilon \mathbf{p}. \end{aligned} \quad (7)$$

Tau tomography

In depth tomography we must constantly deal with the depth-velocity ambiguity problem. Put another way, we are simultaneously trying to estimate both a focusing (S_f) and a mapping (S_m) slowness. Biondi et al.(1997) showed that by mapping (z, x) into (τ, x) through

$$\tau(z, x) = \int_0^z 2S(z', x) \delta z' \quad (8)$$

we can write a **focusing eikonal** equation which only indirectly depends on the mapping velocity

$$4 \left(\frac{\partial t(\tau, x)}{\partial \tau} \right) + S_f(\tau, x)^{-2} \left(\frac{\partial t(\tau, x)}{\partial x} + \sigma_m(\tau, x) \frac{\partial t(\tau, x)}{\partial \tau} \right)^2 = 1 \quad (9)$$

where σ_m is the differential mapping operator defined as

$$\sigma_m(\tau, x) = \int_0^\tau S_m^{-1}(\tau', x) \frac{\partial}{\partial x} S_f(\tau', x) \delta \tau'. \quad (10)$$

From this eikonal equation we can derive a new relation for the change in travel time due to a change in the focusing velocity:

$$\delta t = \left(\frac{\delta \tilde{x}^2}{\tilde{l}} \right) \delta s - (\delta \tilde{x} (\delta \tilde{\tau} - \delta \tilde{x} \tilde{\sigma}_0(x, \tau))) \delta \sigma \quad (11)$$

where $\delta \tilde{x}$ and $\delta \tilde{\tau}$ are the change in x and τ position of the ray segment, $\tilde{\sigma}_0(x, \tau)$ is the differential mapping factor of our initial slowness model at the ray location, and $\delta \sigma$ is defined as

$$\delta \sigma(\tau, x) = \frac{\sigma_0(\tau, x)}{s_0(\tau, x)} \delta s(\tau, x) - 2s_0(\tau, x) \frac{\delta(\delta z(\tau, z))}{\delta x} \quad (12)$$

where

$$\delta z(\tau, z) = \int_0^\tau \frac{\delta s(\tau', x)}{2s_0^2(\tau', x)} \delta \tau. \quad (13)$$

We now have a way to back project travel time errors and can write a new set of fitting goals,

$$\begin{aligned} \Delta \mathbf{t} &\approx (\mathbf{T}_{\tau, \text{ray}} + \mathbf{T}_{\tau, \text{ref}}) \mathbf{A}^{-1} \mathbf{p} \\ -\mathbf{A} \mathbf{s}_0 &\approx \epsilon \mathbf{p} \end{aligned} \quad (14)$$

where $\mathbf{T}_{\tau, \text{ray}}$ and $\mathbf{T}_{\tau, \text{ref}}$ use (11) rather than (2) to back project.

SYNTHETIC TESTS

To test the effectiveness of the method we create a simple anticline synthetic velocity model, Figure 2. We simulated six reflectors, one on top, four in the anticline, and a basement reflector. We calculated travel times to all the reflectors for an offset range of 4 kms. These travel times represent our ‘recorded travel times’. For our initial model we created a $v(z)$ model by taking the lateral average of the velocity field, Figure 2. From this initial model we attempted

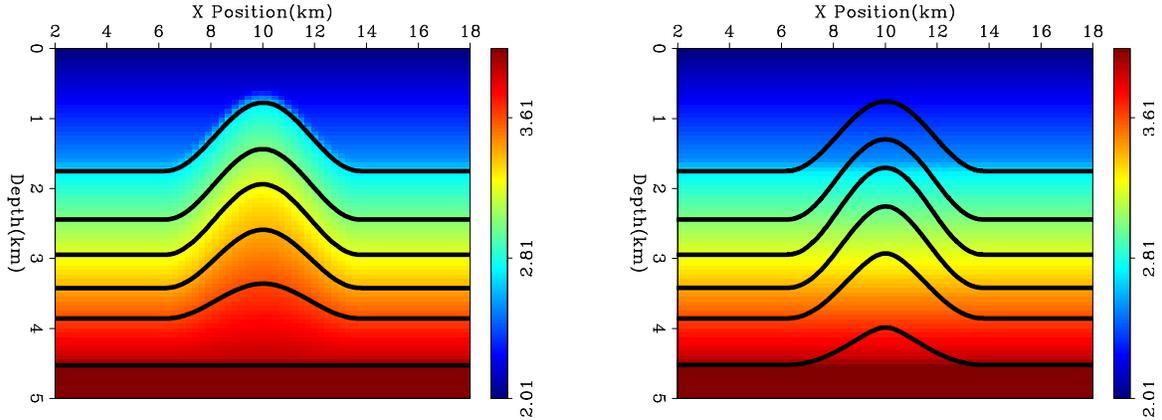


Figure 2: Left panel is our synthetic model superimposed by the six reflectors. The right panel is our starting guess for our velocity function and the map migrated reflector position using this initial velocity estimate. `bob1-model` [CR]

to invert the velocity function by three progressively advanced methods:

Depth-Standard : Inverted for a depth model, using an inverse Laplacian preconditioner (Claerbout, 1998b) for \mathbf{A}^{-1} in our depth fitting goals (7)

Depth-Steering : Inverting for a depth model, using a steering filter operator for our preconditioner in our depth fitting goals (7)

Tau-Steering : Inverting for a tau model, using a steering filter operator for our preconditioner in tau fitting goals (14)

Figure 3 shows the result of one non-linear iteration for all three inversions schemes. All three methods were able to recover the dome shape after one iteration. When using a Laplacian smoother the velocity increase is spread too far both laterally and vertically. As a result, the

bottom reflector is located too deep throughout the model. When using steering filters we still have a significant velocity-depth ambiguity problem, but we have done a little better job position the bottom reflector. In the case of tau tomography with steering filters we have done almost a perfect job after a single iteration. We have not perfectly recovered the lower portion of the anticline structure but we have almost completely flattened the bottom reflector. To see

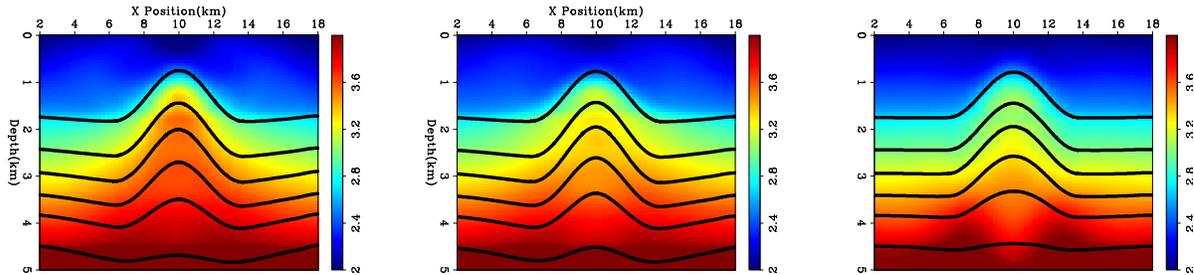


Figure 3: Left panel is Depth-Standard, middle is Depth-Steering, right is Tau-Steering. All after 1 non-linear iteration. `bob1-model1-iter1` [CR]

if, and how fast, we could converge to the correct solution in depth we performed several more non-linear iterations. As Figure 4 shows we did a decent job recovering the anticline with all three methods. In this case, where the model is fairly simple and we have good travel time coverage, the big advantage seems to be speed. We got a high quality result with steering filter tau tomography in a single iteration, while it took three with steering filter depth tomography, and four when using the Laplacian and depth tomography. The smoothness of the anticline was well suited for the Laplacian so we decided on a slightly more difficult challenge that could better differentiate between a Laplacian and steering filter regularization. Our new model keeps the same basic shape for the model but adds a low velocity layer within the anticline. Figure 5 shows the correct, initial, and the result of 4 iterations using both the Laplacian and steering filters to precondition the problem. After 4 iterations the steering filters have done a much better job recovering the low velocity layer.

TEST ON REAL DATA

We next decided to test the method on real data. For this initial test we decided to work with a relatively clean data which still had some residual move-out in the common reflection point (CRP) gathers. The data is from the Blake Outer Ridge as was used by Ecker(1998) to characterize methane hydrate structures. For our initial velocity model we used Ecker's Dix (1955) derived model, Figure 6. Our general philosophy was to limit human time as much as possible. Therefore we chose to do tau migration (Alkhalifah, 1998) using a generic Kirchhoff package(Biondi, 1998). By using tau rather than depth migration, we were quickly able to compare CRPs from iteration to iteration and it allowed us to pick reflector positions only once. After performing the migration we picked six reflectors, Figure 7. We picked the sea floor, a strong reflector above the bottom simulating reflector (BSR), the BSR itself, the flat reflector below the BSR, and two deeper reflectors. Rather than pick move-out differences we decided to create residual semblance panels at each reflector location, Figure 8. The panels

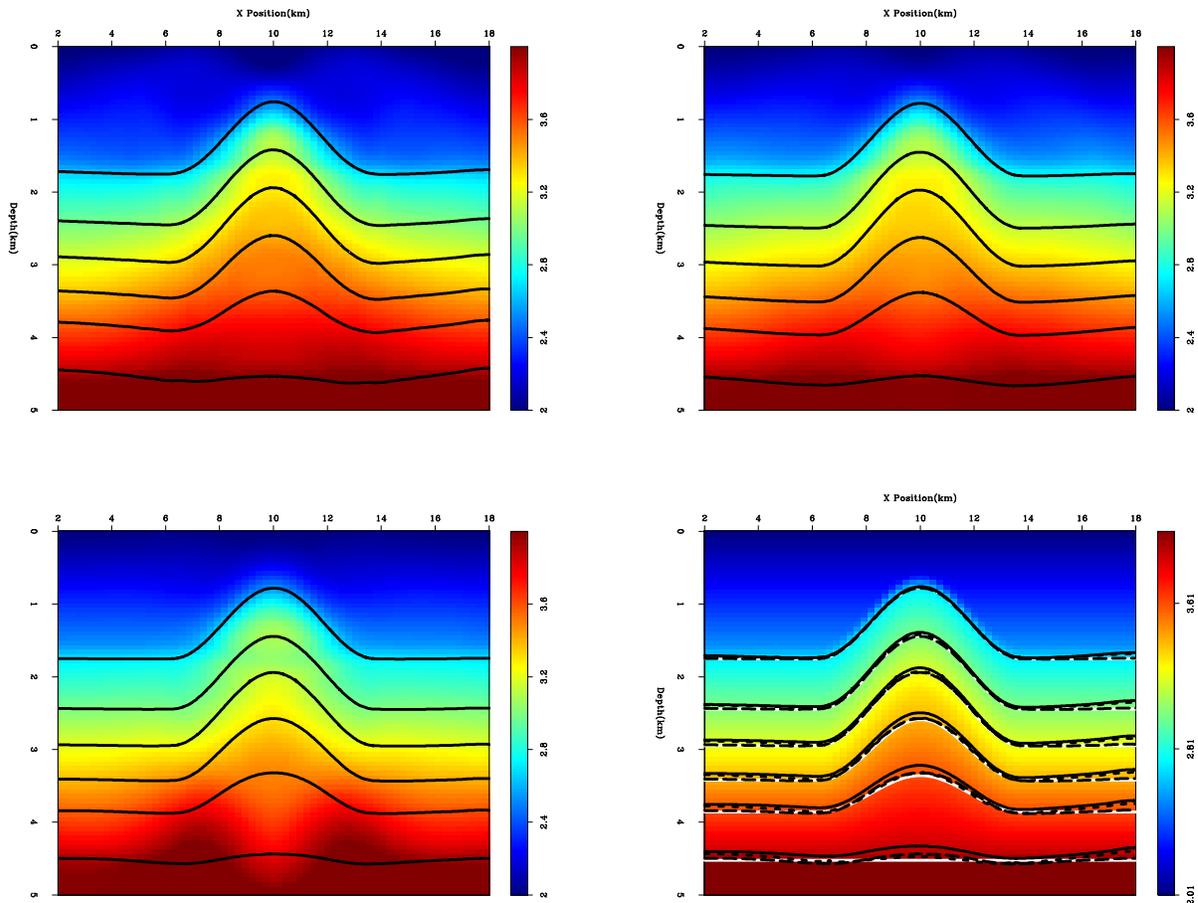


Figure 4: Top-left: Depth-Standard, after 4 non-linear iterations; top-right: Depth-Steering after 3 iterations; bottom-left: Tau-steering after 1 iteration; and bottom-right: a comparison of the reflector positions using all 3 methods. The solid, white line is correct reflector position, the small dashes represent Tau-Steering; large-dashes:Depth-Steering; and the solid black line is Depth-Standard. `bob1-model1-best` [CR]

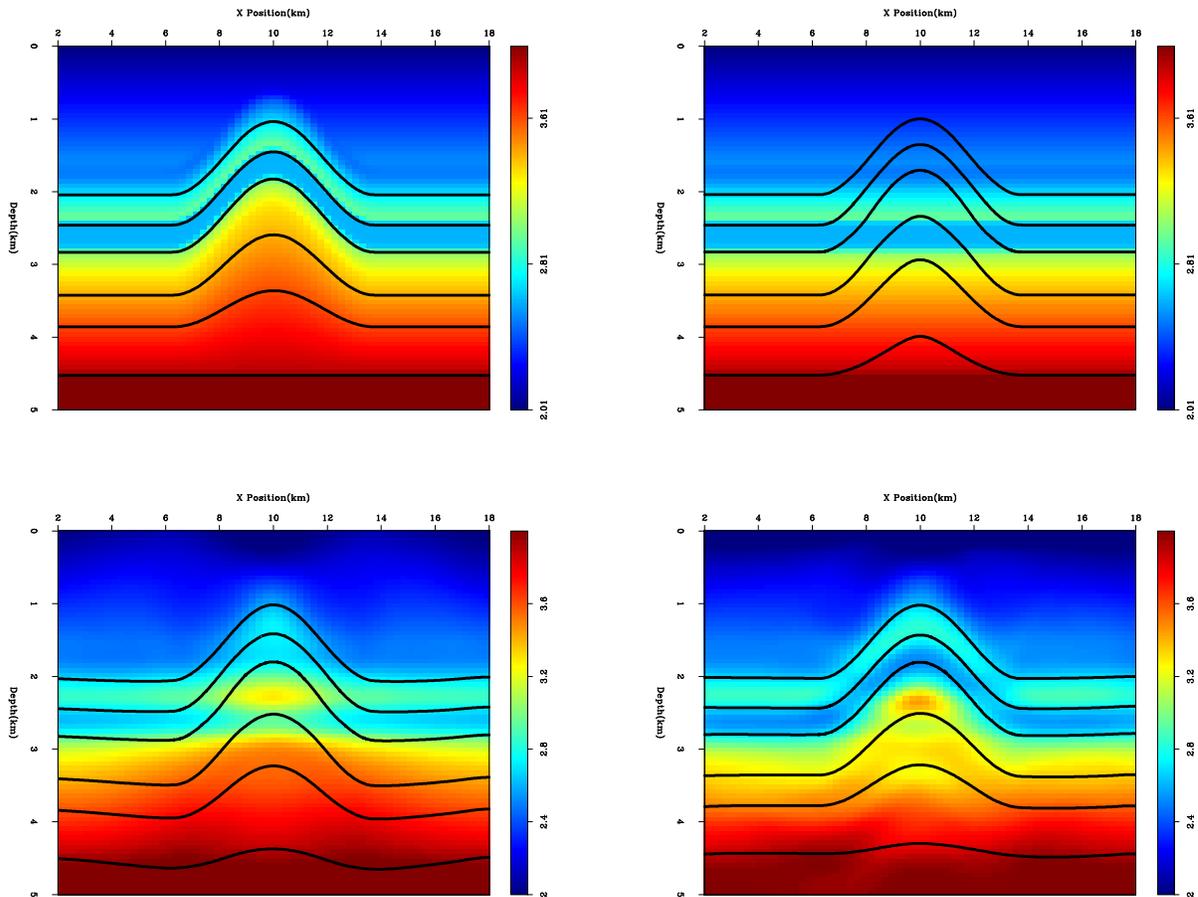


Figure 5: Top, left Our new model with a low velocity layer within the anticline; top-right, our starting model; bottom-left, Depth-Standard after 4 iterations; bottom-right Depth-Steering after 4 iterations. `bob1-model2` [CR]

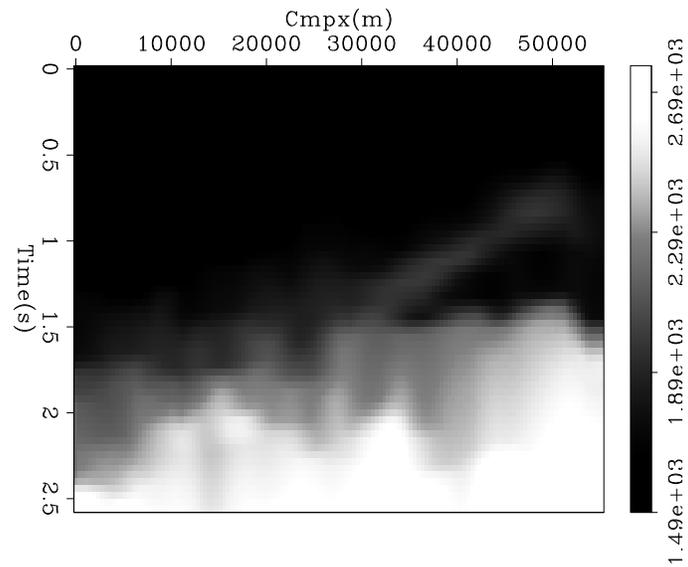


Figure 6: Initial velocity model in depth. Note the low velocity zone caused by the gas hydrate starting at approximately 32000 kms and extending to the end of the section. `bob1-christine-vel0` [ER]

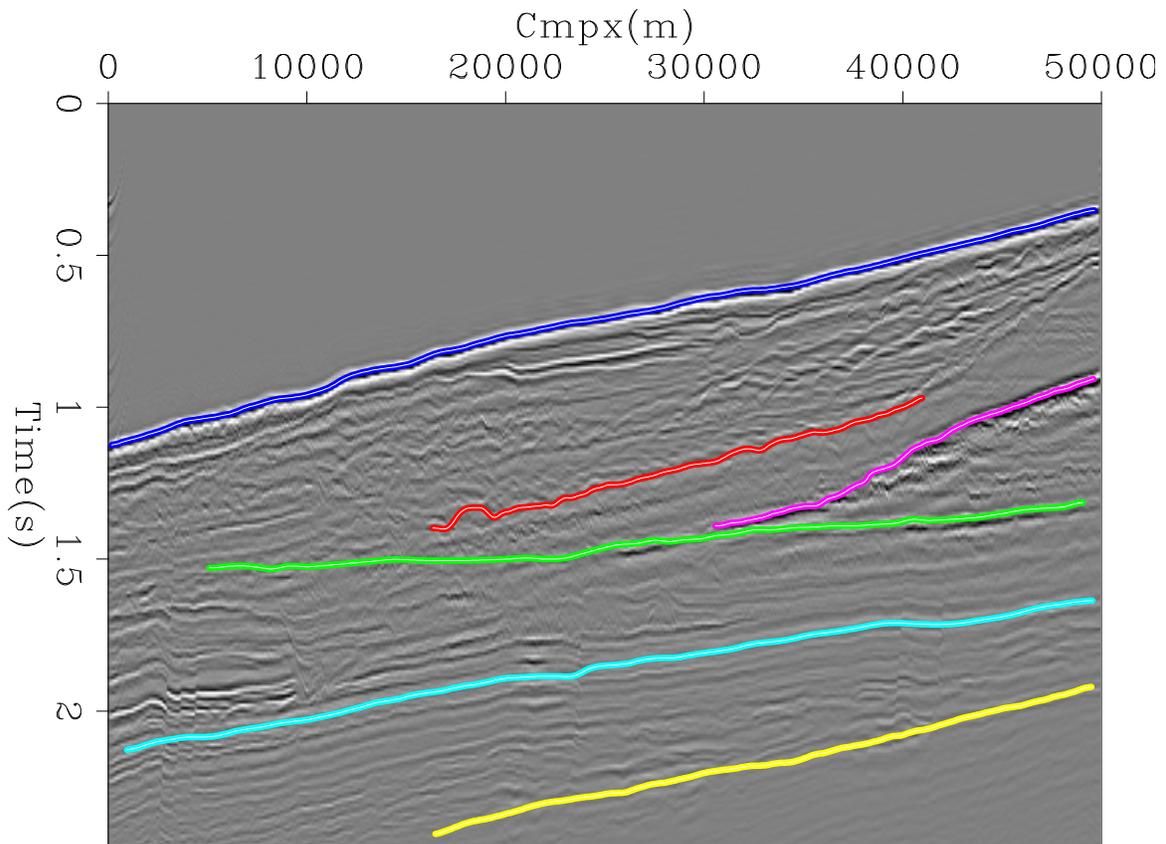


Figure 7: Initial stack overlaid by reflectors picked for tomography. `bob1-stack` [CR]

indicate that there is significant residual curvature, especially where the BSR meets the lower reflector. From these semblance panels we picked smooth curves at approximately the maximum semblance at each reflector. To check to see if a single parameter adequately described the move-outs we back projected the picked semblance into our CRP gathers. Figure 9 shows that the semblance picks did a fairly good job describing the move-out. We used our picked

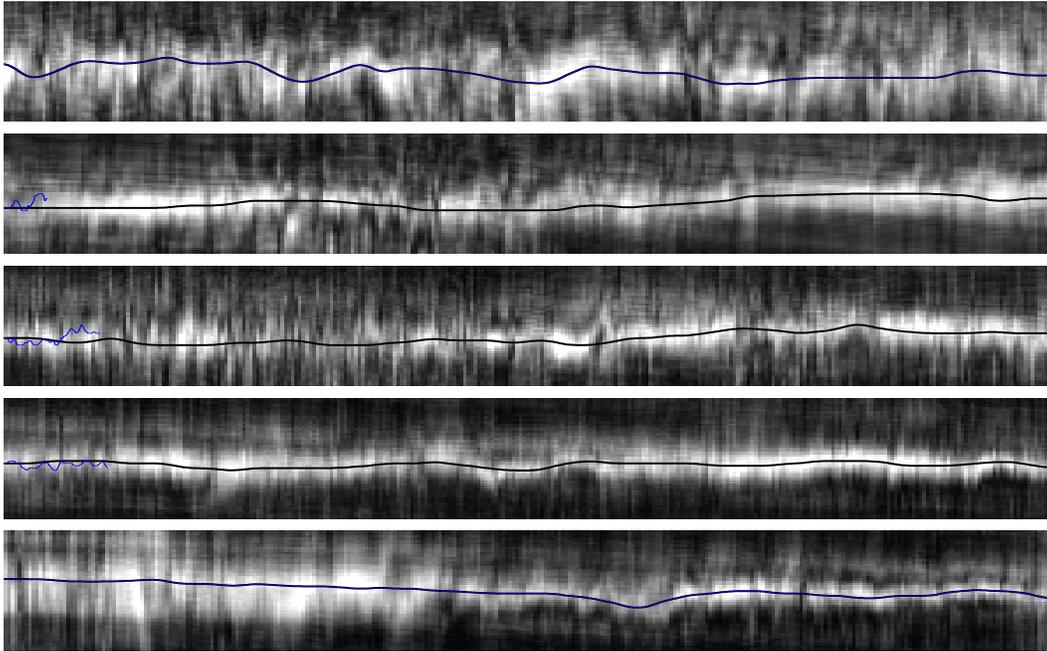


Figure 8: Residual semblance panels for the bottom 5 reflectors. The black line in each panel represents the picked maximum. `bob1-sem-vel0` [CR]

reflectors to construct our steering filters and then applied our tau tomography fitting goals (14). Generally, we have increased velocity, Figure 10, but the changes still keep velocity following reflector dip. The next step is to see if our new velocity model flattens our CRP gathers and improves the focusing of the data. Figures 11 and 12 indicate that we have accomplished both of these goals. Figure 9 shows that all of our reflectors are significantly flatter, with only significant curvature left along the BSR. Figure 12 shows a much more continuous BSR reflection along with overall improved focusing of the section above and below.

CONCLUSIONS

By performing tomography in tau space we are able to quickly converge to geologically realistic velocity models. The results on synthetics indicate that the method when velocity is not the smooth function that the a Laplacian regularize will attempt to create. Early tests on field data are encouraging.

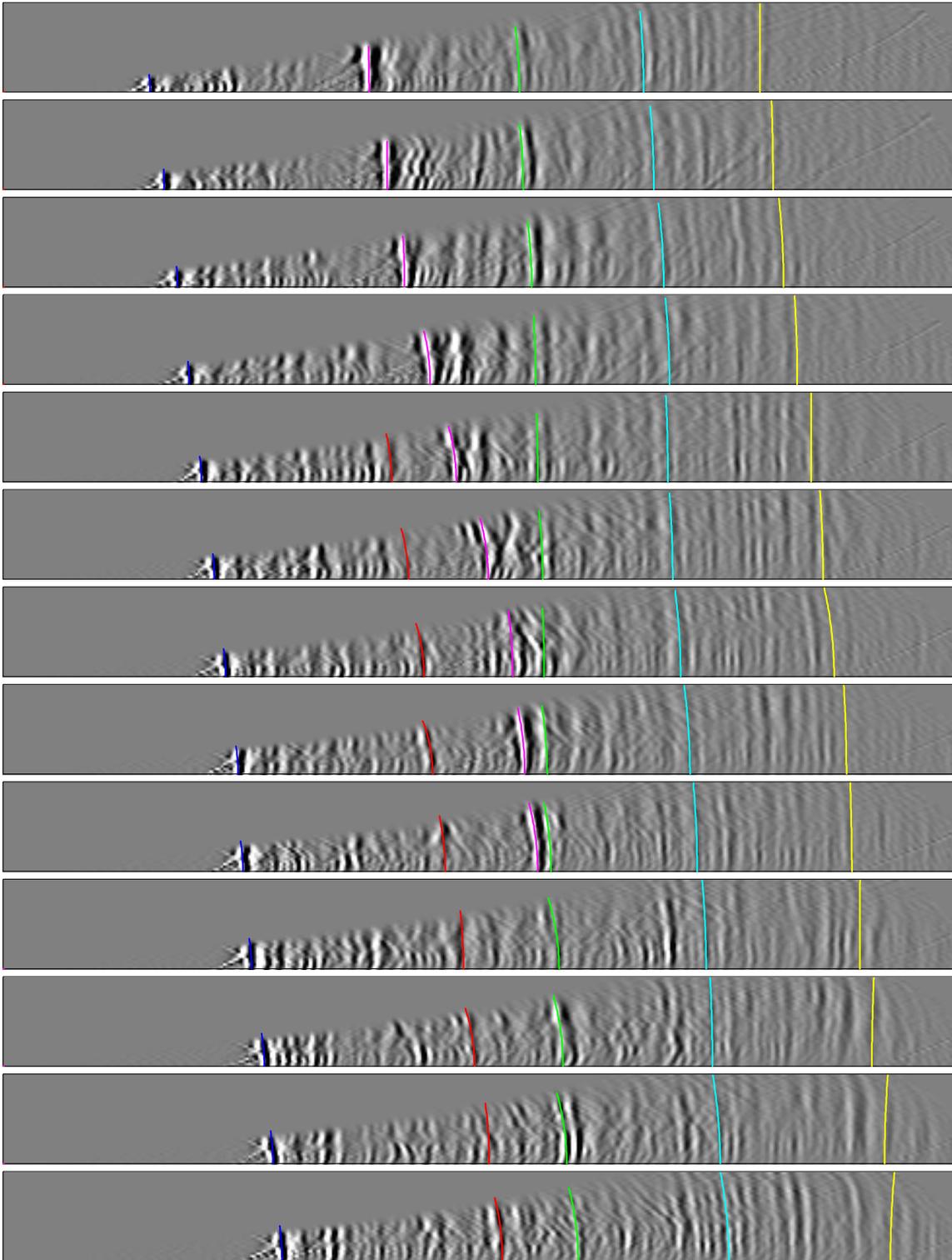


Figure 9: Common reflection point gathers from every 2000 meters starting from 28000. The lines are the result of mapping back the picked residual slowness values. Note how the curves do an excellent job matching the actual reflector move-out. `bob1-overlay.vel0` [CR]

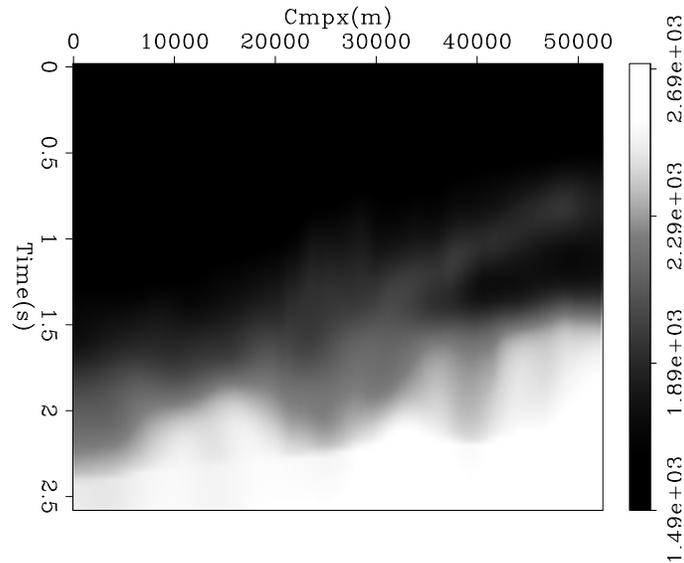


Figure 10: The velocity after 1 iteration of tau-steering tomography `bob1-christine-vel1` [CR]

REFERENCES

- Al-Chalabi, M., 1997, Parameter nonuniqueness in velocity versus depth functions: *Geophysics*, **62**, no. 03, 970–979.
- Alkhalifah, T., 1998, Prestack Kirchhoff time migration for complex media: *SEP-97*, 45–60.
- Bevc, D., 1994, Datuming velocity from traveltimes tomography: *SEP-82*, 145–164.
- Biondi, B., Fomel, S., and Alkhalifah, T., 1997, “Focusing” eikonal equation and global tomography: *SEP-95*, 61–76.
- Biondi, B., 1998, Genkir3D: A toolkit for Kirchhoff imaging: *SEP-97*, 371–402.
- Carrion, P., 1991, Dual tomography for imaging complex structures: *Geophysics*, **56**, no. 9, 1395–1404.
- Claerbout, J. F., 1992, *Earth Soundings Analysis: Processing versus Inversion*: Blackwell Scientific Publications.
- Claerbout, J. *Geophysical Estimation by Example: Environmental soundings image enhancement*: <http://sepwww.stanford.edu/sep/prof/>, 1998.
- Claerbout, J., 1998b, Multidimensional recursive filters via a helix: *SEP-97*, 319–335.
- Claerbout, J. F., 1998c, Multi-dimensional recursive filtering via the helix: *Geophysics*, **63**, no. 5, 1532–1541.
- Clapp, R. G., and Biondi, B. L., 1998, Regularizing time tomography with steering filters: *SEP-97*, 137–146.

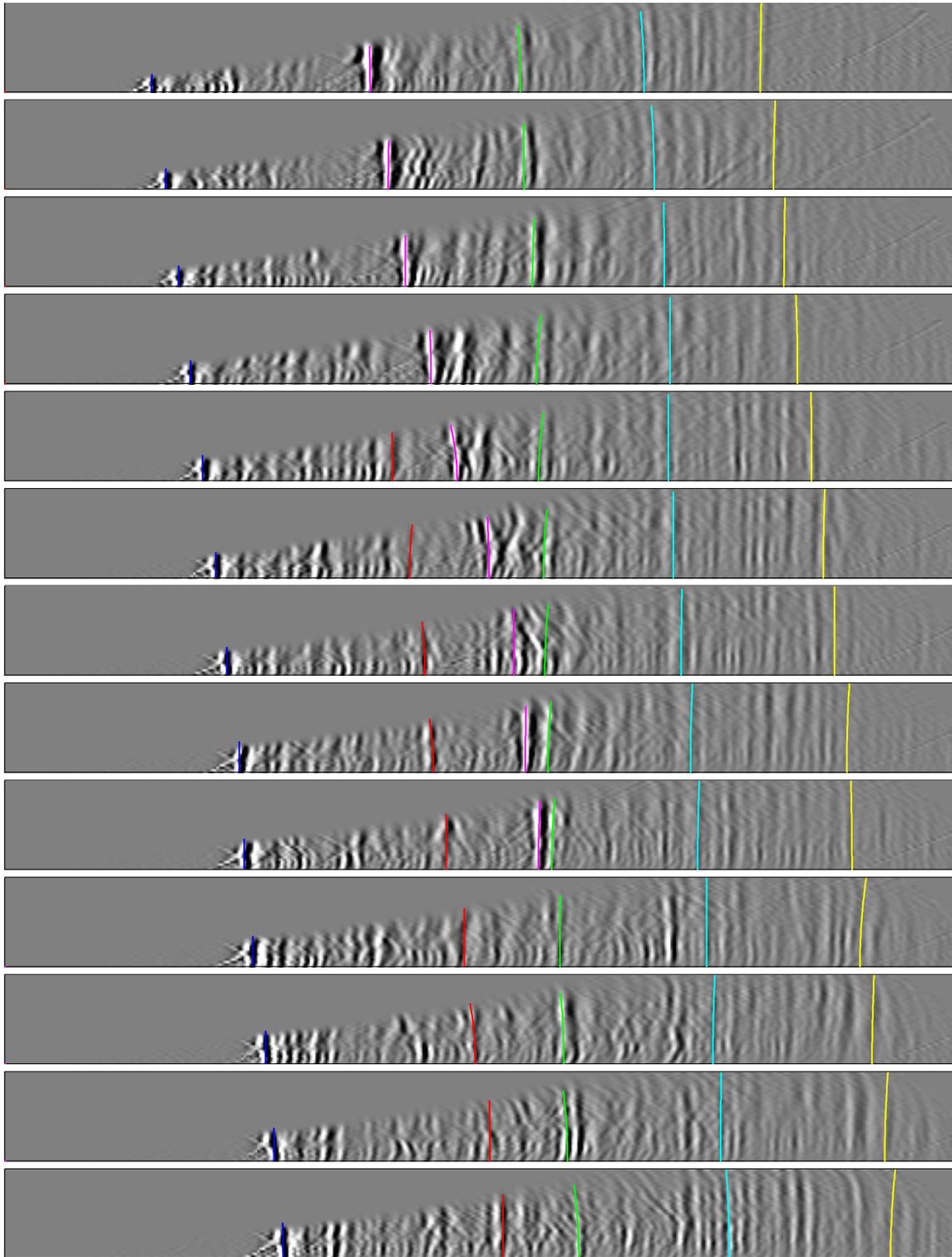


Figure 11: CRP gather from 24000-48000 meters. Note how they are considerably flatter than Figure 9. `bob1-overlay.vel1` [CR]

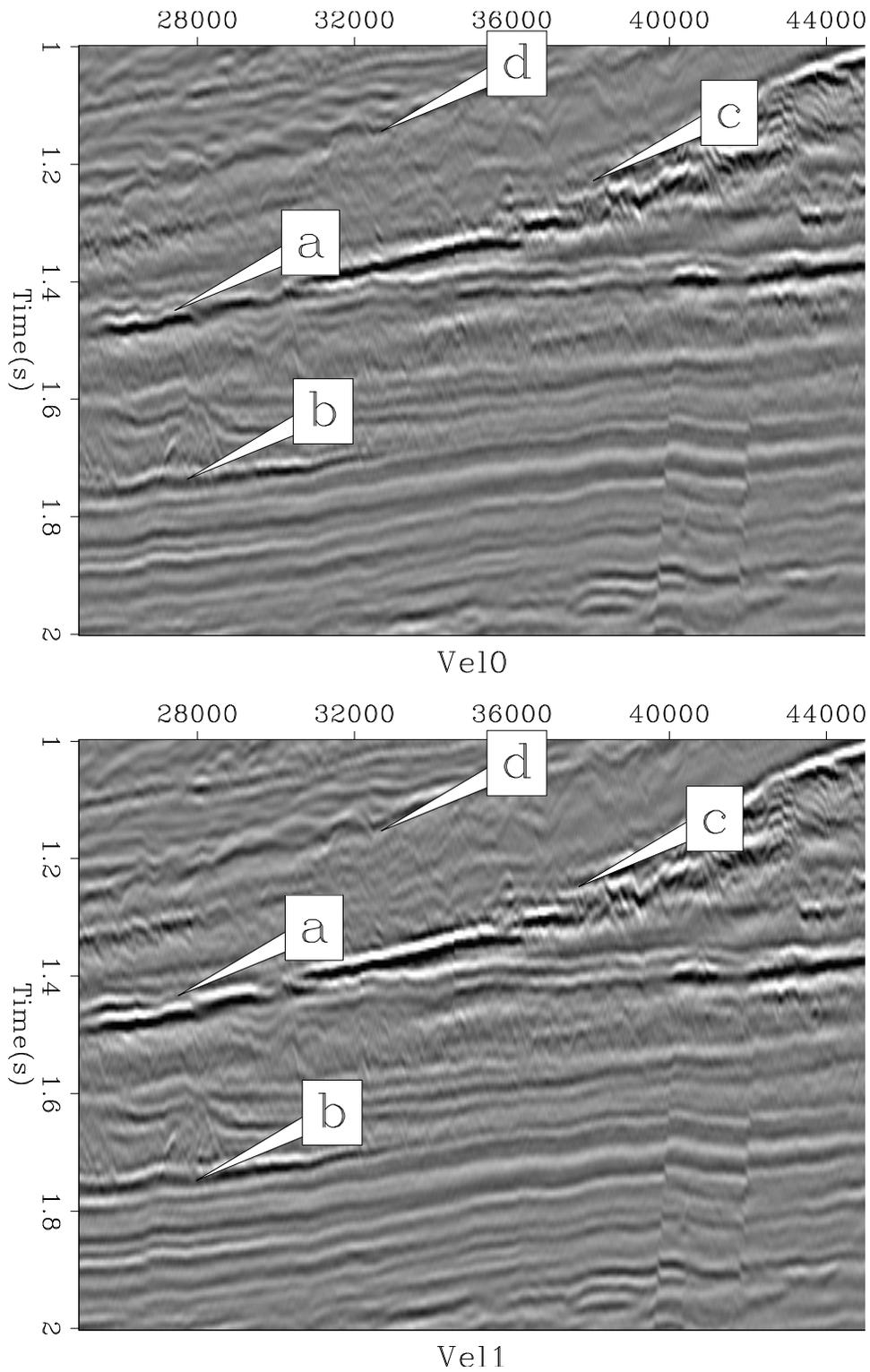


Figure 12: The stack using our initial velocity and the velocity after 1 iteration of tau steering tomography. Note how the reflectors are generally better focused at a, b, c, and d.

`bob1-stack-comp` [CR]

- Clapp, R. G., and Biondi, B., 1999, Why tau tomography is better than depth tomography: SEP-100, 51-58.
- Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP-95, 27-42.
- Clapp, R. G., Fomel, S., Crawley, S., and Claerbout, J. F., 1999, Directional smoothing of non-stationary filters: SEP-100, 197-209.
- Delprat-Jannaud, F., and Lailly, P., 1992, What information on the earth model do reflection travel times provide?: J. Geophysical Research, **97**, 19827-19844.
- Dix, C. H., 1955, Seismic velocities from surface measurements: Geophysics, **20**, no. 01, 68-86.
- Ecker, C., 1998, Seismic characterization of methane hydrate structures: SEP-96.
- Issaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: Oxford University Press.
- Kosloff, D., Sherwood, J., Koren, Z., Machet, E., and Falkovitz, Y., 1996, Velocity and interface depth determination by tomography of depth migrated gathers: *61*, **5**, no. 1511-1523.
- Michelena, R. J., and Harris, J. M., 1991, Tomographic travelttime inversion using natural pixels: Geophysics, **56**, no. 5, 635-653.
- Stork, C., and Clayton, R. W., 1991, An implementation of tomographic velocity analysis: Geophysics, **56**, no. 4, 483-495.
- Stork, C., 1994, Demonstration of mva tomography with controls and constraints for determining an accurate velocity model for prestack depth migration: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1338-1342.
- Tarantola, A., 1987, Inverse problem theory: Elsevier.
- Tieman, H. J., 1995, Migration velocity analysis: Accounting for the effects of lateral velocity variations: Geophysics, **60**, no. 01, 164-175.
- Toldi, J., 1985, Velocity analysis without picking: SEP-43.
- van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: Ph.D. thesis, Stanford University.

Short Note

Why tau tomography is better than depth tomography

Robert G. Clapp and Biondo Biondi¹

INTRODUCTION

Seismic tomography is a non-linear problem. A standard technique is to iteratively assume a linear relation between the change in slowness and the change in travel times (Biondi, 1990; Etgen, 1990) and then re-linearize around the new model. In ray-based methods, this amounts to assuming stationary ray paths and reflection locations to construct a back projection operator (Stork and Clayton, 1991). The change in this back projection operator from non-linear iteration to non-linear iteration can be thought of as an important second order effect.

By formulating our back projection operator in terms of vertical travel-time (τ) rather than depth our reflector locations become more stable (Biondi et al., 1997; Clapp and Biondi, 1998). We show that the corresponding back projection operator is less sensitive to our initial velocity estimate. Therefore, our back projection operator changes less from non-linear to non-linear iteration, making the estimation less likely to get stuck in local minima.

THEORY

Velocity estimation is fundamentally an inverse problem. The correct solution is to do full wave-form inversion (Tarantola, 1986; Mora, 1987), but is generally impractical. Instead we start from the idea that there is a non-linear operator that relates slowness (\mathbf{s}) and travel time (\mathbf{t}),

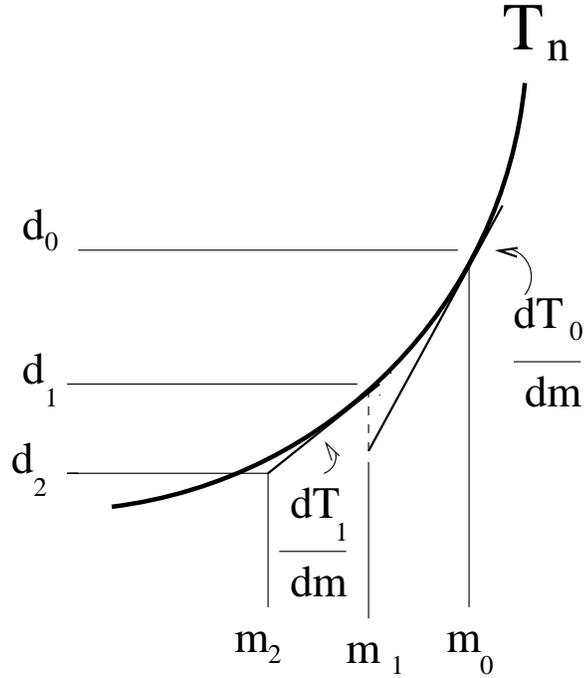
$$\mathbf{t} \approx \mathbf{T}_n \mathbf{s}. \quad (1)$$

We then attempt to approximate \mathbf{T}_n by doing a two term Taylor expansion around our initial guess at the slowness field (a version of Newton's minimization method):

$$\begin{aligned} \mathbf{t} &\approx \mathbf{T}_n \mathbf{s}_0 + \mathbf{T}'_0 \Delta \mathbf{s} \\ \mathbf{t} &\approx \mathbf{t}_0 + \mathbf{T}'_0 \Delta \mathbf{s} \\ \Delta \mathbf{t} &\approx \mathbf{T}'_0 \Delta \mathbf{s}. \end{aligned} \quad (2)$$

¹email: bob@sep.Stanford.EDU, biondo@sep.Stanford.EDU

Figure 1: Newton's method applied to ray based tomography.
 bob2-newton [NR]



where

\mathbf{s}_0 is our initial guess at slowness,

\mathbf{T}'_0 is a linear operator describing the relationship between, slowness and travel times given the initial slowness model. In ray-based methods we usually use some stationary ray paths based on the initial slowness model,

$\Delta \mathbf{s}$ is the change in slowness,

\mathbf{t}_0 are the modeled travel times applying \mathbf{T}'_0 to \mathbf{s}_0 ,

$\Delta \mathbf{t}$ are the difference between the modeled travel times, \mathbf{t}_0 , and the measured travel times, \mathbf{t} .

After inverting for $\Delta \mathbf{s}$, we have a new estimate for our slowness field:

$$\mathbf{s}_1 = \mathbf{s}_0 + \Delta \mathbf{s}. \quad (3)$$

We can then re-linearize around this new model (\mathbf{s}_1), constructing a new tomography operator \mathbf{T}'_1 . We repeat this procedure until $\Delta \mathbf{t}$ is negligible. Figure 1 is a graphical representation of the method. There are two problems with this approach. First, Newton's method is only guaranteed to converge to a local minima, we hope that by applying regularization (Clapp and Biondi, 1999) we can avoid this problem. And second, we are only using the first term in our Taylor expansion, which means that when our higher order derivatives are large, the descent direction will be wrong, and we will converge at a much slower rate. When using rays, this problem occurs when the initial guess at ray paths and reflector locations are too far from their *correct* locations.

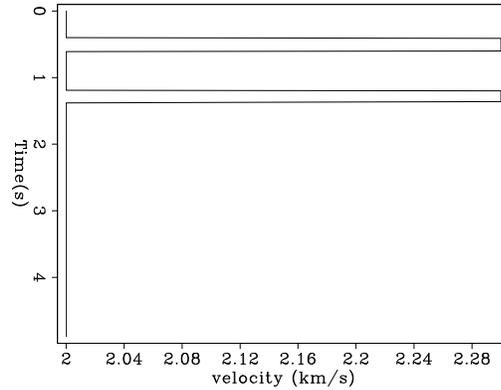


Figure 2: Synthetic 1-D velocity function in τ . `bob2-cor` [ER]

We can obtain a measure of how inaccurate our linear approximation is by looking at how much our linearized tomography operator changes from non-linear iteration to non-linear iteration (the difference between \mathbf{T}'_0 and \mathbf{T}'_1). The smaller the difference, the more accurate our linearization, and the less likely our estimate will diverge. By forming our tomography in (τ, x) rather than (z, x) space, we reduce the change in \mathbf{T}'_1 from \mathbf{T}'_0 . The fundamental reason is that our data is in time rather than depth. In depth, reflector positions and layer boundaries change significantly from iteration to iteration, while in tau, they hardly change at all (Biondi et al., 1997).

SIMPLE TEST

To demonstrate how the tau back projection operator is less affected by our initial slowness model, we constructed a simple 1-D synthetic. The model, Figure 2, is composed of two 2.3 km/s zones in a constant 2 km/s background. For this test we assumed that our slowness model had correctly resolved the bottom anomaly in vertical travel time. Our choice of vertical travel time is quite important, as when doing velocity estimation, we must always preserve zero-offset travel time. In this simple 1-D synthetic, that means that the vertical travel-time to the layer boundaries and to the reflector must be kept constant. Therefore, in depth, we will misplace the location of the bottom high velocity zone but preserve the correct vertical travel times to the layer top and bottom. After constructing the model we found the ray that hit the reflector at 2 km depth, 2 km away from source in both (τ, x) and (z, x) space (Figure 3.) Following the method outlined in Clapp and Biondi (1998), we built the tomography operator for both tau ($\mathbf{T}'_{0,\tau}$) and depth ($\mathbf{T}'_{0,z}$), Figure 4. For comparison, we ray traced through the 'correct' velocity model in both spaces (Figure 5) and calculated the corresponding operators. By comparing the correct and initial operator for tau and depth, or by looking at the difference between the two operators (Figure 7), we can clearly see that our initial guess for our tau operator is overall better than our initial guess for our depth operator. In the upper layer, we see marginally more change in the tau operator but at the lower reflector boundaries (which move in the case of depth but remain constant in tau) we see significantly more error in depth. In addition, the change in reflector position has caused a spike in the difference panel for the depth case. Finally, the change in the tau operator is smooth, while the change in the depth

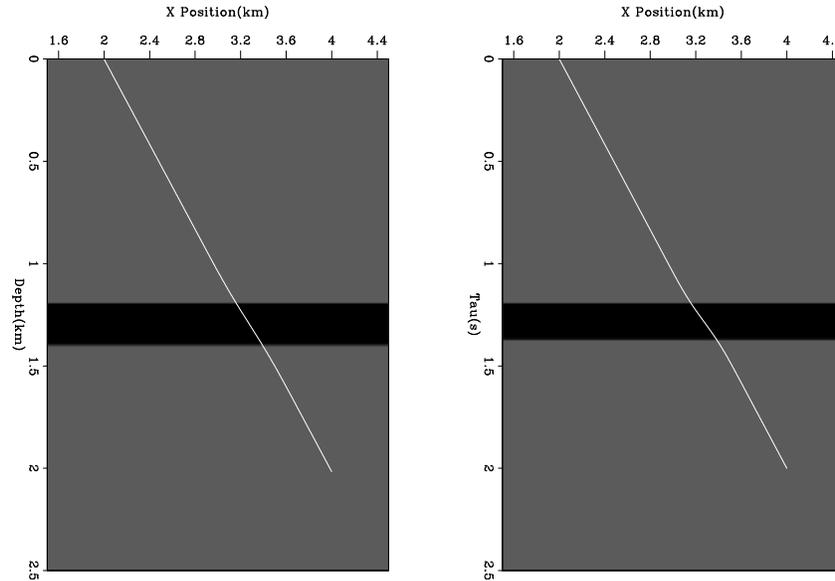


Figure 3: Initial guess at the velocity function overlaid by ray hitting reflector at 4 km with a half-offset of 2 km. Left panel is in depth, right panel is in tau. `bob2-vel0` [ER]

operator shows dramatic jumps. Our successive relinearizing have an underlying assumption that we are smoothly converging to the correct operator. In tau space, this assumption seems to be more valid. With a more complicated model our positioning of layer boundaries, will be subject to more change, making the tau compared to depth difference even more dramatic.

CONCLUSIONS

We showed that for this simple model the tau back projection operator is less affected by our initial velocity estimate than a depth back projection operator. We hypothesize that makes tau tomography to some extent *more* linear and, therefore, less likely to diverge.

REFERENCES

- Biondi, B., Fomel, S., and Alkhalifah, T., 1997, “Focusing” eikonal equation and global tomography: SEP-95, 61–76.
- Biondi, B., 1990, Velocity analysis using beam stacks: Ph.D. thesis, Stanford University.
- Clapp, R. G., and Biondi, B. L., 1998, Regularizing time tomography with steering filters: SEP-97, 137–146.
- Clapp, R. G., and Biondi, B., 1999, Preconditioning tau tomography with geologic constraints: SEP-100, 35–50.

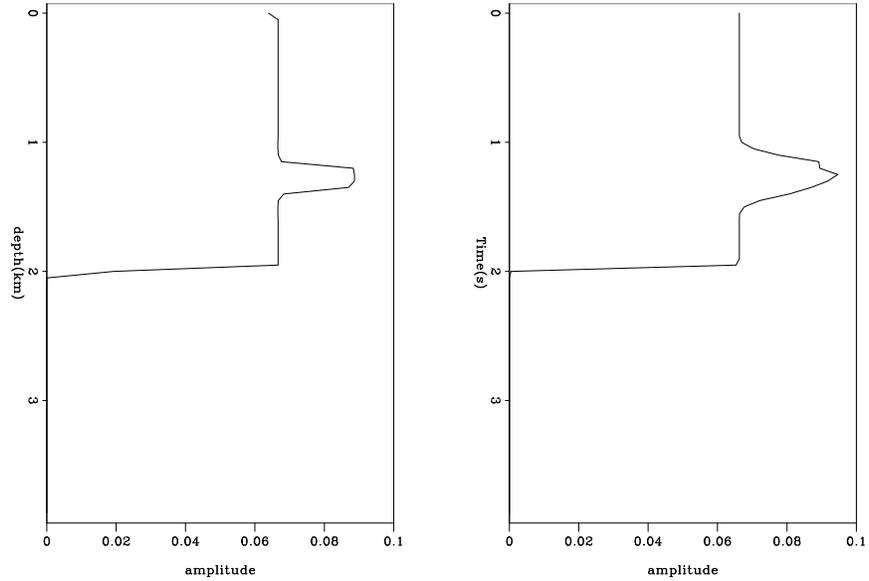


Figure 4: The operator calculated from our initial guess at velocity and the resulting ray paths in depth (left) and tau (right). `bob2-operator0` [ER]

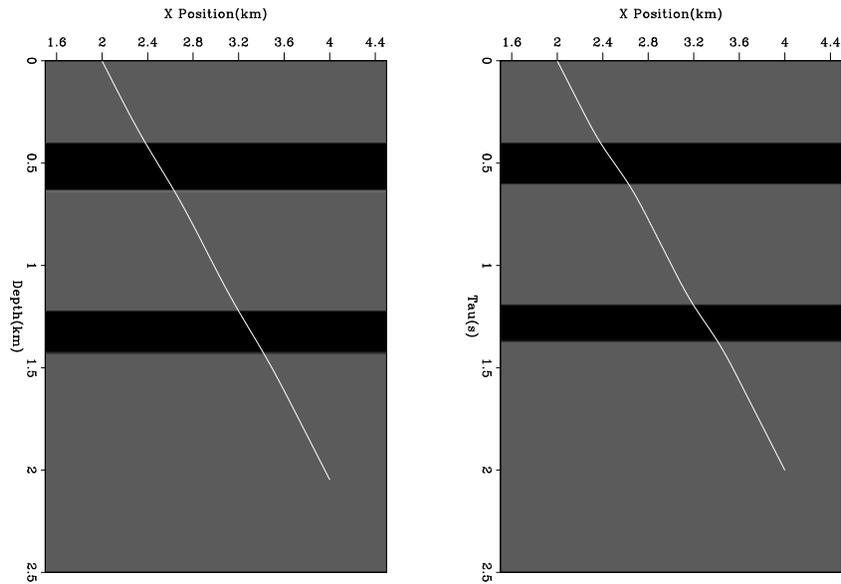


Figure 5: “Correct” velocity function overlaid by ray hitting reflector at 4 km with a half-offset of 2 km. Left Panel is in depth, right panel is in tau. `bob2-vel1` [ER]

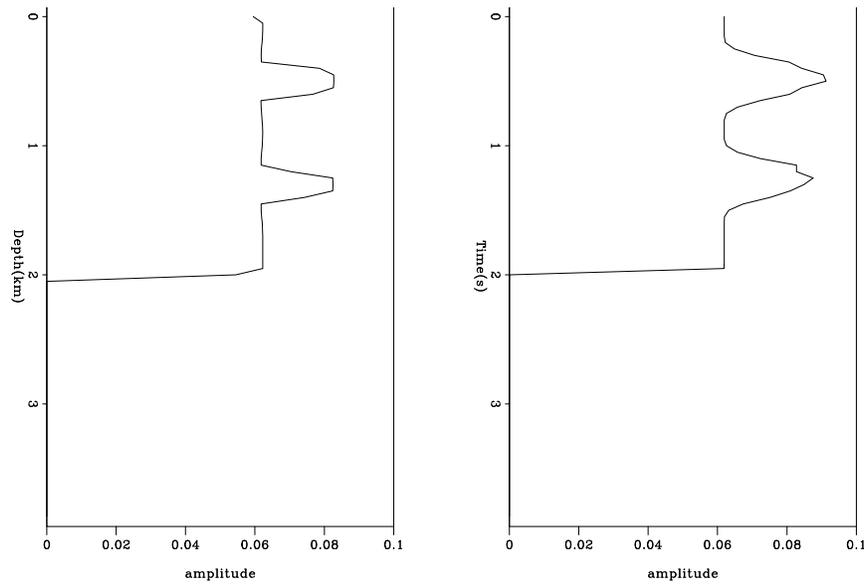


Figure 6: The operator calculated from the “correct” velocity and the resulting ray paths in depth (left) and tau (right). `bob2-operator1` [ER]

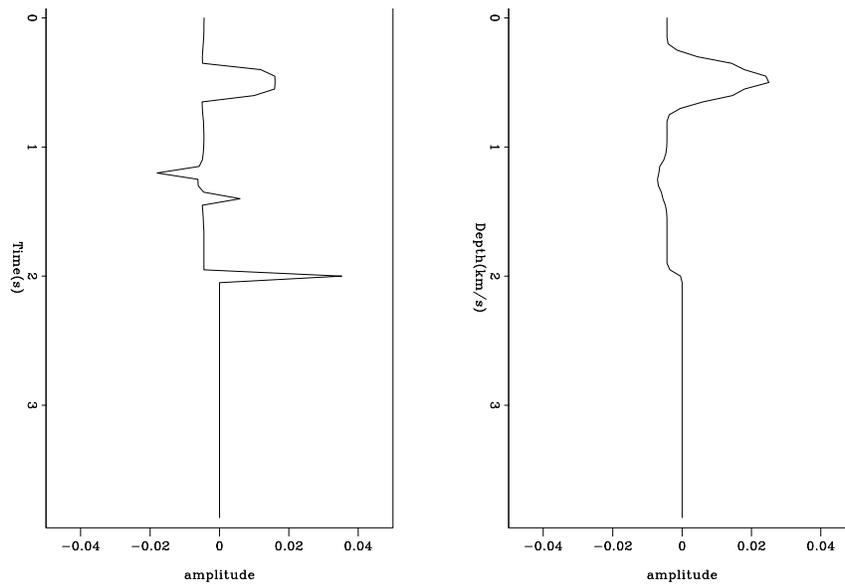


Figure 7: The difference between the operators calculated from the correct and our initial guess at velocity, for depth (left) and tau (right). Note the significant spikes at the reflector and at the lower layer boundary in the depth case. `bob2-diff` [ER]

Clapp, R. G., Sava, P., and Claerbout, J. F., 1998, Interval velocity estimation with a null-space: SEP-97, 147–156.

Etgen, J., 1990, Residual prestack migration and interval velocity estimation: SEP-68, 0.

Mora, P., 1987, Elastic wavefield inversion: Ph.D. thesis, Stanford University.

Stork, C., and Clayton, R. W., 1991, An implementation of tomographic velocity analysis: Geophysics, 56, no. 4, 483–495.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, 51, no. 10, 1893–1903.

Velocity continuation in migration velocity analysis

Sergey Fomel¹

ABSTRACT

Velocity continuation can be applied to migration velocity analysis. It enhances residual NMO correction by properly taking into account both vertical and lateral movements of reflectors caused by the change in migration velocity. I exemplify this fact with simple data tests.

INTRODUCTION

Migration velocity analysis is a routine part of prestack time migration applications. It serves both as a tool for velocity estimation (Deregowski, 1990) and as a tool for optimal stacking of migrated seismic sections and modeling zero-offset data for depth migration (Kim et al., 1997). In the most common form, migration velocity analysis amounts to residual moveout correction on CRP (common reflection point) gathers. However, in the case of dipping reflectors, this correction does not provide optimal focusing of reflection energy, since it does not account for lateral movement of reflectors caused by the change in migration velocity. In other words, different points on a stacking hyperbola in a CRP gather can correspond to different reflection points at the actual reflector. The situation is similar to that of the conventional NMO velocity analysis, where the reflection point dispersal problem is usually overcome with the help of DMO (Deregowski, 1986; Hale, 1991). An analogous correction is required for optimal focusing in the post-migration domain. In this paper, I propose and test velocity continuation as a method of migration velocity analysis. The method enhances the conventional residual moveout correction by taking into account lateral movements of migrated reflection events.

Velocity continuation is an artificial process of transforming time migrated images according to the changes in migration velocity. This process has wave-like properties, which have been described in my earlier papers (Fomel, 1994, 1996, 1997). Hubral et al. (1996) and Schleicher et al. (1997) use the term *image waves* to introduce a similar concept. Velocity continuation extends the theory of residual and cascaded migrations (Rothman et al., 1985; Larner and Beasley, 1987). In practice, the continuation process can be modeled by finite-difference or spectral methods (Fomel and Claerbout, 1997; Fomel, 1998).

Applying velocity continuation to migration velocity analysis involves the following steps:

¹email: sergey@sep.stanford.edu

1. prestack common-offset (and common-azimuth) migration - to generate the initial data for continuation,
2. velocity continuation with stacking across different offsets - to transform the offset data dimension into the velocity dimension,
3. picking the optimal velocity and slicing through the migrated data volume - to generate an optimally focused image.

In this paper, I demonstrate all three steps, using a simple two-dimensional dataset. For the implementation of velocity continuation, I chose the Fourier spectral method. The method has its limitations (Fomel, 1998), but looks optimal in terms of the accuracy versus efficiency trade-off. It is important to note that although the velocity continuation result could be achieved in principle by using prestack residual migration in Kirchhoff (Etgen, 1990) or Stolt (Stolt, 1996) formulation, the first is evidently inferior in efficiency, and the second is not convenient for velocity analysis across different offsets, because it mixes them in the Fourier domain (Sava, 1999).

PUTTING TOGETHER PRESTACK VELOCITY CONTINUATION

Velocity continuation in the zero-offset (post-stack case) can be performed with a simple Fourier-domain algorithm (Fomel, 1998):

1. Input an image, migrated with velocity v_0 .
2. Transform the time axis t to the squared time coordinate: $\sigma = t^2$.
3. Apply a fast Fourier transform (FFT) on both the squared time and the midpoint axis. The squared time σ transforms to the frequency Ω , and the midpoint coordinate x transforms to the wavenumber k . We can safely assume that in the post-migration domain seismic images are uniformly sampled in x , which allows us to use the FFT technique. In the case of 3-D data, FFT should be applied in both midpoint coordinates.
4. Apply a phase-shift operator to transform to different velocities v :

$$\hat{P}(\Omega, k, v) = \hat{P}_0(\Omega, k) e^{\frac{ik^2(v_0^2 - v^2)}{4\Omega}}. \quad (1)$$

5. Apply an inverse FFT to transform from Ω and k to σ and x .
6. Apply an inverse time stretch to transform from σ to t .

The computational complexity of this algorithm has the same order as that of the Stolt migration (Stolt, 1978), but in practice it can be even faster because of the very simple inner computation.

To generalize algorithm (1) to the prestack case, we first need to include the residual NMO term (Fomel, 1996). Residual normal moveout can be formulated with the help of the differential equation:

$$\frac{\partial P}{\partial v} + \frac{h^2}{v^3 t} \frac{\partial P}{\partial t} = 0, \quad (2)$$

where h stands for the half-offset. The analytical solution of equation (2) has the form of the residual NMO operator:

$$P(t, h, v) = P_0 \left(\sqrt{t^2 + h^2 \left(\frac{1}{v_0^2} - \frac{1}{v^2} \right)}, h \right). \quad (3)$$

After transforming to the squared time $\sigma = t^2$ and the corresponding Fourier frequency Ω , equation (2) takes the form of the ordinary differential equation

$$\frac{d\hat{P}}{dv} + i\Omega \frac{2h^2}{v^3} \hat{P} = 0 \quad (4)$$

with the analytical frequency-domain phase-shift solution

$$\hat{P}(\Omega, h, v) = \hat{P}_0(\Omega, h) e^{i\Omega h^2 \left(\frac{1}{v_0^2} - \frac{1}{v^2} \right)}. \quad (5)$$

To obtain a Fourier-domain prestack velocity continuation algorithm, we just need to combine the phase-shift operators in equations (1) and (5) and to include stacking across different offsets. The algorithm takes the following form:

1. Input a set of common-offset images, migrated with velocity v_0 .
2. Transform the time axis t to the squared time coordinate: $\sigma = t^2$.
3. Apply a fast Fourier transform (FFT) on both the squared time and the midpoint axis. The squared time σ transforms to the frequency Ω , and the midpoint coordinate x transforms to the wavenumber k .
4. Apply a phase-shift operator to transform to different velocities v :

$$\hat{P}(\Omega, k, v) = \sum_h \hat{P}_0(\Omega, k, h) e^{i \frac{k^2(v_0^2 - v^2)}{4\Omega}} e^{i\Omega h^2 \left(\frac{1}{v_0^2} - \frac{1}{v^2} \right)}. \quad (6)$$

To save memory, the continuation step is immediately followed by stacking.

5. Apply an inverse FFT to transform from Ω and k to σ and x .
6. Apply an inverse time stretch to transform from σ to t .

One can design similar algorithms by using finite differences or Chebyshev spectral methods (Fomel, 1998).

The complete theory of prestack velocity continuation also requires a residual DMO operator (Etgen, 1990; Fomel, 1996, 1997). However, the difficulty of implementing this operator is not fully compensated by its contribution to the full velocity continuation. For simplicity, I decided not to include residual DMO in the current implementation.

Figure 1 shows impulse responses of prestack velocity continuation. The input for producing this figure was a time-migrated constant-offset section, corresponding to an offset of 1 km and a constant migration velocity of 1 km/s. In full accordance with the theory (Fomel, 1996), three spikes in the input section transformed into shifted ellipsoids after continuation to a higher velocity and into shifted hyperbolas after continuation to a smaller velocity.

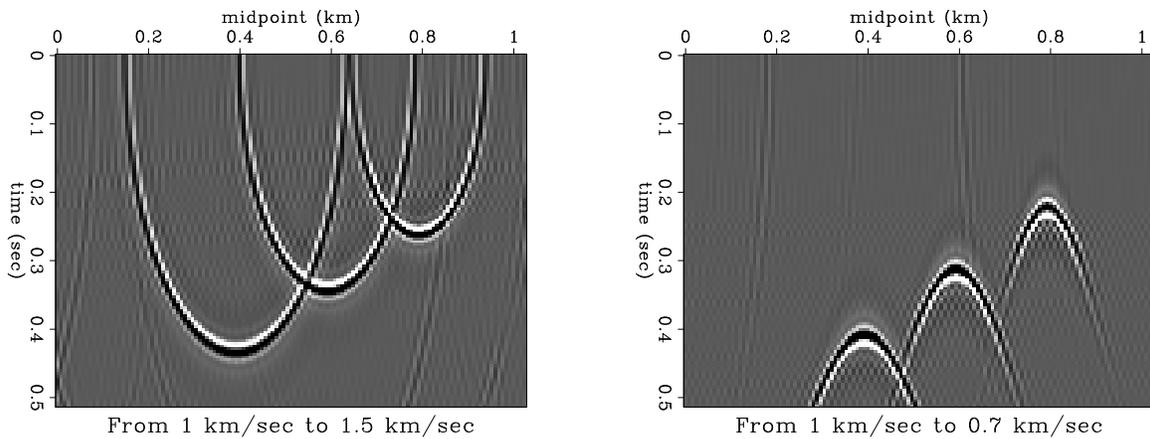


Figure 1: Impulse responses of prestack velocity continuation. Left plot: continuation from 1 km/s to 1.5 km/s. Right plot: continuation from 1 km/s to 0.7 km/s. Both plots correspond to the offset of 1 km. `sergey2-velimp` [ER]

Figure 2 compares the result of a constant-velocity prestack migration with the velocity of 1.8 km/s, applied to the infamous Gulf of Mexico dataset from *Basic Earth Imaging* (Claerbout, 1995) and the result of velocity continuation to the same velocity from a migration with a smaller velocity of 1.3 km/s. The differences in the top part of the images are explained by differences in muting. In the first case, muting was applied after migration, and in the second case, muting was applied prior to velocity continuation. The other parts of the sections look very similar, as expected from the theory.

Velocity continuation creates a time-midpoint-velocity cube (four-dimensional for 3-D data), which we can use for picking RMS velocities in the same way as we would use the result of common-midpoint or common-reflection-point velocity analysis. The important difference is that velocity continuation provides an optimal focusing of the reflection energy by properly taking into account both vertical and lateral movements of reflector images with changing migration velocity. Figure 3 compares velocity spectra (semblance panels) at a CRP location of about 11.5 km after residual NMO and after prestack velocity continuation. Although the overall difference between the two panels is small, the velocity continuation panel

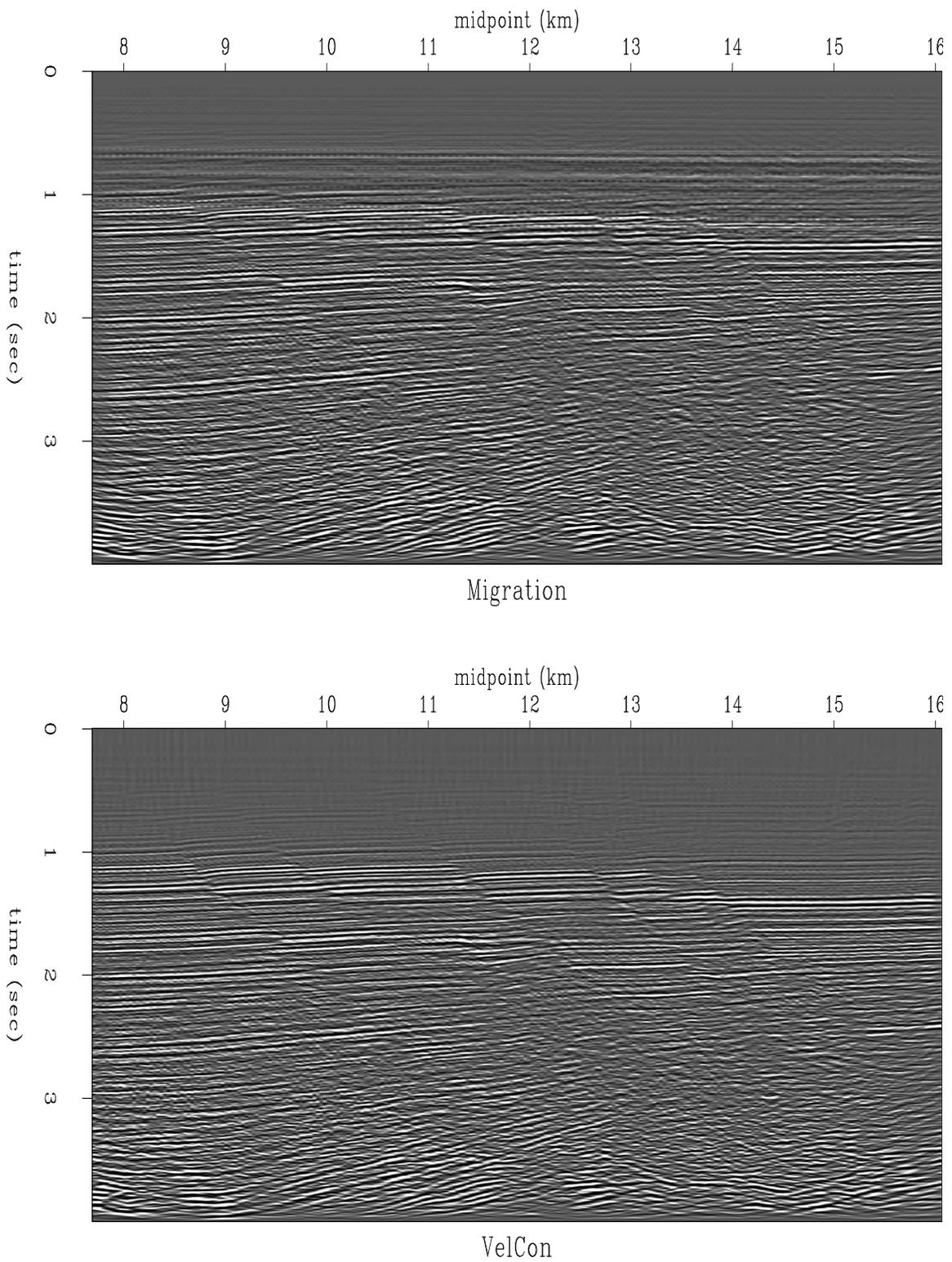


Figure 2: Top: The Gulf of Mexico dataset from *BEI* after prestack migration with the constant velocity of 1.8 km/s. Bottom: The same data after after velocity continuation from 1.3 km/s to 1.8 km/s. [sergey2-velmigr](#) [ER]

shows a noticeably better focusing, especially in the region of conflicting dips between 1 and 2 seconds. The next section discusses the velocity picking step in more details.

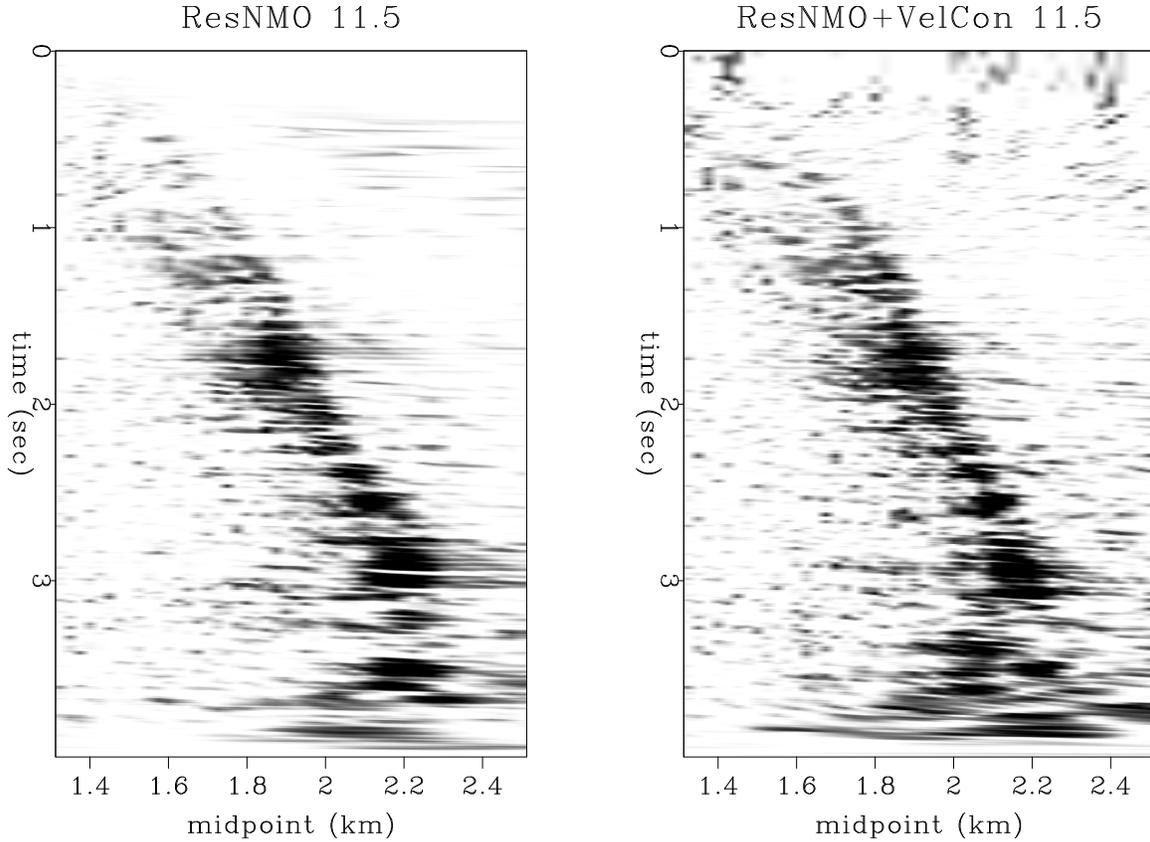


Figure 3: Velocity spectra around 11.5 km CRP after residual NMO (left) and after prestack velocity continuation (right). The right plot shows improved focusing in the region between 1 and 2 seconds. [sergey2-consmb](#) [CR]

VELOCITY PICKING AND SLICING

After the velocity continuation process has created a velocity cube in the prestack common-offset migration domain, we can pick the best focusing velocity from that cube. To automatize the velocity picking procedure, I have designed a simple algorithm. The algorithm based on solving the following regularized least-square system:

$$\begin{cases} \mathbf{W}\mathbf{x} \approx \mathbf{W}\mathbf{p} \\ \epsilon\mathbf{D}\mathbf{x} \approx \mathbf{0} \end{cases} \quad (7)$$

Here \mathbf{p} are blind maximum-semblance picks (possibly in a predefined fairway), \mathbf{x} is the estimated velocity picks, \mathbf{W} is the weighting operator with the weight corresponding to the semblance values at \mathbf{p} , \mathbf{D} is a roughening operator, and ϵ is the scalar regularization parameter. The first least-square fitting goal in (7) states that the estimated velocity picks should match

the measured picks where the semblance is high enough². The second fitting goal tries to find the smoothest velocity function possible. The least-square solution of problem (7) takes the form

$$\mathbf{x} = (\mathbf{W}^2 + \epsilon^2 \mathbf{D}^T \mathbf{D})^{-1} \mathbf{W}^2 \mathbf{p}, \quad (8)$$

where \mathbf{D}^T denotes the adjoint operator. In the case of picking a one-dimensional velocity function from a single semblance panel, we can simplify the algorithm by choosing \mathbf{D} to be the a convolution with the derivative filter $(1, -1)$. It is easy to notice that in this case the inverted matrix in formula (8) has a tridiagonal structure and therefore can be easily inverted with a linear-time algorithm. The regularization parameter ϵ controls the amount of smoothing of the estimated velocity function. Figure 4 shows a velocity spectrum and two automatic picks for different values of ϵ .

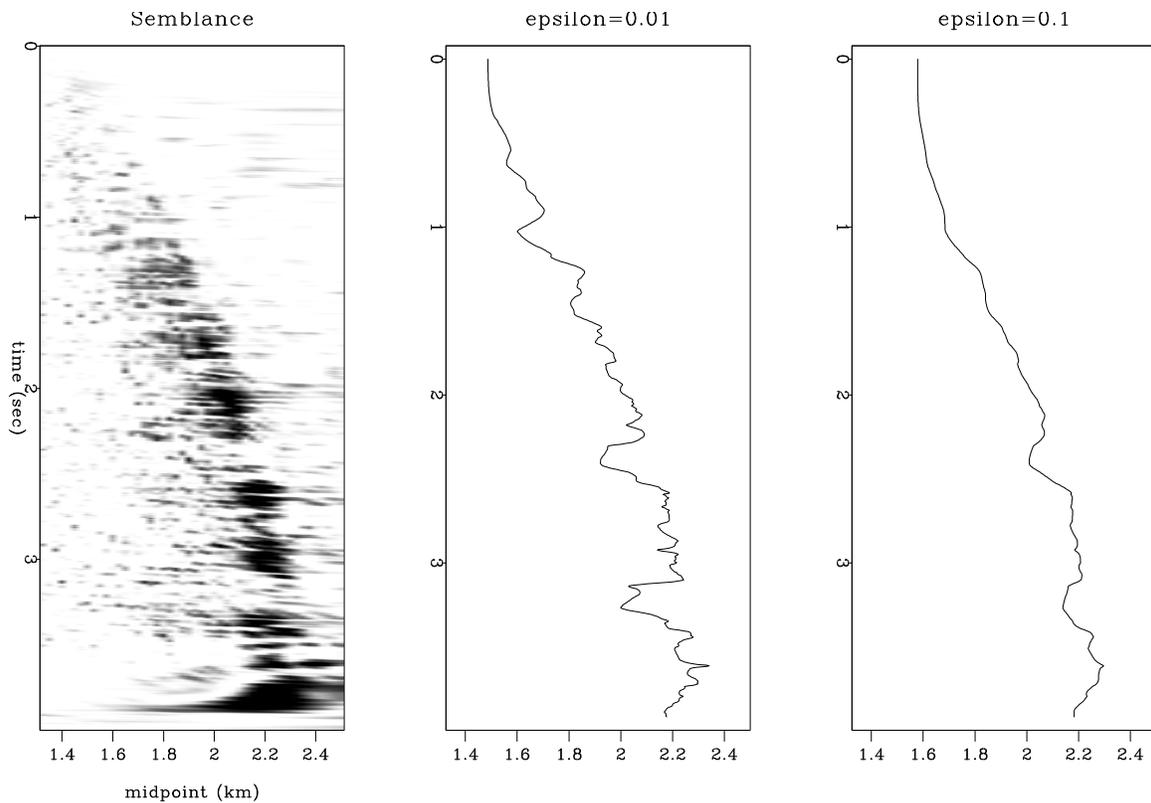


Figure 4: Semblance panel (left) and automatic velocity picks for different values of the regularization parameter. Center: $\epsilon = 0.01$, right: $\epsilon = 0.1$. Higher values of ϵ lead to smoother velocities. [sergey2-velpick](#) [ER]

In the case of picking two- or three-dimensional velocity functions, one could generalize problem (7) by defining \mathbf{D} as a 2-D or 3-D roughening operator. I chose to use a more

²Of course, this goal might be dangerous, if the original picks \mathbf{p} include regular noise (such as multiple reflections) with high semblance value (Toldi, 1985). For simplicity, and to preserve the linearity of the problem, we assume that this is not the case.

simplistic approach. I transform system (7) to the form

$$\begin{cases} \mathbf{W}\mathbf{x} \approx \mathbf{W}\mathbf{p} \\ \epsilon\mathbf{D}\mathbf{x} \approx \mathbf{0} \\ \lambda\mathbf{x} \approx \lambda\mathbf{x}_0 \end{cases}, \quad (9)$$

where \mathbf{x} is still one-dimensional, and \mathbf{x}_0 is the estimate from the previous midpoint location. The scalar parameter λ controls the amount of lateral continuity in the estimated velocity function. The least-square solution to system (9) takes the form:

$$\mathbf{x} = (\mathbf{W}^2 + \epsilon^2\mathbf{D}^T\mathbf{D} + \lambda^2\mathbf{I})^{-1} (\mathbf{W}^2\mathbf{p} + \lambda^2\mathbf{x}_0), \quad (10)$$

where \mathbf{I} denotes the identity matrix. Formula (10) also reduces to an efficient tridiagonal matrix inversion. Figure 5 shows a result of two-dimensional velocity picking after velocity continuation. I used values of $\epsilon = 0.1$ and $\lambda = 0.1$. The first parameter controls the vertical smoothing of velocities, while the second parameter controls the amount of lateral continuity. Figure 6 shows the final result of velocity continuation: an image, obtained by slicing through

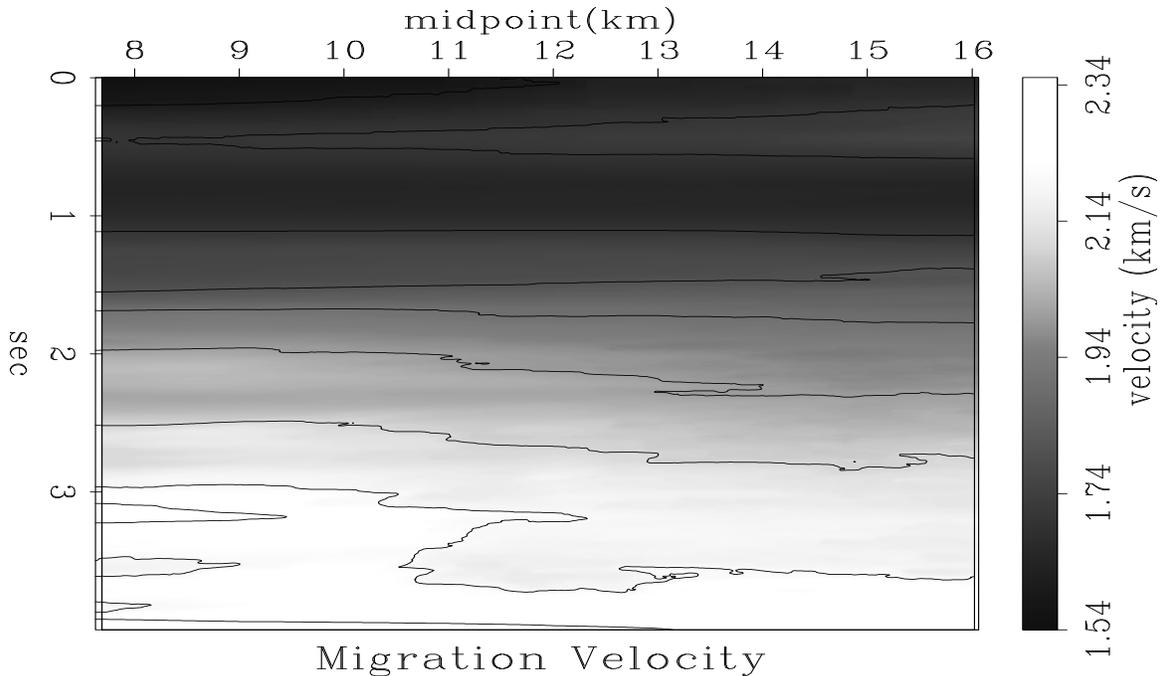


Figure 5: Automatic picks of 2-D RMS velocity after velocity continuation. The contour spacing is 0.1 km/s, starting from 1.5 km/s. [sergey2-beifpk](#) [CR]

the velocity cube with the picked RMS velocity. Different parts of the image have been properly positioned and focused by the velocity continuation process. One way to further improve the image quality is *hybrid migration*: demigration to zero-offset, followed by post-stack depth migration (Kim et al., 1997). This step requires constructing an interval velocity model from the picked RMS velocities. Without repeating the details of the procedure, Figures 7 and 8 show picked RMS velocities and the velocity continuation image for the Blake Outer Ridge data, shown in many other papers in this report.

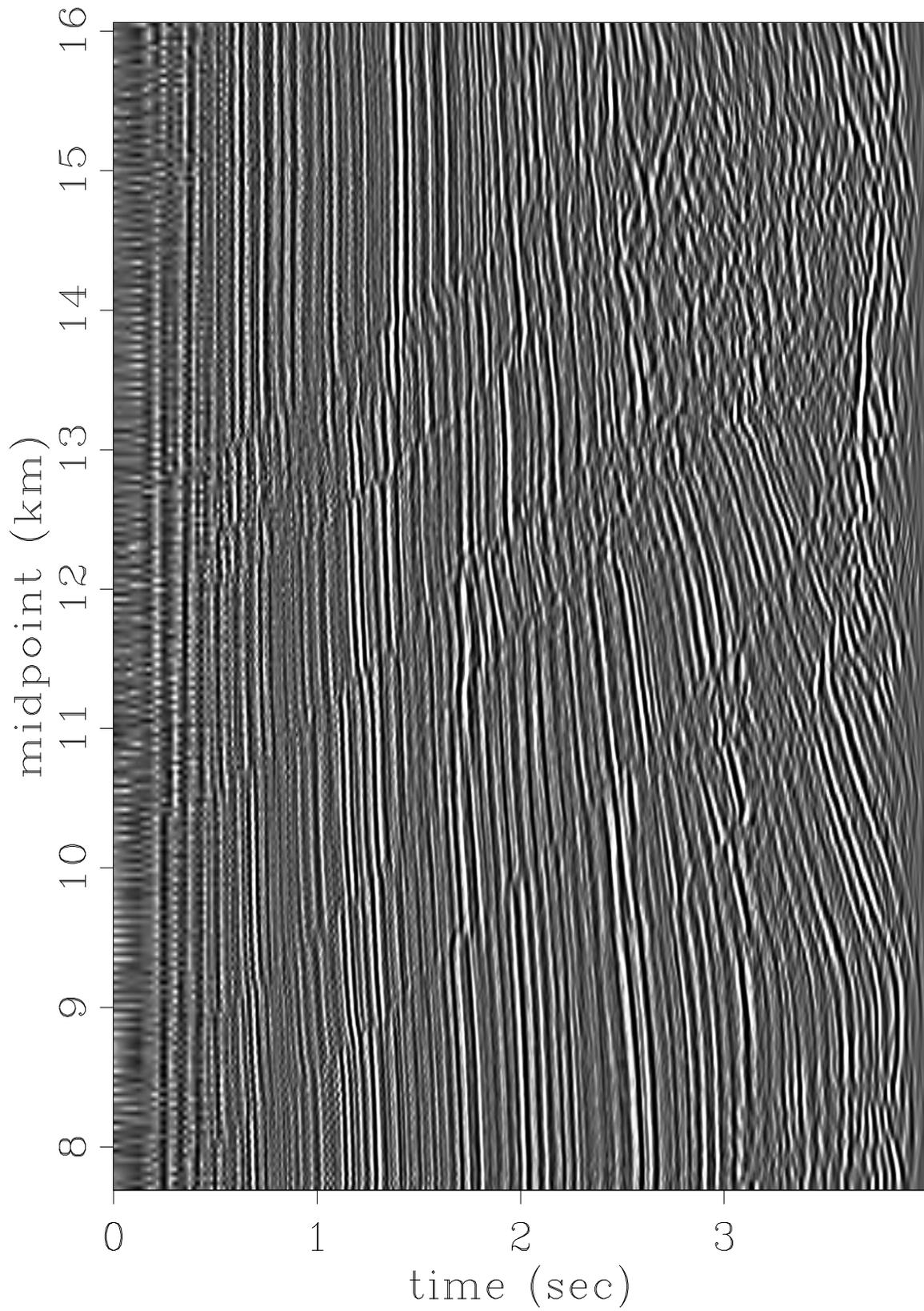


Figure 6: Final result of velocity continuation: seismic image, obtained by slicing through the velocity cube. [sergey2-beifmg](#) [CR]

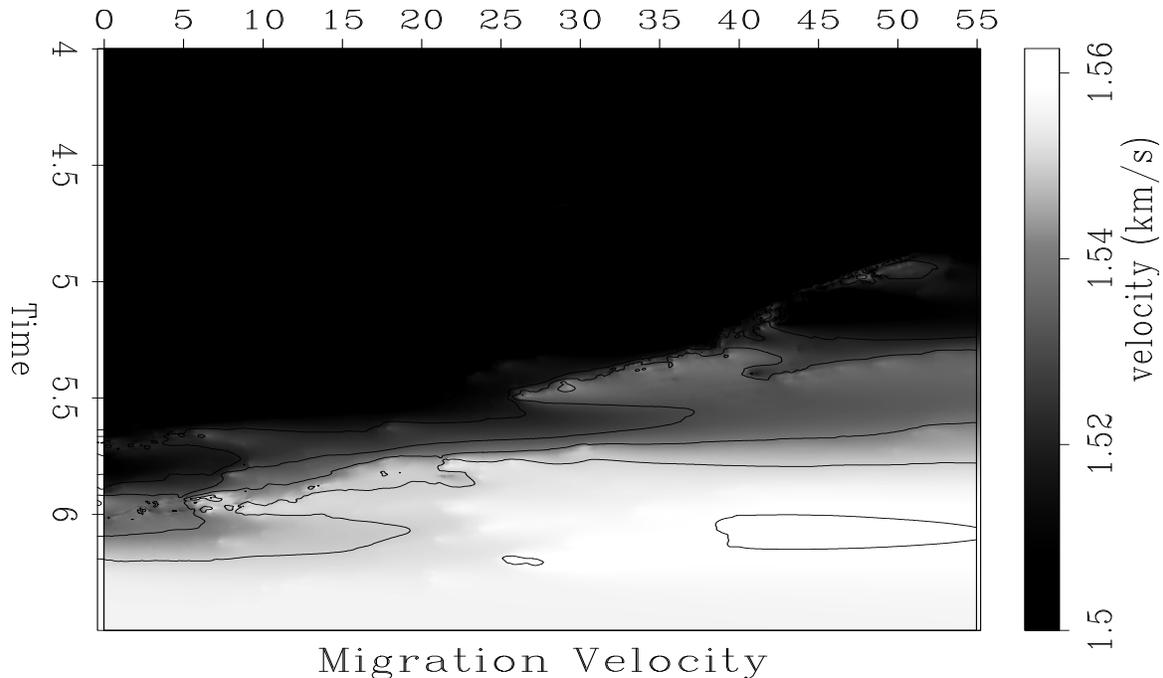


Figure 7: Blake Outer Ridge data. Automatic picks of 2-D RMS velocity after velocity continuation. The contour spacing is 0.01 km/s, starting from 1.5 km/s. `sergey2-pck` [CR]

CONCLUSIONS

I have demonstrated an application of velocity continuation to migration velocity analysis on simple data sets. The first results look promising and encourage further real data tests, hopefully with 3-D data.

ACKNOWLEDGMENTS

I thank Jon Claerbout, Biondo Biondi, and Bill Symes for useful and stimulating discussions.

REFERENCES

- Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.
- Deregowski, S. M., 1986, What is DMO: First Break, **4**, no. 7, 7–24.
- Deregowski, S. M., 1990, Common-offset migrations and velocity analysis: First Break, **8**, no. 6, 224–234.
- Etgen, J., 1990, Residual prestack migration and interval velocity estimation: Ph.D. thesis, Stanford University.

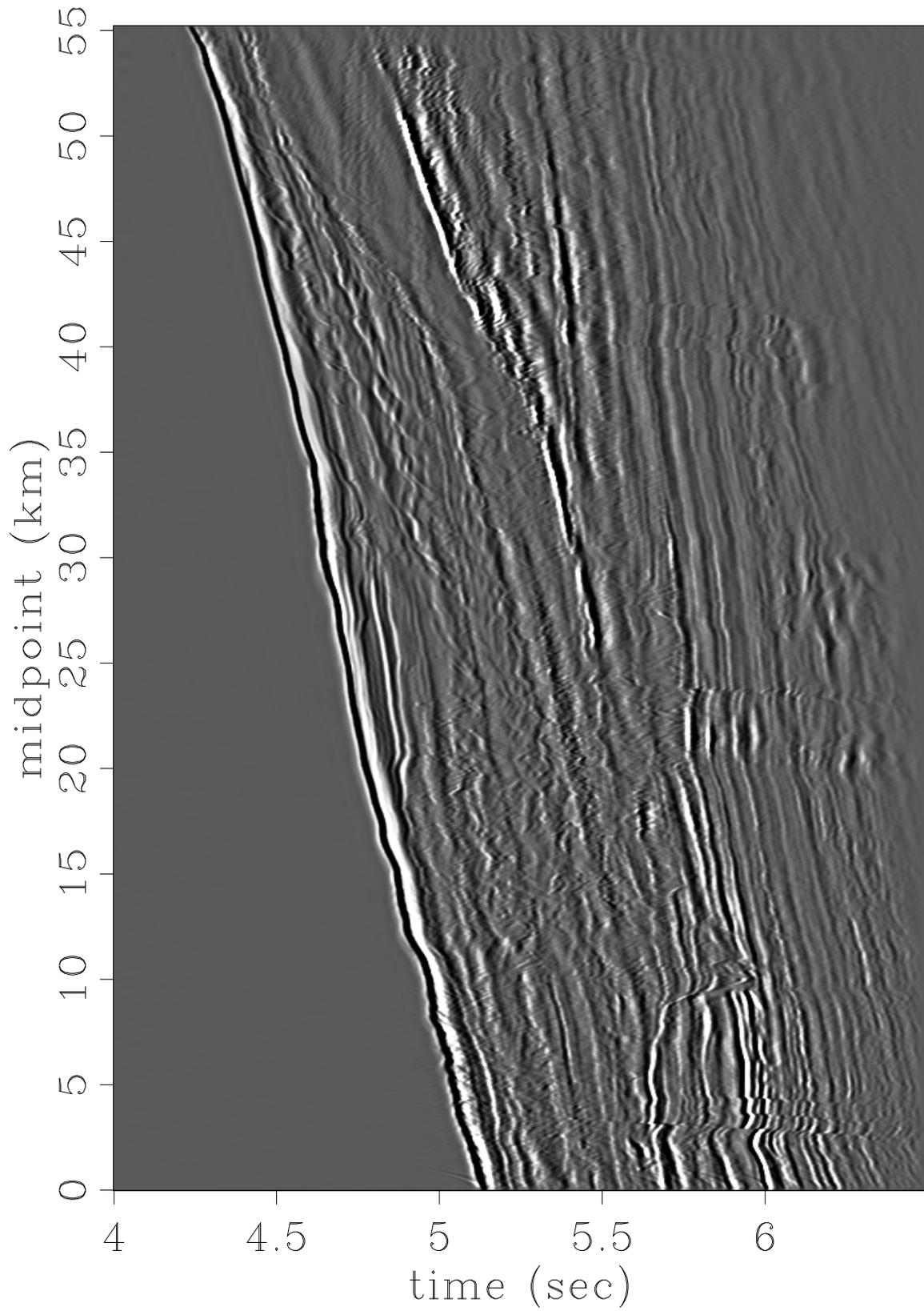


Figure 8: Blake Outer Ridge data. Final result of velocity continuation: seismic image, obtained by slicing through the velocity cube. [sergey2-img](#) [CR]

- Fomel, S., and Claerbout, J. F., 1997, Exploring three-dimensional implicit wavefield extrapolation with the helix transform: *SEP-95*, 43–60.
- Fomel, S. B., 1994, Method of velocity continuation in the problem of temporal seismic migration: *Russian Geology and Geophysics*, **35**, no. 5, 100–111.
- Fomel, S., 1996, Migration and velocity analysis by velocity continuation: *SEP-92*, 159–188.
- Fomel, S., 1997, Velocity continuation and the anatomy of residual prestack migration: 67th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1762–1765.
- Fomel, S., 1998, Velocity continuation by spectral methods: *SEP-97*, 157–172.
- Hale, D., 1991, Course notes: Dip moveout processing: Soc. Expl. Geophys.
- Hubral, P., Tygel, M., and Schleicher, J., 1996, Seismic image waves: *Geophysical Journal International*, **125**, 431–442.
- Kim, Y. C., Hurt, W. B., Maher, L. J., and Starich, P. J., 1997, Hybrid migration: A cost-effective 3-D depth-imaging technique: *Geophysics*, **62**, no. 02, 568–576.
- Larner, K., and Beasley, C., 1987, Cascaded migrations—Improving the accuracy of finite-difference migration: *Geophysics*, **52**, no. 5, 618–643.
- Rothman, D. H., Levin, S. A., and Rocca, F., 1985, Residual migration – applications and limitations: *Geophysics*, **50**, no. 1, 110–126.
- Sava, P., 1999, Short note—on Stolt prestack residual migration: *SEP-100*, 151–158.
- Schleicher, J., Hubral, P., Hocht, G., and Liptow, F., 1997, Seismic constant-velocity remigration: *Geophysics*, **62**, no. 02, 589–597.
- Stolt, R. H., 1978, Migration by Fourier transform: *Geophysics*, **43**, no. 1, 23–48.
- Stolt, R. H., 1996, Short note - A prestack residual time migration operator: *Geophysics*, **61**, no. 02, 605–607.
- Toldi, J., 1985, Velocity analysis without picking: Ph.D. thesis, Stanford University.

All stationary points of differential semblance are asymptotic global minimizers: Layered acoustics¹

William W. Symes²

ABSTRACT

Differential semblance velocity estimators have well-defined and smooth high frequency asymptotics. A version appropriate for analysis of CMP gathers and layered acoustic models has no secondary minima. Its structure suggests an approach to optimal parametrization of velocity models.

INTRODUCTION

The core problem of primaries-only (linearized, Born approximation) modeling, imaging, and inversion is that of finding an accurate reference velocity. Since the typical survey is highly redundant, predictions of reflectivity are redundant, and unlikely to be consistent (or *flat* in common image panels) unless the velocity field used to make them is essentially correct. This concept of *semblance of redundant images* underlies velocity estimation methods in widespread use (Taner and Koehler, 1969; Yilmaz and Chambers, 1984; Reshef, 1997). A number of researchers have cast velocity analysis as an *optimization* problem: that is, they propose an objective function to be minimized or maximized at a correct velocity model (Toldi, 1985; Al Yahya, 1989; Fowler, 1986; Kolb et al., 1986; Cao et al., 1990; Clément and Chavent, 1993; Sevink and Herman, 1993; Martinez and McMechan, 1991; Sen and Stoffa, 1991). Extremization of the objective is then an automatic process, to be accomplished through numerical optimization algorithms. The most widely investigated objectives - variants of stack power or RMS data fit error ("output least squares") - are velocity dependent quadratic forms in the data. These functions are believed to be multimodal and very ill-conditioned (Gauthier et al., 1986; Scales et al., 1991). The presumed existence of many local minima appears to mandate global search methods such as simulated annealing. These usually require orders of magnitude more function evaluations than do gradient-based methods which find local minima. The computational cost of global search methods renders them unsuitable for industrial scale velocity estimation.

The mechanism underlying these features of stack power and similar objectives is *asymptotic instability*: none of these functions have limiting shapes as source and data bandwidth become infinite. Besides accounting for multimodality, saturation, ill conditioning, and other un-

¹This report will also appear in the TRIP 1999 Annual Report

²email: symes@caam.rice.edu

desirable mathematical properties, asymptotic instability of stack power, output least squares, and similar objectives also inhibits analysis of local and global features *via* of high frequency asymptotics.

These observations lead to the question: do there exist velocity dependent quadratic forms in the data, extremized by the correct velocity with model-consistent data, which also have stable high frequency asymptotics? The answer is “yes”, and the nature of such forms is specified completely as part of the answer: they express semblance through comparison of *neighboring traces*. In contrast, stack power and output least squares objectives measure semblance (explicitly or implicitly) by comparing of traces with widely differing offset and/or midpoint, and this fact accounts for the asymptotic instability of these functions. In the ideal limit of continuous sampling, traces to be compared should be infinitely near, so I have used the phrase *differential semblance* to describe these asymptotically stable forms. For mathematical details of the connection between differential semblance and stable asymptotics in the context of the simpler but similar *plane wave detection* problem, see (Kim and Symes, 1998), also (Claerbout, 1992), pp. 93 ff.

The asymptotic stability of differential semblance opens up the possibility of analysing its global shape by asymptotic methods. This paper presents such an analysis for a simple special case applicable to field data, based on the convolutional model of CMPs for layered acoustic Earth response. The layered medium assumption leads to simple explicit expressions for all quantities figuring in this approach to velocity estimation. The analysis shows that, for noise-free (model consistent) data in the continuous sampling limit and velocities limited to natural admissible sets, the length of the gradient bounds the objective, up to an error which vanishes in the high frequency limit. Therefore every stationary point is asymptotically a global minimum of the differential semblance objective function in this case. That is, differential semblance does not suffer from the local minima which plague other optimization formulations of velocity inversion.

Simple estimates bound the effect of noise. Numerical experiments have shown that random noise has virtually no effect on the location of DS stationary points, whereas strong coherent noise, such as multiple reflections, has a maximal effect. In any case the influence of noise is bounded, i.e. the differential semblance velocity estimate “degrades gracefully” as noise of any sort is added to the data.

These conclusions - stable asymptotics, unimodality, bounded influence of noise, significance of coherent noise - conform to the results of many numerical experiments with field and synthetic data. The present paper concerns only analysis: details of computational implementation and results appear elsewhere (Araya et al., 1996; Gockenbach and Symes, 1997; Symes, 1997, 1998; Chauris et al., 1998b,a).

The ubiquitous presence of various “multis” - multiple reflection, multiple refraction (transmission caustics), multiple wave modes, and of course multidimensional geometry - necessarily limits the practical importance of this or any other technique for velocity estimation (or imaging) based on primaries only layered acoustic modeling. Note that the differential semblance concept is *not* in any way limited to layered medium models or 1D velocity functions, any more than stack power is limited to NMO-based velocity analysis. The abstract (Chauris

et al., 1998a) and earlier work of this author (Symes, 1993; Kern and Symes, 1994) present examples of multidimensional velocity estimation by differential semblance optimization.

I begin with an abstract definition of differential semblance. After defining the convolutional model for layered acoustics, I discuss various types of error inherent in this approximation, the construction of mutes, and natural admissible model sets. This groundwork supports an asymptotic analysis of the differential semblance objective, which reveals that in the case of noise free data it is essentially a data-weighted mean square error in RMS square slowness. This observation leads directly to the main result, and to some convenient estimates of the influence of noise in general data. It also suggests an approach to optimal parametrization of velocity profiles.

AN ABSTRACT FORMULATION OF DIFFERENTIAL SEMBLANCE

The general definition of differential semblance presented here owes much to ideas introduced by Hua Song and Mark Gockenbach in their theses (Song, 1994; Gockenbach et al., 1995).

The (reference or background) velocity v includes the slowly varying components of velocity and perhaps other fields. The reflectivity r is a field (or vector of fields) encompassing the rapidly varying components of the model. Linearized scattering treats r as a perturbation of v . Thus in the forward modeling operator $F[v]r$ the dependence on r is linear, whereas the dependence on v is (quite!) nonlinear.

Minimal data sets are those on which the kinematic relations in the data are bijective. Minimal data sets include common shot and common offset gathers, and - for layered models - single traces. For a few models, such as constant density acoustics, the forward modeling operator $F[v]$ is invertible (modulo smoothing operators) on minimal data sets. This paper deals only with constant density acoustics.

Denote by $G[v]$ an approximate inverse operator for $F[v]$ on each (minimal) data bin (common source, common offset, single trace,...). Thus $G[v]$ applied to the data produces a *prestack reflectivity volume*. Similarly, understand by $F[v]$ the application of forward modeling independently for each reflectivity bin.

Each binning scheme also implies a notion of *neighboring bins*: that is, neighboring source positions, offsets,... Denote by W an operator approximating the derivative or gradient in the bin parameter(s). Generally the definition of $F[v]$ necessarily incorporates a cutoff or mute, as does that of $G[v]$. Differentiation in the bin direction across this mute produces edge artifacts. To control these, introduce an additional mute ϕ slightly more severe than the mutes built into F and G . Since the edge effects are localized, application of this secondary mute ϕ eliminates them.

Differentiation enhances high frequency content. To keep the spectrum of the differential semblance output comparable to that of the data, employ a smoothing operator H . An appropriate choice is the inverse square root of the Helmholtz operator $(I - \nabla^2)^{-\frac{1}{2}}$ in all of the variables on which the data depends, i.e. both within-bin and cross-bin variables.

With these notations, define differential semblance $J_0[v]$ by:

$$J_0[v] = \frac{1}{2} \|H\phi F[v]WG[v]S\|^2$$

Here S denotes the data, and the vertical double bars denote the L^2 norm or root mean square, i.e. summation of the square of the quantity inside over all variables, followed by square root.

THE CONVOLUTIONAL MODEL FOR LATERALLY HOMOGENOUS ACOUSTICS

Linearization of the acoustic model for a layered fluid and application of high frequency asymptotics leads to the *convolutional model* of primaries-only reflection seismograms. The convolutional model of offset traces is one of the simplest models of the reflection process within which to pose the velocity analysis problem. A similar model for plane wave traces is almost equally simple, and was the subject of earlier work on differential semblance (Symes and Carazzone, 1991; Minkoff and Symes, 1997). However synthesis of accurate plane wave traces is a nontrivial task. Accordingly the version of the model developed here uses offset domain data.

A natural binning scheme for this model is the common midpoint gather. Since all midpoint gathers are in principle the same for a layered model, the data consists of a single CMP. The bins contain single traces, parameterized by offset x .

The velocity parameter is simply the interval velocity $v(z)$, whereas the reflectivity is $r = \frac{\delta v}{v}$ and is regarded as bin-dependent, i.e. $r = r(z, x)$; this section plays the role of a common image gather, as every trace represents reflectivity below the same midpoint. Thus successful velocity estimation will produce a “flat” (x -independent) $r(z, x)$.

The simple version of DS presented here will assume that source signature deconvolution has been applied to the data, so that it is essentially impulsive.

It will be convenient to parametrize velocity and reflectivity by *vertical two-way time*

$$t_0 = 2 \int_0^z \frac{dz}{v}$$

rather than depth: thus $v = v(t_0), r = r(t_0, x)$.

With these conventions, the forward modeling operator is

$$F[v]r(t, x) = a(t, x)r(T_0(t, x), x)$$

where a is the geometric amplitude and $T_0(t, x)$ is the inverse function of the *two-way travel-time function* $T(t_0, x)$.

ERROR, ERROR EVERYWHERE!

As an approximate predictor of seismic traces, the convolutional model exhibits several types of error:

- physics error: seismic waves are not small amplitude pressure waves in a fluid;
- linearization error: neglect of multiple reflections and other nonlinear effects;
- deconvolution error: complete removal of the source signature is not possible;
- asymptotic error: the convolutional model becomes more accurate as the frequency content of $r(t_0, x)$ moves away from zero Hz.

The practical meaning of asymptotic error is that the convolutional model predicts the higher frequency components of the data more accurately, so that the prediction error can be reduced by more aggressive low-cut filtering. Of course this discarding of low-frequency data is only possible to a limited extent as actual data is bandlimited.

The following computations will introduce yet more sources of asymptotic error - and, with one exception, only asymptotic error. Therefore I will identify asymptotic error explicitly, and treat other types of modeling error as data noise. It is possible to estimate every asymptotic error explicitly, but experience suggests that these explicit estimates are not particularly useful. So instead I will use the symbol “ $O(\lambda)$ ” to suggest proportionality of the asymptotic error to a dominant wavelength in the data. Thus

$$F[v]r(t, x) = a(t, x)r(T_0(t, x), x) + O(\lambda)$$

The single important lesson to learn from the explicit error estimates of geometric optics is that they are uniform over C^∞ -bounded sets of coefficients (meaning in this case the velocity v). Therefore the velocities appearing in the sequel are restricted to vary over such a C^∞ -bounded set. A byproduct of the analysis will suggest explicit finite dimensional subspaces of smooth functions in which it is advantageous to seek v .

MUTES

The linearized model accurately predicts only precritical primary reflections. For layered media, precritical reflections have downgoing incident rays. Along downgoing rays, time is an increasing function of depth. It follows that if t_0 is to be a depth variable, then T must be an increasing function of t_0 . This is generally true only in a subset of the t, x plane, i.e. only part of this plane contains data accurately modeled by linearized acoustics. Therefore the rest of the data must be muted out.

Define the *stretch factor*

$$s(t, x) = \frac{\partial T_0}{\partial t}(t, x) = \left(\frac{\partial T}{\partial t_0}(T_0(t, x), x) \right)^{-1}$$

Then the condition that $T(t_0, x)$ be monotone increasing as a function of t_0 is equivalent to demanding that for large enough t

$$0 < s(t, x) < C_{\text{stretch}}$$

where C_{stretch} is a user-specified parameter larger than one. Define $T_{\text{mute}}(x)$ (the *mute boundary*) to be the infimum of all t for which the above inequality is satisfied on the interval (t, T_{max}) . Then the support of the mute function ϕ should be contained in the set $\{(t, x) : t \geq T_{\text{mute}}(x)\}$.

Define a corresponding t_0, x domain mute by $\phi_0(t_0, x) = \phi(T(t_0, x), x)$.

ADMISSIBLE MODELS

In this section I introduce *admissible sets* \mathcal{A} of models, on which the convolutional model as defined above is reasonably well behaved. Note that the constraints imposed on the models by membership in the admissible sets are very natural from the physical or geological point of view.

First of all, the velocity must be smooth, as noted above in the section on errors. The restriction of v to a bounded subset of C^∞ implies bounds (maximum absolute value, mean square,...) on any derivative of v .

Second, impose smooth upper and lower “envelope” velocities as hard constraints: $v_{\min}(t_0) \leq v(t_0) \leq v_{\max}(t_0)$. It is natural to assume that the velocity is known at the surface, so assume that $v_{\min}(0) = v_{\max}(0)$. These bounds derive from geophysical measurements and general knowledge about rock physics, so should be regarded as distinct from the bounds implied by the first condition (membership in a bounded set in C^∞).

The set of velocities satisfying the constraints just outlined form the *admissible set* \mathcal{A} .

An important consequence is that the mute $\phi \in C_0^\infty(\mathbf{R}^2)$ may be chosen uniform over \mathcal{A} , as uniform bounds then exist for every value of the stretch factor $s(t, x)$. These bounds follow from the equations of geometric optics. However they are even more simply derived for the hyperbolic moveout approximation to traveltime, which I will eventually adopt, so I do not give a derivation here.

ASYMPTOTIC APPROXIMATION OF DIFFERENTIAL SEMBLANCE

The convolutional offset trace model is one of those for which the forward modeling operator on a minimal gather, ie. a single trace, is invertible. The inverse operator is

$$G[v]S(t_0, x) = \frac{S(T(t_0, x), x)}{a(T(t_0, x), x)}$$

The operator measuring semblance differentially is

$$W = \frac{\partial}{\partial x}$$

Then

$$\begin{aligned} F[v]WG[v]S(t,x) &= a(t,x) \left[\frac{\partial}{\partial x} \frac{S(T(t_0,x),x)}{a(T(t_0,x),x)} \right]_{t_0=T_0(t,x)} \\ &= \left(\frac{\partial T}{\partial x}(T_0(t,x),x) \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) S(t,x) + \dots \\ &= \left(p(t,x) \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) S(t,x) + \dots \end{aligned}$$

where

$$p(t,x) = \frac{\partial T}{\partial x}(T_0(t,x),x)$$

is the arrival (horizontal) slowness of the ray passing offset x at time t , and the elided terms involve the amplitude a , but do *not* involve derivatives of the data S . Thus these terms are of lower frequency content than the leading term (explicitly displayed), and are of the same relative order in frequency as terms neglected in the derivation of the convolutional model from the acoustic wave equation. Therefore they can be dropped: this leads to the remarkable conclusion that the differential semblance objective is *independent of the amplitude* at least to leading order in frequency.

This observation is due to Hua Song. As a result, within accuracy limitations already built into the asymptotic linearized model, a might as well be replaced by 1!. That is, to leading order in frequency, differential semblance is insensitive to wave dynamics (amplitude), and responds only to kinematic model changes, i.e. changes in traveltime. Thus minimization of differential semblance will amount to a sort of traveltime tomography.

Fons ten Kroode (personal communication) has pointed out that replacement of $G[v]$ by an asymptotically unitary operator with the same kinematics also yields an asymptotically identical objective without leading order amplitude dependence, and without application of the forward modeling operator, thus at lower computational cost.

The computations above are correct when the map $(t_0, x) \mapsto (t, x)$ is smooth and invertible. This is so inside the mute zone defined above, uniformly for $v \in \mathcal{A}$. Therefore application of the inverse square root Helmholtz operator following will bring the spectral content back into alignment with that of the data, uniformly over $v \in \mathcal{A}$. Thus

$$H\phi F[v]WG[v]S = H\phi \left(\frac{\partial S}{\partial x} + p \frac{\partial S}{\partial t} \right) + O(\lambda)$$

The ray slowness p is locally a smooth function of the velocity v in any fixed open subset of the mute zone, hence J_0 (which is the mean square of the above expression) is a smooth function of $v \in \mathcal{A}$ as well.

NOISE FREE DATA

Assume that the data S are model-consistent, that is

$$S(t, x) = r^*(T_0^*(t, x)) + O(\lambda)$$

for target offset independent reflectivity $r^*(t_0)$ and velocity $v^*(t_0)$. [Since differential semblance does not depend to leading order on the amplitude, as noted above, I set the amplitude to 1 in the following, for simplicity - it can be reintroduced with almost no change in the results to follow.]

Note that

$$0 = \frac{\partial}{\partial x} T(T_0(t, x), x) = \frac{\partial T}{\partial t_0}(T_0(t, x), x) \frac{\partial T_0}{\partial x}(t, x) + \frac{\partial T}{\partial x}(T_0(t, x), x)$$

so

$$\frac{\partial T_0}{\partial x}(t, x) = -s(t, x)p(t, x)$$

(s being the stretch factor, defined above). Thus

$$\frac{\partial}{\partial x} r^*(T_0^*(t, x)) = -s^*(t, x)p^*(t, x) \left(\frac{\partial r^*}{\partial t_0} \right) (T_0^*(t, x))$$

(s^* is the stretch factor belonging to v^*) whence

$$\begin{aligned} F[v]WG[v]S(t, x) &= \left(\frac{\partial}{\partial x} + p(t, x) \frac{\partial}{\partial t} \right) r^*(T_0^*(t, x)) \\ &= s^*(t, x)(p(t, x) - p^*(t, x)) \frac{\partial r^*}{\partial t_0}(T_0^*(t, x)) \end{aligned}$$

According to the calculus of pseudodifferential operators,

$$\begin{aligned} H\phi F[v]WG[v]S &= \\ &= (I - \nabla^2)^{-\frac{1}{2}} \phi \left(s^*(p - p^*) \frac{\partial r^*}{\partial t_0}(T_0^*) \right) \\ &= (I - \nabla^2)^{-\frac{1}{2}} \phi \left(s^*(p - p^*) \frac{\nabla T_0^* \cdot \nabla}{\nabla T_0^* \cdot \nabla T_0^*} r^*(T_0^*) \right) \\ &= \phi \frac{s^*(p - p^*)}{\sqrt{1 + s^{*,2}(1 + p^{*,2})}} r^*(T_0^*) + O(\lambda) \end{aligned}$$

where you get from the next to the last line to the last by substituting ∇T_0^* for ∇ , and using previously derived formulas for the partial derivatives of T_0 .

Thus

$$J_0[v] = \iint dt dx B^*(t, x)(p(t, x) - p^*(t, x))^2 [r^*(T_0^*(t, x))]^2 + O(\lambda)$$

where

$$B^*(t, x) = \phi^2 \frac{s^{*,2}}{1 + s^{*,2}(1 + p^{*,2})}$$

is independent of v , i.e. depending only on v^* and \mathcal{A} .

In the next section I introduce the so called hyperbolic moveout approximation to travel-time. Note that up to this point the development is entirely independent of this approximation. In particular the formulas worked out in this section have precise analogues for versions of differential semblance based on multidimensional seismic models.

HYPERBOLIC MOVEOUT

Claim: To good approximation, for “small” offsets,

$$T(t_0, x) = \sqrt{t_0^2 + \frac{x^2}{v_{\text{RMS}}^2(t_0)}}, \quad T_0(T(t_0, x), x) \equiv t_0$$

where the RMS velocity is

$$v_{\text{RMS}}(t_0) = \sqrt{\frac{1}{t_0} \int_0^{t_0} v^2}$$

Justification [Continuum derivation of the hyperbolic moveout approximation]: The 2-way traveltime $T(t_0, x) = \tau_2(z, x)$ from the surface at $z = 0$ to depth z and back at offset x is related to the solution of the eikonal equation $\tau(z, x)$ with point source at $z = x = 0$ by

$$\tau_2(z, x) = 2\tau\left(z, \frac{x}{2}\right)$$

Thus

$$\frac{1}{4} \left(\frac{\partial \tau_2}{\partial z} \right)^2 + \left(\frac{\partial \tau_2}{\partial x} \right)^2 = \frac{1}{v^2}$$

Differentiate this twice with respect to x and use the vanishing of odd-order x derivatives at $x = 0$ (implied by symmetry) to conclude that the second x derivative

$$q(z) = \frac{\partial^2 \tau}{\partial x^2}(z, 0)$$

satisfies

$$\frac{1}{2v} \frac{dq}{dz} + q^2 = 0$$

Introduce temporarily a new depth coordinate

$$\sigma(z) = 2 \int_0^z v$$

Then in terms of σ , q satisfies the Ricatti equation

$$\frac{dq}{d\sigma} + q^2 = 0$$

The solution which is singular at $\sigma = 0$, i.e. $z = 0$, is

$$q(\sigma) = \frac{1}{\sigma} = \frac{1}{2 \int_0^z v}$$

Since $dz = \frac{1}{2} v dt_0$, you can also write this as

$$q(\sigma) = \frac{1}{\int_0^{t_0} v^2}$$

Thus

$$\begin{aligned} T(t_0, x) &= t_0 + \frac{x^2}{2} \frac{\partial^2 T(t_0, 0)}{\partial x^2} + \dots \\ &= t_0 + \frac{x^2}{2 \int_0^{t_0} v^2} + \dots \end{aligned}$$

Since

$$\frac{\partial^2}{\partial x^2} (T(t_0, x))^2_{x=0} = 2t_0 \frac{\partial^2 T}{\partial x^2}(t_0, 0)$$

the above can be rewritten

$$\begin{aligned} T(t_0, x)^2 &= t_0^2 + \frac{x^2}{\frac{1}{t_0} \int_0^{t_0} v^2} + \dots \\ &= t_0^2 + \frac{x^2}{v_{\text{RMS}}^2(t_0)} + \dots \end{aligned}$$

which reveals that the hyperbolic moveout approximation is just the second order Taylor expansion of T^2 in x , which should be good for “small” x .

This report adopts the hyperbolic moveout approximation, i.e. truncate the Taylor expansion above and take

$$T(t_0, x)^2 = t_0^2 + \frac{x^2}{v_{\text{RMS}}^2(t_0)}$$

This amounts to assuming that all events in the data have precisely hyperbolic moveout. Of course this assumption is not entirely consistent with geometric optics. It has been suggested that the deviation of actual two-way time from the hyperbolic moveout approximation may be mistaken for evidence of anisotropy in some cases. In any case the error caused by replacing actual two way time by its hyperbolic moveout approximation is *not* an asymptotic error in the sense of the last section, so I will treat it as a component of data noise.

The reciprocal square RMS velocity, or *RMS square slowness* is the primary expression of velocity in the above formula. It occurs so often as to warrant its own notation:

$$u(t_0) \equiv (v_{\text{RMS}}(t_0))^{-2}$$

The conditions defining the mute can be restated: since

$$\frac{\partial T}{\partial t_0}(t_0, x) = \frac{t_0 + \frac{x^2}{2} \frac{\partial u}{\partial t_0}(t_0)}{T(t_0, x)}$$

the quantity on the right hand side of this equation must be bounded away from zero. Since v generally increases with depth, hence u decreases, such a lower bound will only be possible for t_0 exceeding a threshold for each x , which is the mute boundary mentioned before. In the data, i.e. (t, x) , coordinates, the stretch factor condition becomes

$$s(t, x) = \frac{\partial T_0}{\partial t}(t_0, x) = \left(\frac{\partial T}{\partial t_0}(T_0(t, x), x) \right)^{-1} = \frac{t}{T_0(t, x) + \frac{x^2}{2} \frac{\partial u}{\partial t_0}(T_0(t, x))} < C_{\text{stretch}}$$

and as before the mute ϕ must be supported in the set specified by this condition.

The upper and lower velocity envelopes implied by membership of the velocity in \mathcal{A} imply corresponding envelope mean square slownesses (u_{\min} corresponding to v_{\max} and vis-versa) so that $u_{\min}(t_0) \leq u(t_0) \leq u_{\max}(t_0)$.

It is usually reasonable to assume the lower velocity bound to be constant (independent of t_0) - for example, equal to sound velocity in water, or close to it. Then u_{\max} is also constant, so you can explicitly estimate a lower bound for T_0 :

$$T^2(t_0, x) = t_0^2 + x^2 u(t_0) \leq t_0^2 + x^2 u_{\max}$$

so

$$T_0(t, x) \geq \sqrt{t^2 - x^2 u_{\max}}$$

The velocity bounds also imply a bound on the derivative of u :

$$\frac{du}{dt_0} = -\frac{u^2}{t_0} \left(v^2 - \frac{1}{u} \right) = -\frac{u^2}{t_0} \left(v^2 - v^2(0) - \frac{1}{t_0} \int_0^{t_0} (v^2 - v^2(0)) \right)$$

The bounds on v , the known value of v at the surface, and the maximum two way time imply bounds on the slope

$$\frac{v^2(t_0) - v^2(0)}{t_0}$$

whence a bound u'_{\max} on the derivative of u follows immediately.

Since both the lower bound on T_0 and the upper bound on the derivative of u are uniform over \mathcal{A} , a \mathcal{A} -uniform bound on the stretch factor follows:

$$s(t, x) \leq \frac{t}{\sqrt{t^2 - x^2 u_{\max}} - \frac{x^2}{2} u'_{\max}}$$

From this you can derive a \mathcal{A} -uniform mute boundary. Therefore assume henceforth that ϕ is a \mathcal{A} -uniform mute.

GLOBAL ANALYSIS OF STATIONARY POINTS IN HYPERBOLIC MOVEOUT APPROXIMATION

Until further notice regard F etc. as depending on RMS square slowness u rather than on interval velocity v . Dependence on v , through the relatively easily analyzed map $v \mapsto u$, will be reintroduced at the end.

A short calculation shows that

$$p(t, x) = \frac{x}{t} u(T_0(t, x)).$$

Introduce the quantity Γ , with units of time:

$$\Gamma(t_0, x) = T_0(T^*(t_0, x), x)$$

That is, $\Gamma(t_0, x)$ is the zero offset time for which the time at offset x is the same in the slowness u as the time one obtains for t_0, x in slowness u^* .

Then introducing the expression for p , and changing variables from t to t_0 in the integral above, yields

$$J_0[u] = \int \int dt_0 dx B_0(t_0, x) (u(\Gamma(t_0, x)) - u^*(t_0))^2 (r^*(t_0))^2 + O(\lambda)$$

where

$$B_0(t_0, x) = \phi(T^*(t_0, x), x) B^*(T^*(t_0, x), x) \left(\frac{x}{T^*(t_0, x)} \right)^2$$

depends only on u^* and \mathcal{A} .

It is now straightforward to compute the first order perturbation δJ_0 of J_0 with respect to u . First,

$$\begin{aligned} \delta T(t_0, x) &= \frac{\frac{x^2}{2} \delta u(t_0)}{T(t_0, x)} \\ 0 &= \delta(T(T_0(t, x), x)) = \delta T(T_0(t, x), x) + \frac{\partial T}{\partial t_0}(T_0(t, x), x) \delta T_0(t, x) \\ &= \frac{\frac{x^2}{2} \delta u(T_0(t, x))}{t} + \frac{\delta T_0(t, x)}{s(t, x)} \end{aligned}$$

so

$$\delta T_0(t, x) = -\frac{\frac{x^2}{2} s(t, x) \delta u(T_0(t, x))}{t}$$

whence

$$\delta \Gamma(t_0, x) = \delta T_0(T^*(t_0, x), x) = -\frac{\frac{x^2}{2} s(T^*(t_0, x), x) \delta u(\Gamma(t_0, x))}{T^*(t_0, x)}$$

and

$$\begin{aligned} \delta(u(\Gamma(t_0, x))) &= \delta u(\Gamma(t_0, x)) + \frac{du}{dt_0}(\Gamma(t_0, x)) \delta \Gamma(t_0, x) \\ &= \delta u(\Gamma(t_0, x)) \left(1 - \frac{\frac{x^2}{2} s(T^*(t_0, x), x) \frac{du}{dt}(\Gamma(t_0, x))}{T^*(t_0, x)} \right) \end{aligned}$$

Recall that

$$s(t, x) = \frac{t}{T_0(t, x) + \frac{x^2}{2} \frac{\partial u}{\partial t_0}(T_0(t, x))}$$

so that

$$s(T^*(t_0, x), x) = \frac{T^*(t_0, x)}{\Gamma(t_0, x) + \frac{x^2}{2} \frac{\partial u}{\partial t_0}(\Gamma(t_0, x))}$$

so

$$\begin{aligned} \delta(u(\Gamma(t_0, x))) &= \delta u(\Gamma(t_0, x)) \left(1 - \frac{\frac{x^2}{2} \frac{du}{dt}(\Gamma(t_0, x))}{\Gamma(t_0, x) + \frac{x^2}{2} \frac{\partial u}{\partial t_0}(\Gamma(t_0, x))} \right) \\ &= \frac{\delta u(\Gamma(t_0, x))}{1 + \frac{x^2}{2\Gamma(t_0, x)} \frac{\partial u}{\partial t_0}(\Gamma(t_0, x))} = \frac{\Gamma(t_0, x)}{T^*(t_0, x)} s(T^*(t_0, x), x) \delta u(\Gamma(t_0, x)) \end{aligned}$$

Putting this all together,

$$\delta J_0[u] = \int \int dt_0 dx B_1(t_0, x) (u(\Gamma(t_0, x)) - u^*(t_0)) (r^*(t_0))^2 \delta u(\Gamma(t_0, x)) + O(\lambda)$$

where

$$B_1(t_0, x) = \frac{B_0(t_0, x)}{1 + \frac{x^2}{2\Gamma(t_0, x)} \frac{\partial u}{\partial t_0}(\Gamma(t_0, x))} = \frac{\Gamma(t_0, x)}{T^*(t_0, x)} B_0(t_0, x) s(T^*(t_0, x), x)$$

depends on u , u^* , and \mathcal{A} . To compute the gradient, change variables again to $t'_0 = \Gamma(t_0, x)$ for each x . Since

$$\begin{aligned} \frac{\partial \Gamma}{\partial t_0}(t_0, x) &= \frac{\partial}{\partial t_0}(T_0^*(T(t_0, x), x)) \\ &= \frac{\partial T_0^*}{\partial t}(T(t_0, x), x) \frac{\partial T}{\partial t_0}(t_0, x) = \frac{s^*(T(t_0, x), x)}{s(T(t_0, x), x)} \end{aligned}$$

and so

$$\frac{\partial \Gamma^{-1}}{\partial t_0}(t_0, x) = \frac{s(T^*(t_0, x), x)}{s^*(T^*(t_0, x), x)}$$

you get

$$\delta J_0[u] = \int \int dt_0 dx B_1^*(t_0, x) (u(t_0) - u^*(\Gamma^{-1}(t_0, x))) (r^*(\Gamma^{-1}(t_0, x)))^2 \delta u(t_0) + O(\lambda)$$

with

$$B_1^*(t_0, x) = B_1(\Gamma^{-1}(t_0, x), x) \frac{s(T^*(t_0, x), x)}{s^*(T^*(t_0, x), x)}.$$

Thus the L^2 gradient of J_0 is

$$\nabla J_0[u](t_0) = \int dx B_1^*(t_0, x) (u(t_0) - u^*(\Gamma^{-1}(t_0, x))) (r^*(\Gamma^{-1}(t_0, x)))^2 + O(\lambda)$$

Both expressions for J_0 and its gradient suggest that these quantities are comparing the trial square slowness u and the target square slowness u^* at different points (eg. t_0 vs. $\Gamma^{-1}(t_0, x)$), and this in turn makes understanding of the implications for determination of u difficult. Fortunately this is not really the case:

Key Lemma: There exists a function $h(t_0, x)$, depending on velocity v (or slowness u) and also on u^* and \mathcal{A} , having the following properties:

- $h(t_0, x) > 0$ over the mute zone, and $\log h(t_0, x)$ is uniformly bounded for t_0, x in the mute zone and $v \in \mathcal{A}$;
- $u(t_0) - u^*(\Gamma^{-1}(t_0, x)) = h(t_0, x)(u(t_0) - u^*(t_0))$

Proof of Key Lemma: Note first that since

$$T^*(T_0^*(t, x), x) = t$$

$$\begin{aligned} \frac{\partial T^*}{\partial x}(T_0^*(t, x), x) + \frac{\partial T^*}{\partial t_0}(T_0^*(t, x), x) \frac{\partial T_0^*}{\partial x}(t, x) &= 0 \\ &= \frac{xu^*(T_0^*(t, x))}{t} + \frac{1}{s^*(t, x)} \frac{\partial T_0^*}{\partial x}(t, x) \end{aligned}$$

so

$$\frac{\partial T_0^*}{\partial x}(t, x) = -s^*(t, x) \frac{xu^*(T_0^*(t, x))}{t}.$$

It follows that since

$$\begin{aligned} \Gamma^{-1}(t_0, x) &= T_0^*(T(t_0, x), x), \\ \frac{\partial \Gamma^{-1}}{\partial x}(t_0, x) &= \frac{\partial T_0^*}{\partial t}(T(t_0, x), x) \frac{\partial T}{\partial x}(t_0, x) + \frac{\partial T_0^*}{\partial x}(T(t_0, x), x) \\ &= s^*(T(t_0, x), x) \frac{xu(t_0)}{T(t_0, x)} - s^*(T(t_0, x), x) \frac{xu^*(T_0^*(T(t_0, x), x))}{T(t_0, x)} \\ &= \frac{x s^*(T(t_0, x), x)}{T(t_0, x)} (u(t_0) - u^*(\Gamma^{-1}(t_0, x))) \end{aligned}$$

Thus

$$\begin{aligned} u(t_0) - u^*(\Gamma^{-1}(t_0, x)) &= u(t_0) - u^*(t_0) + \int_0^x dx' \frac{\partial}{\partial x'} (u(t_0) - u^*(\Gamma^{-1}(t_0, x'))) \\ &= u(t_0) - u^*(t_0) - \int_0^x dx' \frac{\partial u^*}{\partial t_0}(\Gamma^{-1}(t_0, x')) \frac{\partial \Gamma^{-1}}{\partial x'}(t_0, x') \\ &= u(t_0) - u^*(t_0) - \int_0^x dx' \frac{\partial u^*}{\partial t_0}(\Gamma^{-1}(t_0, x')) \frac{x' s^*(T(t_0, x'), x')}{T(t_0, x')} (u(t_0) - u^*(\Gamma^{-1}(t_0, x'))) \\ &= u(t_0) - u^*(t_0) + \int_0^x dx' g(t_0, x') (u(t_0) - u^*(\Gamma^{-1}(t_0, x'))) \end{aligned}$$

where

$$g(t_0, x) = -\frac{\partial u^*}{\partial t_0}(\Gamma^{-1}(t_0, x')) \frac{x' s^*(T(t_0, x'), x')}{T(t_0, x')}$$

This simple integral equation has the solution

$$u(t_0) - u^*(\Gamma^{-1}(t_0, x')) = h(t_0, x)(u(t_0) - u^*(t_0))$$

where

$$h(t_0, x) = \exp\left(\int_0^x dx' g(t_0, x')\right)$$

has the properties claimed for it in the statement of the lemma. **Q.E.D.**

Now changing variables in the asymptotic formula for J_0 , and applying the above relation to both this and the formula for ∇J_0 , you obtain

$$J_0[u] = \int \int dt_0 (u(t_0) - u^*(t_0))^2 dx B_0^*(t_0, x) h(t_0, x) (r^*(\Gamma^{-1}(t_0, x)))^2 + O(\lambda)$$

$$\nabla J_0[u](t_0) = (u(t_0) - u^*(t_0)) \int dx B_1^*(t_0, x) h(t_0, x) (r^*(\Gamma^{-1}(t_0, x)))^2 + O(\lambda)$$

where

$$B_0^*(t_0, x) = B_0(\Gamma^{-1}(t_0, x), x) \frac{s(T^*(t_0, x), x)}{s^*(T^*(t_0, x), x)}.$$

Now B_0^* and B_1^* differ at each point in the mute zone by factors or divisors of s, s^*, T , and the like, and these are bounded over the mute zone uniformly in $v \in \mathcal{A}$. Therefore there exists a constant $C > 0$ depending only on \mathcal{A} for which

$$J_0[u] \leq C \int dt_0 (u(t_0) - u^*(t_0)) \nabla J_0[u](t_0) + O(\lambda)$$

and we have proved the

Theorem: If u , the RMS square slowness for $v \in \mathcal{A}$, is a stationary point of $J_0[u]$, then $J_0[u] = O(\lambda)$.

That is, for noise free data, any stationary point of J_0 is a global minimizer, up to an asymptotically vanishing error.

NOISE: GENERAL CASE

Suppose that the data $S(t, x)$ is the sum of model-consistent data and another field, regarded as noise or error:

$$S(t, x) = S^*(t, x) + E(t, x)$$

where “model-consistent” means as before

$$S^*(t, x) = r^*(T_0^*(t, x)) + O(\lambda)$$

and $E(t, x)$ is arbitrary (but finite “energy” = mean square).

Since there are several data running around in this part of the discussion, include the name of the data in the notation for the differential semblance objective:

$$J_0[v, S] = \frac{1}{2} \|H\phi F[v]WG[v]S\|^2$$

etc. Then

$$J_0[v, S] = J_0[v, S^*] + J_0[v, E] + K[v, S^*, E]$$

where

$$\begin{aligned} K[v, S^*, E] &= \langle H\phi F[v]WG[v]S^*, H\phi F[v]WG[v]E \rangle \\ &= \int \int dx dt \phi \left(\frac{\partial}{\partial x} + p \frac{\partial}{\partial t} \right) S \left[(I - \nabla^2)^{-1} \phi \left(\frac{\partial}{\partial x} + p \frac{\partial}{\partial t} \right) E \right] \end{aligned}$$

satisfies

$$|K[v, S^*, E]| \leq C \|S^*\| \|E\|$$

Here and in the following, C will stand for a constant uniform over $v \in \mathcal{A}$ (though the precise value may vary from display to display).

Likewise,

$$J[v, E] \leq C \|E\|^2$$

Similarly, the gradients *with respect to RMS square slowness* u satisfy

$$\nabla J_0[v, S] = \nabla J_0[v, S^*] + \nabla J_0[v, E] + \nabla K[v, S^*, E]$$

and

$$\|\nabla K[v, S^*, E]\| \leq C \|S^*\| \|E\|$$

$$\|\nabla J_0[v, E]\| \leq C \|E\|^2$$

Suppose that u (or its corresponding v) is a stationary point of $J_0[v, S]$, i.e. $\nabla J_0[v, S] = 0$. Then

$$\begin{aligned} J_0[v, S] &\leq J_0[v, S^*] + C \|E\| (\|S^*\| + \|E\|) \\ &\leq C (\langle u - u^*, \nabla J_0[v, S^*] \rangle + \|E\| (\|S^*\| + \|E\|)) + O(\lambda) \\ &= C (\langle u - u^*, \nabla J_0[v, S] - \nabla J_0[v, E] - \nabla K[v, S^*, E] \rangle + \|E\| (\|S^*\| + \|E\|)) + O(\lambda) \\ &\leq C \|E\| (\|S^*\| + \|E\|) + O(\lambda) \end{aligned}$$

If you presume that the data error is less than 100%, i.e. $\|E\| \leq \|S^*\|$, which seems reasonable (or pick any other fixed percentage, if 100% seems wrong to you - just absorbs in C), then this becomes

$$J_0[v, S] \leq C \|E\| + O(\lambda)$$

That is,

Theorem: At a stationary point of the differential semblance objective, its value is bounded by a \mathcal{A} -uniform multiple of the distance of the data to the set of model-consistent data.

Thus for a family of data converging to model-consistent data, any set of corresponding stationary points of J_0 must have J_0 values which converge to zero, modulo asymptotic errors.

This result may well *not* imply that stationary points for noisy data are global minima. Indeed, substitute the “target” velocity v^* in the expression for $J_0[v, S]$: from the expansion and estimates above you easily see that

$$J[v^*, S] \leq C (\|E\|^2 + O(\lambda) \|E\|)$$

Certainly one hopes that the asymptotic error is no worse than other errors, in particular than the data error E , so this inequality effectively implies that the global minimum value of $J_0[\cdot, S]$

is proportional to $\|E\|^2$ for near consistent data, whereas the theorem shows only that the stationary values are proportional to $\|E\|$, so presumably larger at least in some cases.

In the next section I will show that when the differential semblance minimization is supplemented with proper constraints on the velocity model, in addition to those already imposed, the error in the *RMS square slowness* is proportional to the error in the data. It then follows from the estimates above that stationary values conforming to these constraints are indeed proportional to the square of the error level, hence essentially global minima. It would be interesting to know whether relaxing these constraints actually permits anomalously large stationary values.

DATA DRIVEN MODEL PARAMETRIZATION AND OPTIMAL ERROR ESTIMATES

Up to this point I have imposed only minimal constraints on the RMS velocity, namely those necessary to justify use of the convolutional model. Most velocity analysis imposes far more stringent constraints, either explicitly or implicitly, in the form of *parsimonious parametrization* or *regularization*. In the former case, the choice of parameters (eg. how many spline nodes, where to place them) is *ad hoc*. In the latter, the type of regularization (first derivative, second derivative,...) and the choice penalty weight are also obscure.

In this section I suggest that the differential semblance objective itself supplies a mechanism for constraining the velocity to a parsimoniously parametrized space. I'll propose a choice of subspace within which

- the global minimum is unique for noise free data;
- the error in RMS square slowness is proportional to the error in the data, and so
- any stationary values are proportional to the square of the data error energy, so essentially global minima.

Assume until further notice that the data is free of noise:

$$S(t, x) = r^*(T_0^*(t, x)) + O(\lambda)$$

The Key Lemma proved in the last section then implies that the Hessian $\nabla\nabla J_0$ takes the form

$$\begin{aligned} \nabla\nabla J_0[u]\delta u(t_0) &= \delta u(t_0) \int dx B_1^*(t_0, x) h(t_0, x) (r^*(\Gamma^{-1}(t_0, x)))^2 + O(\lambda, \|u - u^*\|) \\ &= \tilde{R}[u](t_0)\delta u(t_0) \end{aligned}$$

While the expression for \tilde{R} above is not easily computable, the approximation

$$R[u](t_0) = \int dx (r^*(\Gamma^{-1}(t_0, x)))^2$$

is simply the stack of the squared prestack reflectivity estimates, and therefore an inexpensive byproduct of the computation. At $u = u^*$,

$$\nabla \nabla J_0[u^*] \delta u(t_0) = b[u](t_0) R[u^*](t_0) \delta u(t_0)$$

i.e. the Hessian is actually the approximation followed by a positive diagonal scaling.

Now suppose that u^* differs from a reference square slowness u^0 (in practice, an initial estimate) by a member of a space W . Introduce an inner product in the space of W by

$$\langle w_1, w_2 \rangle_2 = \int \frac{d^2 w_1}{dt_0^2} \frac{d^2 w_2}{dt_0^2}$$

To make this inner product positive definite, thus defining a Hilbert space structure, assume furthermore that

$$w(0) = \frac{dw}{dt_0}(0) = w(t_0^{\max}) = \frac{dw}{dt_0}(t_0^{\max}) = 0$$

Thus W is a subspace of the Sobolev space $H_0^2([0, t_0^{\max}])$.

Since the interval velocities, hence the RMS square slownesses, are supposed to vary over a bounded set in C^∞ , membership in \mathcal{A} entails a bound on the W norm of $u - u^0$.

Let $g(t_0^1, t_0^2)$ be the Green's function for the operator

$$\frac{d^4}{dt_0^4}$$

with the boundary conditions stated above. Then the W gradient of J_0 restricted to $u^0 + W$ is

$$\nabla_W J_0[u] = \mathcal{G} \nabla J_0[u]$$

in which \mathcal{G} denotes the operator with kernel g . Similarly,

$$\nabla_W \nabla_W J_0[u] = \mathcal{G} \nabla \nabla J_0[u]$$

Next suppose that $H[u]$ is uniformly positive definite for all $u \in \mathcal{A}$. That is, there exist $0 < h_* \leq h^*$ for which

$$h_* \|w\|^2 \leq \langle w, H[u]w \rangle_2 \leq h^* \|w\|_2^2$$

for all $w \in W, u \in \mathcal{A}$.

Then there exists a similar uniform bound for $\nabla_W \nabla_W J[u]$, since the latter differs from $H[u]$ by a diagonal scaling operator with uniform upper and lower bounds over \mathcal{A} . For the same reason,

$$\|\nabla_W J_0[u]\|_2 = \|\mathcal{G} b[u] H[u](u - u^*)\|_2 \geq l_* \|u - u^*\|_2$$

for a suitable $l_* > 0$.

That is: within $u^0 + W$, u^* is the unique stationary point of J_0 .

Moreover, consulting the estimates of the last section, you see that if the search is limited to $u^0 + W$, then at a stationary point u ,

$$J_0[u, S] \leq C(\|E\|^2)$$

as claimed, since the cross-term $K[u, S^*, E]$ in the notation of that section is bounded by a multiple of $\|u - u^*\|_2$.

Finally, how does one lay hands on such a paragon of a function space as W with the properties supposed here? The operator $H[u]$ is symmetric positive semidefinite on $H_0^2([0, t_0^{\max}])$. An optimal choice for W is the direct sum of eigenspaces of $\nabla_W \nabla_W J_0[u^*]$ corresponding to the eigenvalues above the cutoff level h_* . A computable estimate of this space is the corresponding direct sum of eigenspaces of $H[u]$. A basis consists of eigenfunctions of the Sturm-Liouville problem

$$\frac{d^4 w}{dt_0^4} = \frac{R[u]}{\lambda} w,$$

$$w(0) = \frac{dw}{dt_0}(0) = w(t_0^{\max}) = \frac{dw}{dt_0}(t_0^{\max}) = 0$$

To construct W , find the eigenfunctions of this problem, and choose those whose eigenvalues lie above a “suitable” cutoff.

Note that if there is little data in a t_0 interval, $R[u]$ will be small in that interval and eigenfunctions of the 4th derivative operator will smoothly interpolate values to either side. Thus my suggested space implicitly “picks events” with significant energy, pins the RMS velocity down at those places, and interpolates between “events” - just as a human velocity analyst would.

It remains to analyse this “picking” effect, and to devise good algorithms for choosing the eigenvalue cutoff as a function of data quality and success in fitting moveout (i.e. minimizing J_0), so as to justify the assumption that $u^* \in u^0 + W$. But that’s another story...

ACKNOWLEDGEMENT

The author performed part of the work reported here while a guest of the Stanford Exploration Project. The author thanks the Director of SEP, Prof. Jon Claerbout, and the other project members for their hospitality and for the stimulating intellectual atmosphere of SEP.

This work was partially supported by the National Science Foundation under grant numbers DMS 9404283 and DMS 9627355; the Office of Naval Research under grant number N00014-96-1-0156; the U. S. Department of Energy under grant number DE-FG07-97 ER14827; and The Rice Inversion Project. TRIP Sponsors for 1998 are Amerada Hess, Amoco Production Co., ARCO Oil and Gas, Conoco Inc., Discovery Bay, Exxon Production Research Co., Mobil Research and Development Corp., and Shell International Research.

REFERENCES

- Al Yahya, K., 1989, Velocity analysis by iterative profile migration: *Geophysics*, **54**, 718–729.
- Araya, K., Kim, S., Nolan, C. J., and Symes, W. W., 1996, Velocity inversion and high frequency asymptotics: *Mathematical Methods in Geophysical Imaging: Society of Photo-optical Instrumentation Engineers*.
- Cao, D., Singh, S., and Tarantola, A., 1990, Simultaneous inversion for background velocity and impedance maps: *Geophysics*, **55**, 458–469.
- Chauris, H., Noble, M., and Podvin, P., 1998a, Testing the behaviour of differential semblance for velocity estimation: 68th Annual International Meeting and Exposition, Society of Exploration Geophysicists, 1305–1308.
- Chauris, H., Noble, M., and Podvin, P., 1998b, Testing the behaviour of differential semblance for velocity optimization: European Association for Geoscientists and Engineers Annual Meeting.
- Claerbout, J. F., 1992, *Earth sounding analysis: Processing versus inversion*: Blackwell, Boston.
- Clément, F., and Chavent, G., 1993, Waveform inversion through MBTT formulation *in* Kleinman, E., Angell, T., Colton, D., Santosa, F., and Stakgold, I., Eds., *Mathematical and Numerical Aspects of Wave Propagation*:: SIAM.
- Fowler, P., 1986, Migration velocity analysis by optimization: linear theory: Proc. 56th Annual International Meeting, Society of Exploration Geophysicists, 660–662.
- Gauthier, O., Tarantola, A., and Virieux, J., 1986, Two-dimensional nonlinear inversion of seismic waveforms: *Geophysics*, **51**, 1387–1403.
- Gockenbach, M. S., and Symes, W. W., 1997, Duality for inverse problems in wave propagation *in* Biegler, L., Coleman, T., Santosa, F., and Conn, A., Eds., *Large Scale Optimization*:: Springer Verlag.
- Gockenbach, M. S., Symes, W. W., and Tapia, R. A., 1995, The dual regularization approach to seismic velocity inversion: *Inverse Problems*, **11**, no. 3, 501–531.
- Kern, M., and Symes, W. W., 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples: *Geophysical Prospecting*, **99**, 565–614.
- Kim, S., and Symes, W. W., 1998, Smooth detectors of linear phase: *Inverse Problems*, **14** (1), 101–112.
- Kolb, P., Collino, F., and Lailly, P., 1986, Prestack inversion of a 1D medium: *Proceedings of IEEE*, **74**, 498–506.

- Martinez, R., and McMechan, G., 1991, τ -p seismic data for viscoelastic media – Part I: Modelling; Part II: Linearized Inversion: *Geophysical Prospecting*, **39**, 141–181.
- Minkoff, S. E., and Symes, W. W., 1997, Full waveform inversion of marine reflection data in the plane-wave domain: *Geophysics*, **62**, no. 2, 540–553.
- Reshef, M., 1997, The use of 3D prestack depth imaging to estimate layer velocities and reflector positions: *Geophysics*, **62**, 206–210.
- Scales, J., Smith, M., and Fischer, T., 1991, Global optimization methods for highly nonlinear inverse problems *in* Cohen, G., Halpern, L., and Joly, P., Eds., *Mathematical and Numerical Aspects of Wave Propagation Phenomena*: SIAM, 434–444.
- Sen, M., and Stoffa, P., 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: *Geophysics*, **56**, 1624–1636.
- Sevink, A., and Herman, G., 1993, *in* Kleinman, R., Angell, T., Colton, D., Santosa, F., and Stakgold, I., Eds., *Mathematical and Numerical Aspects of Wave Propagation*: SIAM.
- Song, H., 1994, On a transmission inverse problem: Ph.D. thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.
- Symes, W. W., and Carazzone, J., 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, no. 5, 654–663.
- Symes, W. W., 1993, A differential semblance criterion for inversion of multioffset seismic reflection data: *J. Geoph. Res.*, **98**, 2061–2073.
- Symes, W. W., 1997, High frequency asymptotics and velocity estimation: *Depth Imaging Symposium*: Geophysical Society of Tulsa.
- Symes, W. W., 1998, Differential semblance estimation of velocities: Progress report: TRIP 1998 Annual Report (<http://www.trip.caam.rice.edu>).
- Taner, M., and Koehler, F., 1969, Velocity spectra: digital computer derivation and application of velocity functions: *Geophysics*, **34**, 859–881.
- Toldi, J., 1985, Velocity analysis without picking: Proc. 55th Annual International Meeting, Society of Exploration Geophysicists.
- Yilmaz, O., and Chambers, R., 1984, Migration velocity analysis by wavefield extrapolation: *Geophysics*, **49**, 1664–1674.

Coherent noise suppression in velocity inversion¹

William W. Symes²

ABSTRACT

Data components with well-defined moveout other than primary reflections are sometimes called *coherent noise*. Coherent noise makes velocity analysis ambiguous, since no single velocity function explains incompatible moveouts simultaneously. Contemporary data processing treats the control of coherent noise influence on velocity as an interpretive step. Dual regularization theory suggests an alternative, automatic inversion algorithm for suppression of coherent noise when primary reflection phases dominate the data. Experiments with marine data illustrate the robustness and effectiveness of the algorithm.

INTRODUCTION

Velocity analysis quantifies and parametrizes moveout in terms of velocity functions. In common with most other parts of conventional processing, velocity analysis rests on the linearized model of reflections, which treats short scale components of Earth mechanical structure function as perturbations of the large scale components. In fact, most processing derives from the acoustic version of this model, which predicts only one family of moveout curves, or phases, those of so-called compressional wave primary reflections.

Some seismic data exhibit the characteristics predicted by this *acoustic primaries-only model* to good approximation. Other data exhibit several phases, however; while one of these usually appears to be a compressional wave primary phase, others may represent multiple reflections, mode conversions, 3D reflection phenomena in data treated as 2D, and so forth. These other phases may carry considerable energy. Multiple reflection energy is suppressed to some extent by various multiple removal techniques, but none is universally effective in removing all phases but the primary. Residual non-primary phases pose an obstacle to velocity estimation, in that a single velocity function cannot predict several moveout families simultaneously (within the linearized acoustic approximation). The conventional approach to moveout ambiguity is visual and interpretive: the processor is expected to reject coherent noise by interactively updating velocity functions to recognize and flatten selectively the primary events in image gathers, recognize and fit only primary reflection peaks, and so on. It seems odd that the most robust information in seismic data must in the end be teased out by hand, especially as the size of 3D datasets precludes visual inspection of all but a small fraction of the prestack data.

¹This report will also appear in the TRIP 1999 annual report

²email: symes@caam.rice.edu

This paper presents an alternative approach to coherent noise rejection, based on a formulation of velocity analysis as an inverse problem, when primary reflection energy dominates the (preprocessed) data. The idea is quite simple. Flatness of image gathers diagnoses the success of a velocity analysis. Image gathers created from data containing multiple phases are impossible to flatten. Therefore creation of flat image gathers requires data perturbation. If the primary phase is dominant, then *the smallest data perturbation permitting flat image gathers will be that which removes the non-primary phases.*

The mathematical embodiment of this idea is the *dual regularization theory* of velocity inversion, introduced in Gockenbach et al. (1995); Gockenbach and Symes (1997). Given a relative noise level σ , we seek the velocity function v and the data perturbation of root mean square relative size at most σ which together yield the flattest image gather. To measure flatness, we use the *differential semblance* criterion, introduced in Symes (1986) and developed in a series of papers [for example Symes (1998b); Chauris et al. (1998)]. In fact, differential semblance is the only semblance measure providing the good theoretical properties needed to ensure the reliability of coherent noise rejection (Kim and Symes, 1998; Symes, 1998a).

This next section describes a simple algorithm for solution of this constrained optimization problem. An example using a marine CMP and layered acoustic modeling demonstrates the coherent noise rejection permitted by reasonable estimates of noise level σ . The final section summarizes our conclusions and formulates a few directions for further research.

ALGORITHM

Given a relative noise level σ , we seek the velocity function v and the data perturbation of root mean square relative size at most σ which together yield the flattest image gather. Transformation of this idea into an implementable algorithm requires definition of operators, functions, and optimization methods. This section gives a sketch of these mathematical details.

The *forward map* $F[v]$ is a linear operator depending on a velocity function v . The velocity function depends on all or part of the subsurface coordinates; the examples presented below use depth-dependent velocity. $F[v]$ is a *prestack* forward modeling operator; it takes an image volume or bin-dependent reflectivity as input, and outputs a seismic data volume. In the examples presented below, the data will be a common midpoint gather, each bin will contain a single trace, and the bin parameter is offset. Thus the input reflectivity also has the appearance of a common midpoint gather, and can be identified with the image gather in this setting.

The *inverse map* $G[v]$ is an approximate inverse to $F[v]$. That is, if data d and reflectivity r satisfy $d = F[v]r$, then $r \simeq G[v]d$. For multioffset data and multidimensional models, Beylkin (1985) showed how to build such operators as weighted diffraction sums. For layered modeling, $G[v]$ is essentially moveout correction, after compensation for amplitude and wavelet deconvolution; $F[v]$ inverts these steps.

Differential semblance measures nonflatness by comparing neighboring image bins. That is, if the image is $r = G[v]d$, and the bin index is i , then the differential semblance is the mean square power of $(Dr)_i = \text{const.} \times (r_{i+1} - r_i)$. A convenient notation for root mean

square of a field, say r , is $\|r\|^2$. The dot product of two fields (viewed as vectors of samples), say r_1 and r_2 , is $\langle r_1, r_2 \rangle$. Thus $\|r\|^2 = \langle r, r \rangle$. The basic (“raw”) semblance operator is $W[v] = F[v]DG[v]$. The application of the modeling operator $F[v]$ after formation of the bin difference makes the power of the output independent of amplitude, up to an error which decays with increasing signal frequency. [This trick was discovered by Hua Song (Song, 1994)]. Since the data is differenced in formation of $W[v]d$, we bring its high frequency content back into consistency with that of the data *via* a *smoothing operator* H of order -2 (k^{-2} filter).

The dual regularization objective function J_σ is then

$$J_\sigma[v; d] = \min_r \frac{1}{2} \langle W[v]r, HW[v]r \rangle \text{ subj } \|r - d\| \leq \sigma$$

Note that the differential semblance objective explored in the above cited references is the special case of this one with $\sigma = 0$. In general, a Lagrange multiplier λ exists for which the solution r satisfies the *normal* and *secular* equations:

$$W[v]^T HW[v]r + \lambda(r - d) = 0, \|r - d\| = \sigma$$

These two equations together determine $\lambda = \lambda[v; \sigma]$ and $r = r[v; \sigma]$. Thus J_σ is

$$J_\sigma[v; d] = \frac{1}{2} \langle W[v]r[v; \sigma], HW[v]r[v; \sigma] \rangle$$

First order perturbation with respect to v , $v \rightarrow v + \delta v$, gives

$$\delta J_\sigma[v; d] = \langle W[v]r[v; \sigma], H\delta W[v]r[v; \sigma] \rangle$$

after simplifications due to the normal and secular equations. From this expression follows a formula for the gradient of J_σ in terms of the first order perturbation of $W[v]$ and its adjoint operator, which may in turn be expressed as products of the operators $F[v]$, D , and $G[v]$, their first order perturbations, and the adjoints of these.

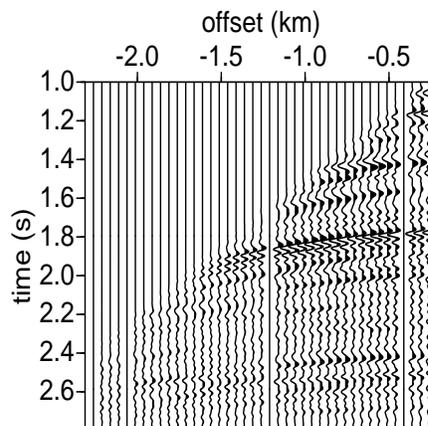
Besides accurate numerical implementations of the operators described above, we require methods for solving the system of normal and secular equations. For the latter, we use the Moré-Hebden algorithm (Björk, 1997) with the linear systems occurring in this method solved approximately by conjugate gradient iteration. The best estimate of v results from gradient-based optimization (a quasi-Newton method) applied to J_σ . In the experiments reported here, we have used the Limited Memory Broyden-Fletcher-Goldfarb-Shanno method as developed in Nocedal (1980).

EXAMPLE: A CMP FROM THE NORTH SEA

Figure 1 shows a common midpoint gather from the Mobil AVO data set (Keys and Foster, 1998). This part of the North Sea covers relatively flat lying sediments to a depth of 2 s, where an unconformity introduces older, more deformed rock which is nonetheless still for the most

part flat lying. Therefore layered modeling seems reasonable for this data, at least to perhaps 3 s and as a first approximation. The work reported here views whatever converted wave energy is present in the data as noise, so acoustic modeling is reasonable. Finally, as the range of offsets in this data is modest (2.5 km maximum), the hyperbolic moveout approximation seemed likely to be adequate, at least when combined with an aggressive mute as displayed in Figure 1.

Figure 1: CMP from Mobil AVO data. `bill2-cmpfig` [NR]



Strong surface related multiple energy is characteristic of this region. The main preprocessing steps were hyperbolic Radon filtering and bandpass filtering. The Radon filter suppressed but did not entirely remove coherent noise: it seemed reasonable to hope that primary energy dominates the filtered data, as is required by the dual regularization strategy. The bandpass filter ensured that the data were not spatially aliased, so that the differential semblance could be computed accurately.

Minimization of $J_0[v;d]$ produced the RMS velocity displayed in Figure 2, which exhibits a characteristic feature of the differential semblance function: when faced with contradictory moveout (as for example in the interval 1.8-2.4 s), it averages the apparent velocities to come as close as possible to flattening all events. The moveout corrected data ($G[v]d$ in the notation of the last section) displayed in Figure 3 shows a mixture of overcorrected and undercorrected events. The minimization process (*via* a quasi-Newton algorithm) required approximately 12 s on an SGI Origin2000 processor.

We used the output of the $J_0[v;d]$ minimization as the initial guess for minimization of $J_{0.5}[v;d]$; the latter required approximately 3 min on the Origin. Figure 4 shows that $\sigma = 0.5$ was not a bad guess at the level of coherent noise: the automatic velocity analysis has now essentially ignored slow (multiple reflection) and smaller fast (steeply dipping or out of plane) phases and placed its estimate of RMS velocity squarely in the main corridor of apparent primary reflection phases, as one also sees in the the conventional image gather (= moveout corrected data $G[v]d$, Figure 5). Dual regularization also produces an *inverted reflectivity* ($r[v;\sigma]$ in the last section, Figure 6) or denoised data, which is superior to the conventional

Figure 2: V_{RMS} from velocity inversion, $\sigma = 0.0$, overplotted on velocity spectrum. Note that the estimated RMS velocity navigates between peaks. `bill2-sigdisplay0` [NR]

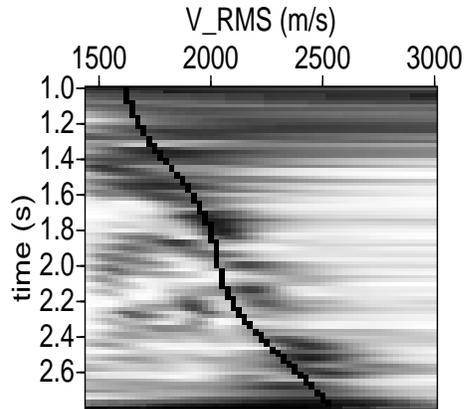


Figure 3: Image gather (= NMO corrected CMP) using V_{RMS} from velocity inversion with $\sigma = 0.0$. Note residual curvature in all events. `bill2-siginvdata0` [NR]

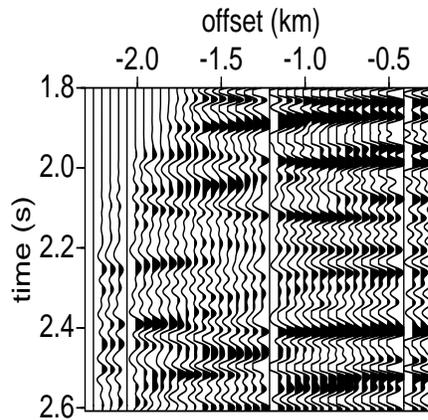


image gather as a basis for further processing.

Figure 4: V_{RMS} from velocity inversion, $\sigma = 0.5$, overplotted on velocity spectrum. Note that the estimated RMS velocity picks apparent primary phases. `bill2-sigdisplay5` [NR]

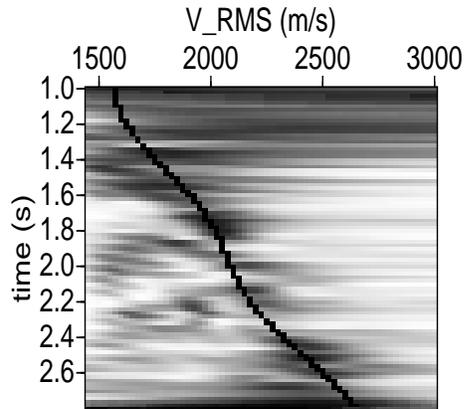
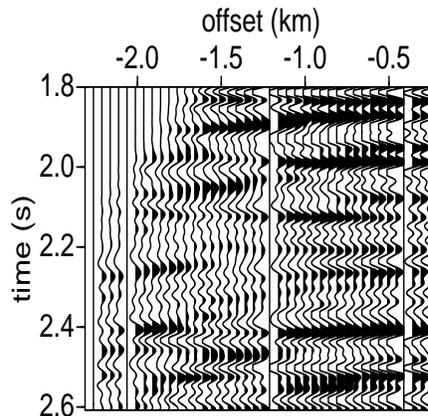


Figure 5: Image gather (= NMO corrected CMP) using V_{RMS} from velocity inversion with $\sigma = 0.5$. Primary reflections are essentially flat. `bill2-siginvdata5` [NR]

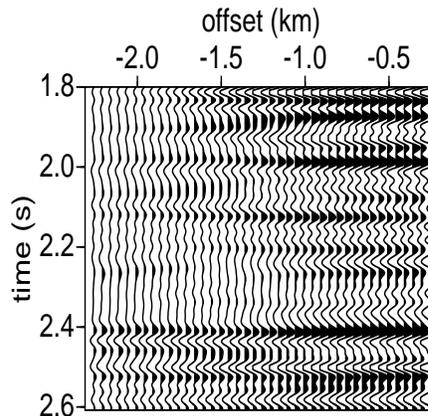


DISCUSSION

The results exhibited in the last section are characteristic of dual regularization: it robustly identifies the dominant moveout trend, so long as the linear solves in the Moré-Hebden algorithm are sufficiently precise. It is also a great deal more expensive than “raw” differential semblance (J_0) minimization, often by a factor of 10-15, due to the need to solve linear systems.

Note that dual regularization does not eliminate the need for preprocessing to reduce coherent noise: the desired primary reflection phase must be energetically dominant, whereas

Figure 6: Inverted reflectivity, $\sigma = 0.5$, essentially a cleaned up version of the data, with non-primary phases suppressed. `bill2-sigestrefl5` [NR]



very strong multiple reflection phases are common in some areas.

Speed improvements should be possible through better heuristics and algorithmic tuning, and perhaps through more effective constrained optimization. Since differential semblance is also effective in estimating laterally heterogeneous models (Symes and Versteeg, 1993; Chauris et al., 1998), dual regularization can also be applied in that context; of course, algorithmic efficiency will then become even more of an issue.

Dual regularization also implies a strategy for determination of σ : it should assume the smallest value for which the minimum of J_σ is (essentially) zero. Application of this noise level determination algorithm requires a method for estimating a tolerance for this minimum, a matter currently under study.

ACKNOWLEDGEMENT

The authors thank the National Science Foundation, the Office of Naval Research, the U. S. Department of Energy, and the sponsors of The Rice Inversion Project for support of this work. The first author (WWS) performed part of the research reported here while a guest of the Stanford Exploration Project; WWS thanks SEP members and director Prof. Jon Claerbout for their hospitality and many stimulating discussions.

REFERENCES

Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform: *J. Math. Phys.*, **26**, 99–108.

- Björk, A., 1997, Numerical methods for least squares problems: Society for Industrial and Applied Mathematics, Philadelphia.
- Chauris, H., Noble, M., and Podvin, P., 1998, Testing the behaviour of differential semblance for velocity estimation: 68th Annual International Meeting and Exposition, Society of Exploration Geophysicists, 1305–1308.
- Gockenbach, M., and Symes, W. W., 1997, Duality for inverse problems in wave propagation *in* Biegler, L., Coleman, T., Santosa, F., and Conn, A., Eds., Large Scale Optimization:: Springer Verlag.
- Gockenbach, M. S., Symes, W. W., and Tapia, R. A., 1995, The dual regularization approach to seismic velocity inversion: *Inverse Problems*, **11**, no. 3, 501–531.
- Keys, R. G., and Foster, D. J., 1998, Comparison of seismic inversion methods on a single real data set: Society of Exploration Geophysicists, Tulsa.
- Kim, S., and Symes, W. W., 1998, Smooth detectors of linear phase: *Inverse Problems*, **14** (1), 101–112.
- Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: *Mathematics of Computation*, **95**, 339–353.
- Song, H., 1994, On a transmission inverse problem: Ph.D. thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.
- Symes, W. W., and Versteeg, R., 1993, Velocity model determination using differential semblance optimization: Society of Exploration Geophysicists 63rd Annual International Meeting, 696–699.
- Symes, W. W., 1986, Stability and instability results for inverse problems in several-dimensional wave propagation *in* Glowinski, R., and Lions, J., Eds., Proc. 7th International Conference on Computing Methods in Applied Science and Engineering:: North-Holland.
- Symes, W. W., All stationary points of differential semblance are asymptotic global minimizers: Layered acoustics:, Technical report, The Rice Inversion Project, <http://www.trip.caam.rice.edu>, 1998.
- Symes, W. W., 1998b, High frequency asymptotics, differential semblance, and velocity estimation: Society of Exploration Geophysicists 68th Annual International Meeting, 1616–1619.

Angle-domain common image gathers by wave-equation migration

Marie L. Prucha, Biondo L. Biondi, and William W. Symes¹

ABSTRACT

Shot- and offset-domain common image gathers encounter problems in complex media. They can place events that come from different points in the subsurface at one subsurface location based on identical arrival times and horizontal slownesses. Angle-domain common image gathers uniquely define ray couples for each point in the subsurface, therefore each event in the data will be associated with only one subsurface location. It is possible to generate angle-domain common image gathers with wave-equation migration methods and these angle-domain common image gathers may be used for velocity analysis and amplitude-versus-angle analysis. Applications of these methods to the Marmousi model are promising.

INTRODUCTION

Current depth imaging technology works very well in areas that have slow velocity variations but may fail in more complex areas (Claerbout, 1985) for a variety of reasons, such as multiple reflections, bad velocities, and spatial aliasing. One potential cause of imaging failure is reflector location ambiguity due to multipathing of reflected energy: it is possible that a single event recorded in the data at one surface location could come from reflectors at two or more subsurface locations. Besides contributing to imaging artifacts, reflector ambiguity contributes non-flat events to common image gathers, thus rendering velocity analysis ambiguous (Nolan and Symes, 1996).

Several authors have suggested *angle domain* imaging as a solution for the reflector ambiguity (Xu et al., 1998; Brandsberg-Dahl et al., 1999). Angle domain sections collect the energy in a data set which has scattered over a specific reflection (“opening”) angle θ . We will argue below that an event in an angle section uniquely determines a ray couple, which in turn uniquely locates the reflector. Thus imaging artifacts and velocity update ambiguity due to multipathing are eliminated in this domain.

Multipathing is better handled by wave-equation migration methods than Kirchhoff ones, therefore the former are a natural choice for producing angle-domain common image gathers (CIGs). We present a simple method for extracting CIGs from 3-D prestack data downward

¹email: marie@sep.stanford.edu, biondo@sep.stanford.edu, symes@sep.stanford.edu

continued using the Double Square Root equation (DSR). The method is based on a slant-stack decomposition of the downward continued wavefield at each *depth* level. Our method is thus different from the method proposed by Ottolini and Claerbout (1984), that applies the DSR to downward continue prestack data slant-stacked at the *surface*. In layered media the two methods should produce equivalent results, but in presence of lateral velocity variations plane-wave downward continuation is not strictly valid and true angle-domain CIGs can only be produced by wavefield decomposition at depth.

Migration methods based on DSR operators have been applied to 2-D prestack migration for long time (Claerbout, 1985). However, the direct application of DSR migration methods to 3-D prestack data have been prevented by the tremendous computational cost. Only recently computationally efficient methods to continue 3-D prestack data have been presented (Biondi and Palacharla, 1996; Mosher et al., 1997). In particular, common-azimuth migration is an attractive alternative to Kirchhoff migration for sub-salt imaging because of its robustness with respect to the complex multipathing that is induced by salt bodies.

We will explain how some widely used common image gathers can contain reflector ambiguity due to multipathing and why angle-domain common image gathers will not. Then we will demonstrate the construction of angle-domain common image gathers from wave equation downward continued data for use in velocity analysis and amplitude-versus-reflection angle analysis. Finally, we apply this method to the Marmousi model.

KINEMATICS OF MULTIARRIVALS IN THE SHOT AND OFFSET DOMAINS

The kinematics of shot domain common image gathers and offset domain image gathers are well understood in constant velocity and $v(z)$ media. Difficulties arise when we begin considering complex subsurfaces with rapid lateral velocity variations. Even in 2-D it is easy to construct a model for which an individual common shot gather or common offset gather can contain two events from two points in the subsurface that arrive at the same time and are indistinguishable. Let us investigate particular cases of these occurrences.

An individual common shot gather is parameterized by the receiver location r . If two raypaths between the same source and receiver exist such that they have the same receiver horizontal slowness p_r and the two-way traveltimes along each is the same, it is impossible to distinguish between the two reflector locations (Nolan and Symes, 1996). Figure 1 shows a very simple case of this.

An individual common offset gather is parameterized by midpoint m . Suppose that the horizontal midpoint slowness $p_m = p_s + p_r$ is the same for two raypaths with the same traveltimes. Once again, the two raypaths represent the same event, and the location of the reflector causing this event is completely ambiguous. Figure 2 shows a case where $p_{s1} = -p_{r1}$ and $p_{s2} = -p_{r2}$ so that $p_{m1} = p_{m2} = 0$.

By shooting a fan of rays from both \mathbf{x} locations in Figure 2, it is possible to obtain travel-time curves for a common offset. These curves are shown in Figure 3. The travel-time curve for the diffractor inside the anomaly is nicely hyperbolic. The curve for the deeper diffractor has

Figure 1: Individual shot gather: the circular lens is a low velocity anomaly so the traveltimes and p_r are identical. `marie1-shot1` [NR]

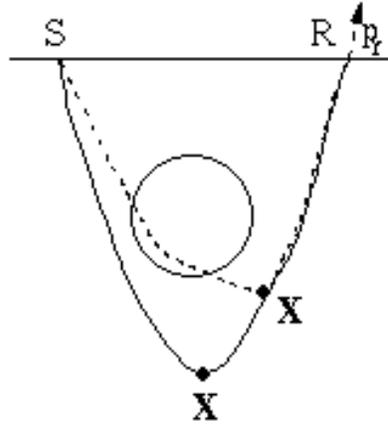


Figure 2: Individual offset gather: the circular lens is a low velocity anomaly so the traveltimes are identical and the midpoint slownesses are equal. `marie1-offset1` [NR]

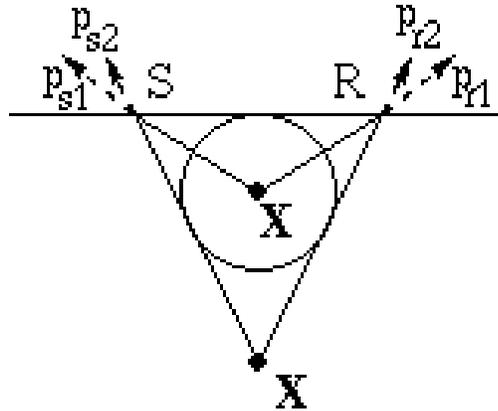
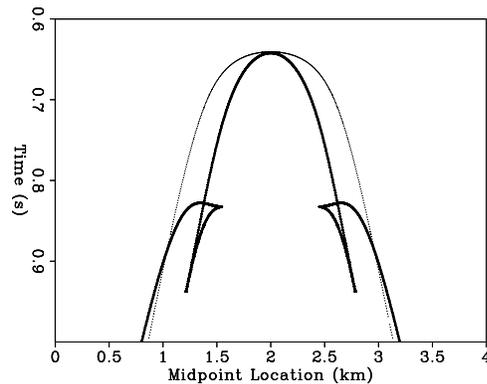


Figure 3: Traveltime curves for the model in Figure 2. The thick curve is for the deeper diffractor, the thin curve is for the diffractor in the center of the anomaly. `marie1-dualcurve1` [ER]



large symmetrical triplications. At this offset, the traveltimes for both curves are identical for the midpoint directly above the diffractors. From the geometry of the model, it is clear that the midpoint slownesses are the same, therefore the events in the data will be indistinguishable.

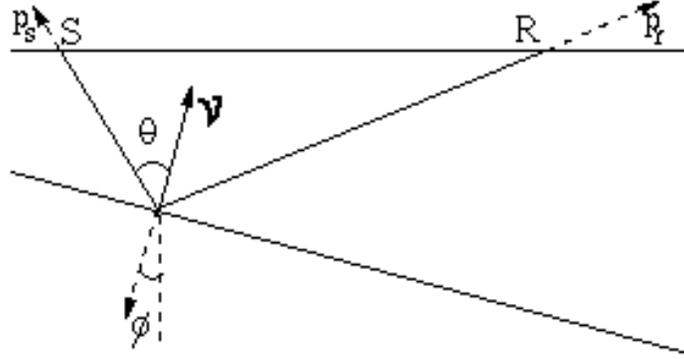


Figure 4: Simple reflection: the vector ν is normal to the reflector, θ is one-half the reflection angle, and ϕ is related to the dip angle. marie1-rangle2 [NR]

KINEMATICS OF MULTIARRIVALS IN THE ANGLE DOMAIN

We confine our discussion to the 2D case; the 3D case is similar, provided that complete surface coverage is available. Specular reflection connects a *reflector element*, consisting of a subsurface position \mathbf{x} containing the midpoint location \mathbf{m} and the depth z and *dip vector* ν representing the reflector normal at that position with an *event element* s, p_x, r, p_r, t consisting of source and receiver positions s and r , source and receiver horizontal slownesses p_s and p_r , and two way time t . The connection *via* incident and reflected rays is depicted in Figure 4, which also shows the *opening angle* θ . Note that given \mathbf{x} , ν , and θ , the incident and reflected rays and the event element s, p_s, r, p_r, t are completely determined: therefore the latter are functions of \mathbf{x} , ν , and θ . The angle transform of a data set $\{d(s, r, t)\}$ is

$$a(\mathbf{x}, \theta) = \int d\nu w(\nu, \theta, \mathbf{x}) d(s, r, t)$$

where $w(\nu, \theta, \mathbf{x})$ is an appropriate weighting function and s, r, t are also functions of ν, θ, \mathbf{x} .

The principle of stationary phase shows that an event in a single angle panel, i.e. a position \mathbf{y} and an *angle domain dip vector* η , arise when incident and reflected rays meet at \mathbf{y} and are bisected by η ; these rays determine once again an event element s, p_s, r, p_r, t , and *this event must have been present in the data for the event in question to be present in the angle domain*. Of course the event element s, p_s, r, p_r, t completely determines the rays in the subsurface carrying the energy of the event. We assume the *Traveltime Injectivity Condition* (ten Kroode et al., 1999): a pair of rays and a total (two-way) time determines at most one reflector element.

In that case, the event in the angle domain is compatible with at most one reflector element (\mathbf{x}, ν) .

Note the contrast with the constant offset domain as described in the preceding section where an event element in the data could correspond kinematically to more than one reflector element.

The velocity field used to generate the rays used in the formation of the angle transform does not necessarily need to be the same as the velocity field which gave rise to the moveout in the data - which is fortunate, as we don't know the latter at the outset of the migration/velocity analysis process, and have only an approximation of it at the end! When the two velocity fields are different, the angle transform events will not necessarily match the reflectors in the Earth: the two will differ by a residual migration. When the two velocity fields are the same, the image is perfect, i.e. $(\mathbf{x}, \nu) = (\mathbf{y}, \eta)$.

ANGLE-DOMAIN CIG BY WAVE-EQUATION MIGRATION

In the previous sections we discussed the advantages of angle-domain CIGs over offset-domain CIGs when a complex velocity function induces multipathing and event triplication. In this section we show how to extract angle-domain CIGs from downward-continued prestack data.

Recorded 3-D seismic data can be organized as a function of midpoint coordinates (\mathbf{m}) and offset coordinates (\mathbf{h}) . Prestack data are efficiently downward continued using the DSR equation in the frequency (ω) domain. Furthermore, since we either use 2-D downward continuation or 3-D common-azimuth downward continuation, the offset space is restricted to the in-line offset h_x , and thus we express the recorded wavefield as $P(\omega, \mathbf{m}, h_x; z = 0)$, where z is depth and $z=0$ indicates data recorded at the surface.

The prestack wavefield at depth is obtained by downward continuing the recorded data using the DSR, and is imaged by extracting the values at zero time

$$P(\omega, \mathbf{m}, h_x; z = 0) \xrightarrow{\text{DSR}} P(\omega, \mathbf{m}, h_x; z) \quad (1)$$

$$P(\omega, \mathbf{m}, h_x; z) \xrightarrow{\text{Imaging}} P(t = 0, \mathbf{m}, h_x; z) \quad (2)$$

The downward-continuation process focuses the wavefield towards zero offset (left panel in Figure 5) and if the continuation velocity is correct, a migrated image can be obtained by extracting the value of the wavefield at zero offset. However, the zero-offset wavefield has limited diagnostic information for velocity updating, and no information on the amplitude of the reflections versus reflection angle (AVA). We therefore perform a slant stack along the offset axis before imaging and obtain an image as a function of the offset ray parameter p_{hx} , as

$$P(\omega, \mathbf{m}, h_x; z = 0) \xrightarrow{\text{DSR}} P(\omega, \mathbf{m}, h_x; z) \quad (3)$$

$$P(\omega, \mathbf{m}, h_x; z) \xrightarrow{\text{Slant stack}} P(\tau, \mathbf{m}, p_{hx}; z) \quad (4)$$

$$P(\tau, \mathbf{m}, p_{hx}; z) \xrightarrow{\text{Imaging}} P(\tau = 0, \mathbf{m}, p_{hx}; z). \quad (5)$$

$$(6)$$

Angle-domain CIGs are subsets of $P(\tau = 0, \mathbf{m}, p_{hx}; z)$ at fixed midpoint location. The right panel in Figure 5 shows the angle-domain CIG gather corresponding to the downward-continued offset gather shown in the left panel. Notice that because in downward-continued offset gathers the energy is concentrated around zero offset, the slant stack decomposition does not suffer from the usual artifacts caused by the boundary conditions.

Strictly speaking, the CIG gathers obtained by the proposed procedure are function of the offset ray parameters p_{hx} and not of the aperture angle θ . However, p_{hx} is linked to θ by the following simple trigonometric relationship

$$\frac{\partial t}{\partial h} = p_{hx} = \frac{2 \sin \theta \cos \phi}{V(z, \mathbf{m})}, \quad (7)$$

where ϕ is the geological dip along the in-line direction and $V(z, \mathbf{m})$ is the velocity function.

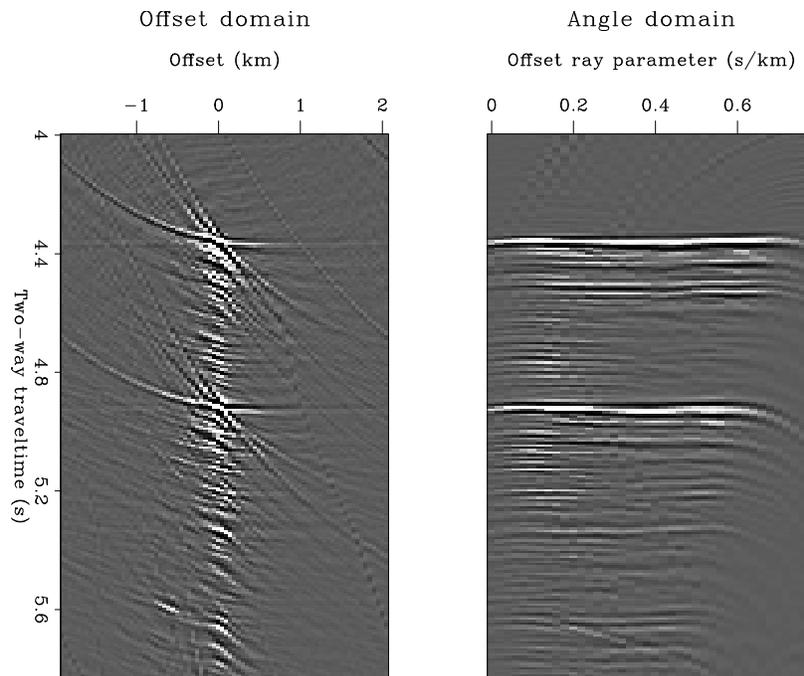


Figure 5: Left: Offset panel after downward continuation. Right: Angle-domain CIG
marie1-AVO-hydrate-off-angle [CR]

Angle-domain CIG and velocity

Angle-domain CIGs can be used to update the velocity function after migration similarly to the way that offset-domain CIGs are currently used (Brandsberg-Dahl et al., 1999) or for wave-equation Migration Velocity Analysis (Biondi and Sava, 1999). As for offset-domain CIGs, if

the velocity function is correct the reflections are aligned along the angle axis. If the velocity function is too low the reflections will smile upward; if the velocity function is too high the reflections will frown downward. This behavior is demonstrated by the analysis of the gathers in Figure 6. The gathers were extracted from a 3-D prestack wavefield focused using common-azimuth downward continuation. The left gather was obtained using the correct velocity. The right gather was obtained using a low constant velocity. Figure 7 shows the inline migrated section that passes through the gather shown in Figure 6. Notice that the CIG gathers show only the first kilometer of the image.

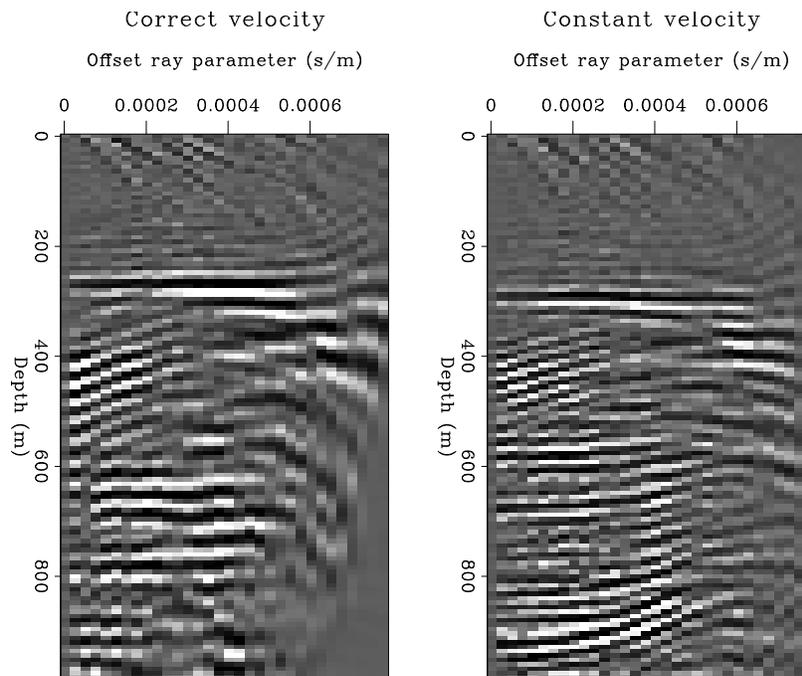


Figure 6: Left: Angle-domain CIG with correct velocity. Right: Angle-domain CIG with too low constant velocity. [marie1-AVO-jupiter-16801](#) [CR]

Angle-domain CIG and AVA

Angle-domain CIG can also be used to analyze the reflectivity as a function of the reflection angle to estimate rock and fluid properties in the subsurface. This potential use is illustrated by the gathers shown in Figure 8. The left panel shows an angle-domain CIG gather while the right panel shows the corresponding offset-domain CIG gather obtained by an amplitude preserving Kirchhoff migration (Ecker et al., 1996). The amplitude behavior as a function of offset ray parameter (left panel) is in qualitative agreement with the the amplitude behavior as a function of offset (right panel).

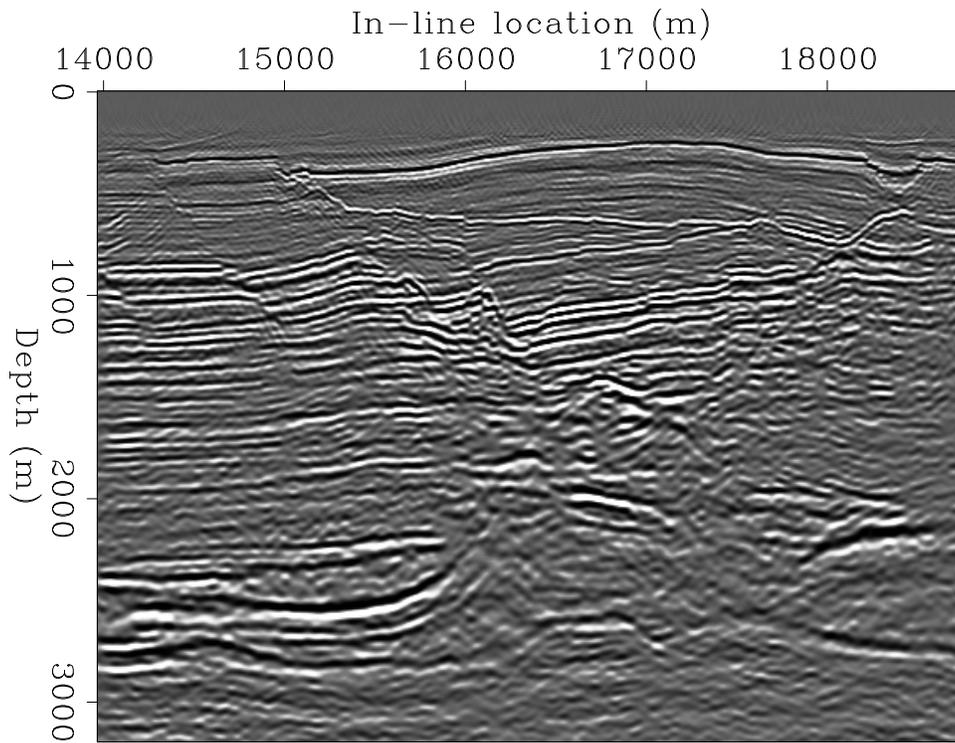


Figure 7: In-line section of 3-D migrated cube `marie1-Wave-jupiter-y20000` [CR]

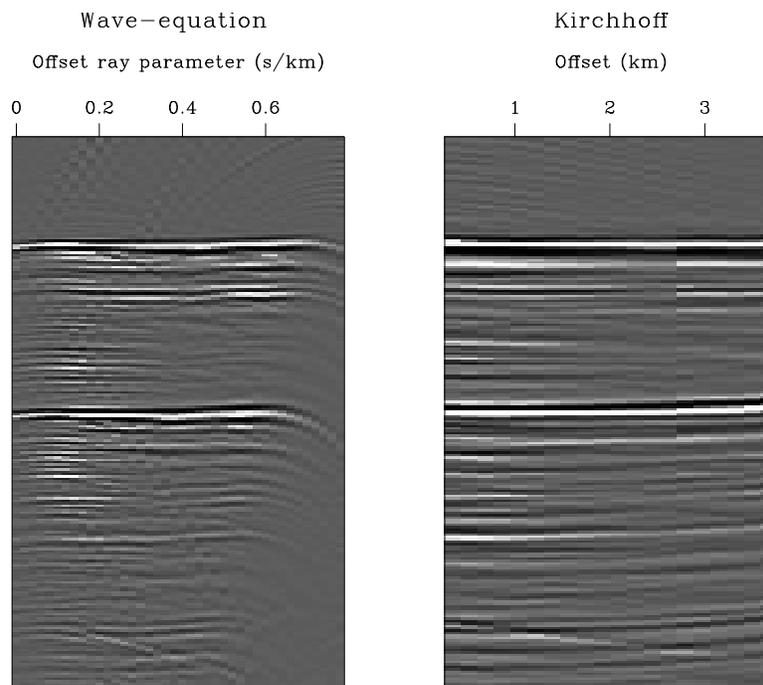


Figure 8: Left: Angle-domain CIG obtain by wave-equation migration. Right: Offset-domain CIG gather obtained by Kirchhoff migration `marie1-AVO-hydrate-angle-kir` [CR]

APPLICATION TO MARMOUSI MODEL

We applied the same method to the Marmousi model. Figure 9 shows a Kirchhoff migrated stack of the Marmousi data. The right panel of Figure 10 shows the offset-domain CIG taken from the surface location at 6336 meters. The left panel shows angle-domain CIG for the same surface location obtained by the wave-equation migration method described earlier. Although the angle-domain panel has more artifacts, it has better quality results. The angle-domain CIG contains higher frequencies. The reflectors at depths 1800 and 2500 meters are stronger and more continuous. The artifacts are caused by difficulties in the sampling and stack out of the image (Figure 11).

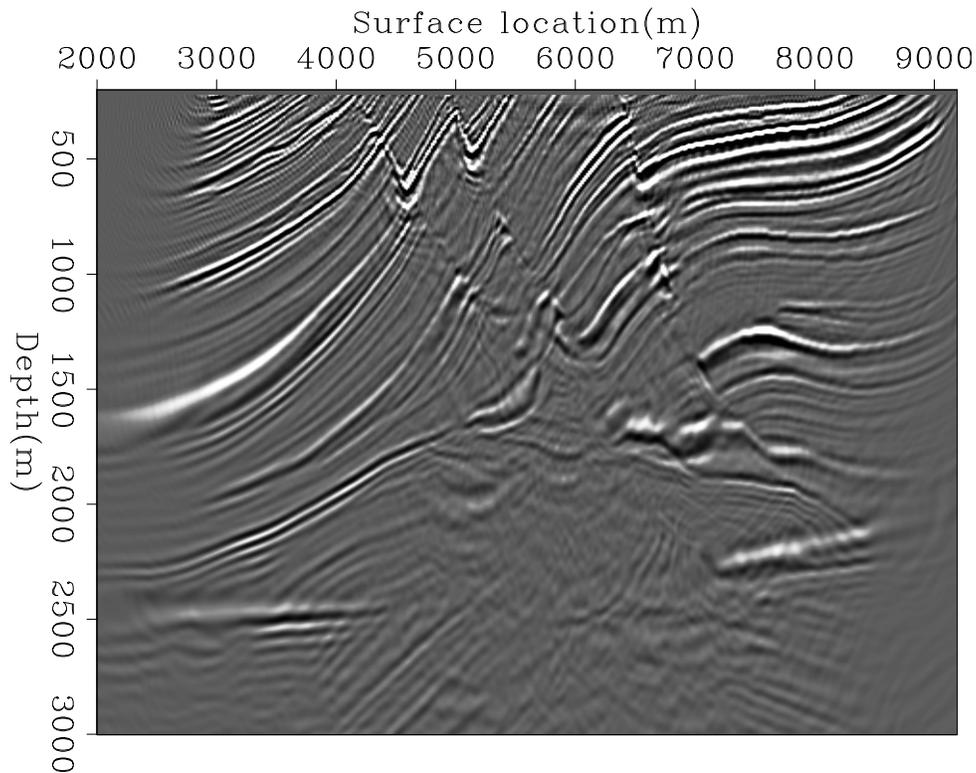


Figure 9: Kirchhoff migrated stack `marie1-stack.marm` [CR]

CONCLUSIONS

Angle-domain CIG gathers have attractive properties when a complex velocity model causes multipathing of reflected energy. They are free of the imaging artifacts caused by reflector-position ambiguity that degrade the image obtained from either shot gathers or common-offset gathers. We presented a simple method for extracting angle-domain CIGs from the prestack wavefield downward-continued using the wave equation. Our method produces high-quality CIG gathers that can be readily used for either velocity analysis or AVA analysis. Experimentation with the Marmousi dataset shows that our methods are valid.

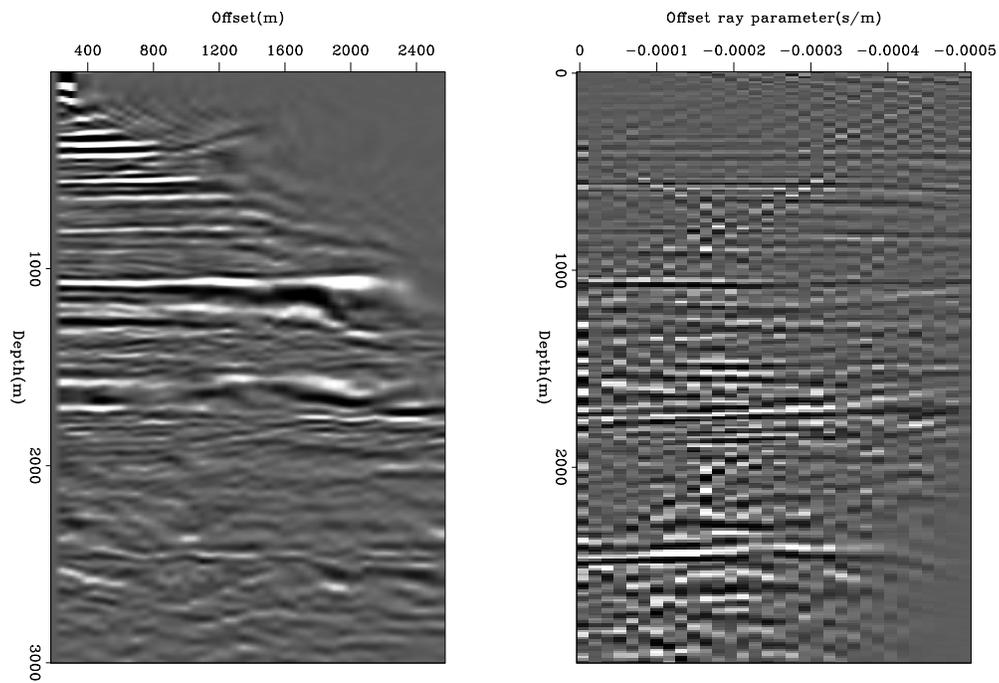


Figure 10: Left: angle-domain CIG. Right: offset-domain CIG. Both are from surface location 6336 meters. `marie1-marm.gathers` [CR]

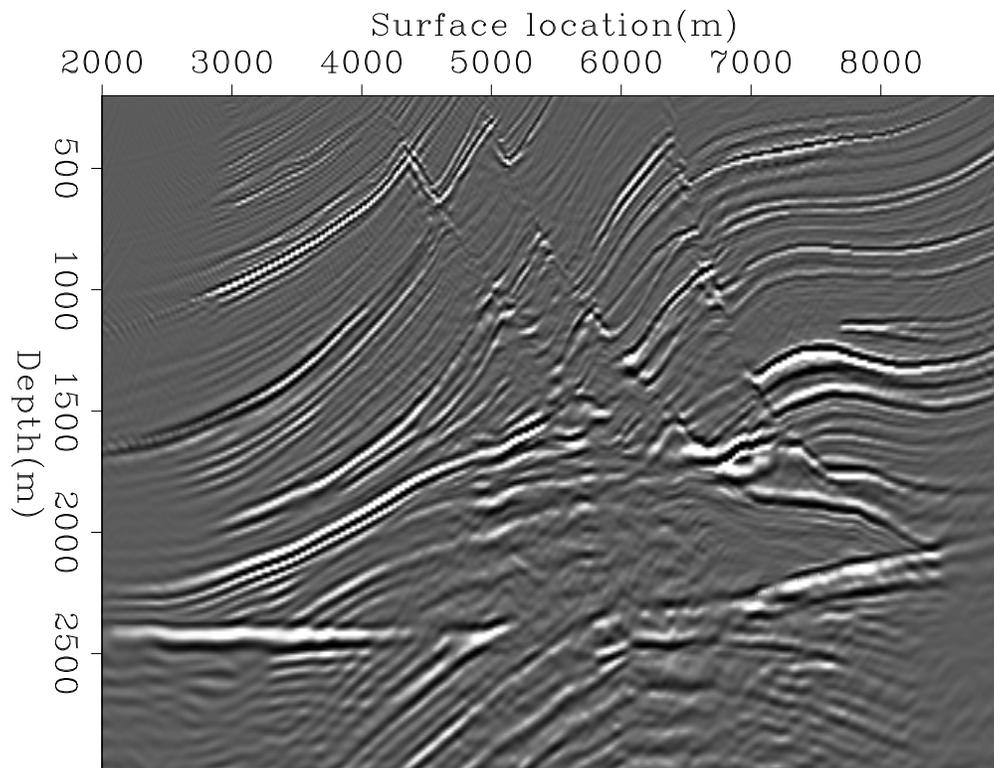


Figure 11: Wave-equation migrated stack `marie1-stack.WE` [CR]

FUTURE WORK

We are currently working on a way to generate angle-domain CIGs with Kirchhoff migration and multiple arrival traveltimes. The angle domain's ability to prevent multipathing should allow cleaner results in areas with complex subsurfaces.

We also intend to explore the use of angle-domain CIGs in shadow zones. Theoretically, reflection angle gathers have a more equalized amplitude versus angle panel than offset gathers have for their amplitude versus offset panels. Therefore, studying areas of low illumination in the angle domain should yield more continuous reflectors.

ACKNOWLEDGMENTS

Bob Clapp is largely responsible for the code used to generate many of the Kirchhoff migrated figures in this paper.

REFERENCES

- Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: *Geophysics*, **61**, 1822–1832.
- Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Brandsberg-Dahl, S., de Hoop, M. V., and Ursin, B., 1999, Sensitivity transform in the common scattering-angle/azimuth domain: EAGE 61st Annual Conference, Expanded Abstracts, submitted.
- Claerbout, J. F., 1985, *Imaging the Earth's Interior*: Blackwell Scientific Publications.
- Ecker, C., Dvorkin, J., and Nur, A., 1996, Sediments with gas hydrates: Internal structure from seismic avo: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1767–1770.
- Mosher, C. C., Foster, D. J., and Hassanzadeh, S., 1997, Common angle imaging with offset plane waves: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1379–1382.
- Nolan, C. J., and Symes, W. W., 1996, Imaging and coherency in complex structure: Proc. 66th Annual International Meeting, 359–363.
- Ottolini, R., and Claerbout, J. F., 1984, The migration of common-midpoint slant stacks: *Geophysics*, **49**, no. 03, 237–249.
- ten Kroode, A. P. E., Smit, D.-J., and Verdel, A. R., 1999, Linearized inverse scattering in the presence of caustics: *Wave Motion*.

Xu, S., Chauris, H., Lambare, G., and Noble, M., 1998, Common angle image gather - a strategy for imaging complex media: Depth Imaging of Reservoir Attributes, X012.

Subsalt imaging by common-azimuth migration

*Biondo Biondi*¹

ABSTRACT

The comparison of subsalt images obtained by common-azimuth migration and single-arrival Kirchhoff migration demonstrates the potential of wave-equation migration when the velocity model causes complex multipathing. Subsalt reflectors are better imaged and the typical Kirchhoff artifacts caused by severe multipathing disappear. A detailed analysis of common-azimuth images indicates that the results of common-azimuth imaging could be improved. It points to opportunities to improve the numerical implementation as well as the downward continuation method.

INTRODUCTION

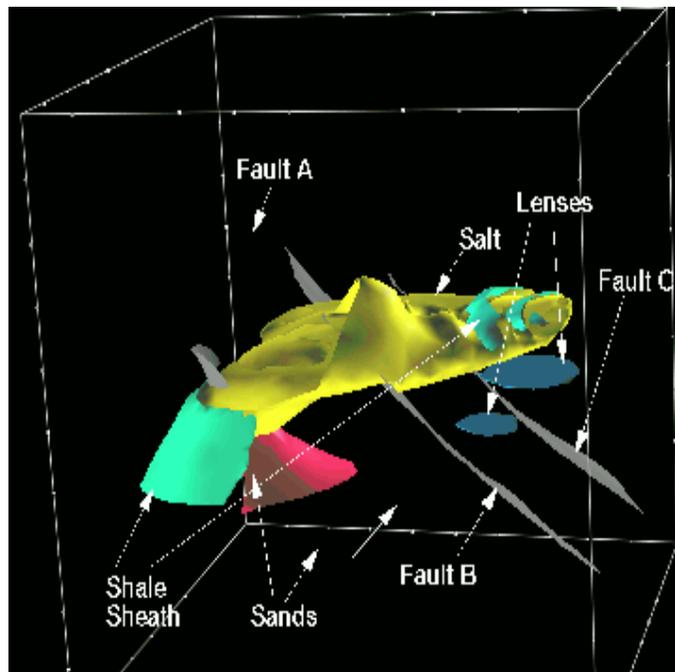
Kirchhoff migration methods often fail to produce satisfactory images of subsalt reflectors because they do not handle correctly multipathing of the reflected energy. When the wavefield is severely distorted by a salt body, or other complex velocity structure, the computation of the multivalued Green functions required by Kirchhoff methods is challenging. Further, even if we were able to compute the Green functions accurately and efficiently, the numerical integration of the wavefield over patchy and multivalued integration surfaces would be a difficult, and probably unreliable, task.

Wave-equation methods are an attractive and robust alternative to the complexities involved in extending Kirchhoff migration to handle correctly multipathing. However, full wave-equation 3-D prestack migration is still too computationally intensive to become a practical tool. Therefore, in the past few years I developed common-azimuth migration that is an approximation to full wave-equation 3-D prestack migration (Biondi and Palacharla, 1996; Biondi, 1997). It exploits the narrow-azimuth nature of marine data to reduce the computational cost by a large factor (20 to 50) with respect to full wave-equation 3-D prestack migration.

I applied common-azimuth migration to the narrow-azimuth subset of SEG-EAGE salt model (Aminzadeh et al., 1996) (known as C3 Narrow-Azimuth classic data set C3-NA (1997)) and compared the results with the results produced by a single-arrival Kirchhoff migration. The data were recorded on the realistic and complex salt-dome structure shown in Figure 1. In the subsalt areas, common-azimuth migration resolves the reflectors better than Kirchhoff

¹email: biondo@sep.stanford.edu

Figure 1: 3-D representation of the SEG-EAGE salt model. [biondo1-saltnew-color](#) [NR]



migration, and yields an image with much less artifacts and spurious reflectors. The run times of common-azimuth migration and Kirchhoff migration were roughly the same. These results confirm the potentiality of wave-equation migration and give new impulse to our efforts to develop a complete wave-equation imaging (migration and velocity analysis) procedure for both structural and stratigraphic imaging.

The lack of a simple and reliable method to extract prestack information from the results of wave-equation migration has been correctly perceived as a serious drawback (Etgen, 1998). It curtailed the usefulness of wave-equation migration for both velocity estimation and Amplitude Versus Angle (AVA) analysis. In Prucha et al. (1999), we show a very simple method to overcome this problem. We slant stack the downward continued wavefield at each depth level and produce high-quality Common Image Gathers (CIG) that display image amplitude as a function of the reflection angle. Sinha and Biondi (1999) discuss an example that compares wave-equation CIGs with the corresponding Kirchhoff CIGs for AVA analysis. Wave-equation CIGs can also be used for estimating migration velocity in a way similar to the common use of Kirchhoff-derived CIGs. Wave-equation CIGs are sensitive to migration velocity errors as the CIG obtained by migrating offset plane waves (Ottolini and Claerbout, 1984; Mosher et al., 1997). However, they are more accurate because they are based on a wavefield decomposition at depth and not at the surface.

Our ultimate goal is to build a complete and self-consistent wave-equation imaging procedure. An indispensable component of wave-equation imaging is a Wave-Equation Migration Velocity Analysis (WEMVA) method. Biondi and Sava (1999) present a WEMVA based on the linearization and inversion of downward continuation operators. WEMVA is more robust and stable than conventional ray-based MVAs because it can easily handle discontinuous velocity function and multipathing. The capability of performing both velocity analysis and

AVA analysis by wave-equation methods is attractive because it opens the possibility, though still far from being reality, to perform AVA analysis in more complex areas than is possible today. For AVA analysis in complex areas, wave-equation methods have a crucial advantage over asymptotic methods; they can model correctly the amplitudes variations related to the focusing and defocusing of bandlimited wavefields caused by velocity variations.

The results that are presented in this paper also show that Kirchhoff migration produced better images than common-azimuth migration in some areas of the model. At the moment of writing, it is not clear whether the degradation of the common-azimuth migration images was caused by the limitations of the numerical method employed to implement common-azimuth migration, or, by the inherent limitations of the common-azimuth migration method itself. Further investigation is needed. Common-azimuth migration has inherent limitations that are directly related to the approximations needed for its derivation and can be explained by a theoretical analysis (Biondi and Palacharla, 1996). Vaillant and Biondi (1999) discuss a promising method, dubbed narrow-azimuth migration, to overcome these limitations. The new method is more expensive than common-azimuth migration but it is still an order of magnitude cheaper than full wave-equation migration.

SEG/EAGE SALT DATA SET AND PREPROCESSING

The Salt Model C3-NA data set simulates a narrow-azimuth marine acquisition with 8 streamers recorded on an area about one quarter of the whole model. The maximum absolute offset is about 2,600 m and the maximum cross-line offset between the sources and the outer streamers is ± 140 m. The in-line direction is East-West, corresponding to the approximate North-South direction in the model as displayed in Figure 1. Notice that the figure displays the model rotated with respect to its "true" orientation; that is, the "true North" of the model points to the West of the figure.

The salt body in the model exhibits steep flanks near the crest and a rough surface on the top of the shelf. These characteristics cause severe distortions in the wavefield propagating through the salt. The reflectors below the salt area are thus poorly illuminated by data acquired with narrow-azimuth marine-like geometry. Consequently, the imaging of subsalt reflector is spotty even when using full-wave equation methods (Ober and Oldfield, 1999). Furthermore, deep dipping reflectors cannot be imaged because of the limited spatial extent of the data set. To reduce the computational cost of the modeling effort, the data were acquired on a dense grid only on a subset of the model. Good reference reflectors are: the bottom of the salt, the flat strong reflector at the bottom of the model (not marked in Figure 1), and the two sand lenses marked as "Lenses" in Figure 1. The bottom of the salt can be imaged pretty well in most of the areas, with the exception of the root proximities, where the interfaces are steeply dipping.

Before common-azimuth migration, the narrow-azimuth data were transformed to effective common-azimuth data by applying Azimuth Moveout (Biondi et al., 1998). The regularized common-azimuth data set was binned with a 20 meter CMP spacing in both the in-line and cross-line directions, and with 100 meter sampling along the in-line offset direction. With

100 meter offset spacing, the moveouts of the shallow events are aliased. However, because the dips along the offset can be safely assumed to be always positive, aliasing by a factor of two can be easily overcome by both wave-equation migration and Kirchhoff migration (?). The data were muted with a “deep” mute because the early arrivals are contaminated by all sorts of modeling noise. This mute affected the imaging of the shallow events. A more careful mute could accomplish both goals of noise removal and shallow events preservation. Kirchhoff migration is more flexible than common-azimuth migration with respect to the input-data geometry. Therefore, the original narrow-azimuth data were migrated by Kirchhoff migration.

MIGRATION RESULTS

Figure 2 shows a typical in-line section through the velocity model, taken at the constant cross-line coordinate $y=6,850$ m. Figure 3 shows the corresponding common-azimuth migration results. Almost perfect images are obtained for the reflectors in the sediment above the salt, the top and the bottom of the salt. The flat ‘basement’ is well imaged in some areas and not in others. The dipping reflectors in the subsalt are not easily distinguishable from the background noise, if they are present at all. The good results above the salt are to be expected from common-azimuth migration. Prestack time migration would have produced similarly good results. However, poststack migration after constant velocity DMO would have had trouble to image correctly the steep faults that terminate at the top of the salt because of NMO-velocity conflicts (Rietveld et al., 1997; Biondi, 1998).

The focus of this paper is on the subsalt region, and on the comparison with Kirchhoff migration results. Figure 4 shows the in-line section taken through the velocity cube at constant cross-line coordinate $y=9,820$ m. This section is interesting because it crosses both sand lenses in the subsalt. Further, between the lenses there is an anticlinal structure broken by converging normal faults that has some chances to be visible in the images because it is flattish. Figure 5 shows the subsalt images obtained by Kirchhoff migration (top) and common-azimuth migration (bottom). The common-azimuth image is superior to the Kirchhoff image in several ways. First, the common-azimuth image lacks the strong coherent artifacts that makes the Kirchhoff image difficult to interpret. These artifacts are caused by partially coherent stacking of multipathing events along wrong trajectories. They are typical of Kirchhoff subsalt images, and can be only partially removed by a “smart” selection of the Kirchhoff summation surfaces, such as the ones suggested by the most-energetic arrival or shortest-path criteria (Nichols et al., 1998). Second, both lenses are interpretable from the common-azimuth image while in the Kirchhoff image they are either lost in the noise (top lens) or completely missing (bottom lens). Third, both the bottom of the salt and the basement are more continuous in the common-azimuth image. On the other hand, the large fault visible on the left part of the section at (1,800-4,000 m) is not perfectly imaged by common-azimuth migration. I will analyze this problem in more detail at the end of this section with the help of cross-line sections (Figure 9, Figure 10 and Figure 11).

Figure 6 shows the cross-line section taken through the velocity cube at constant in-line coordinate $y=7,440$ m. This cross-line section passes through the two subsalt lenses as the in-

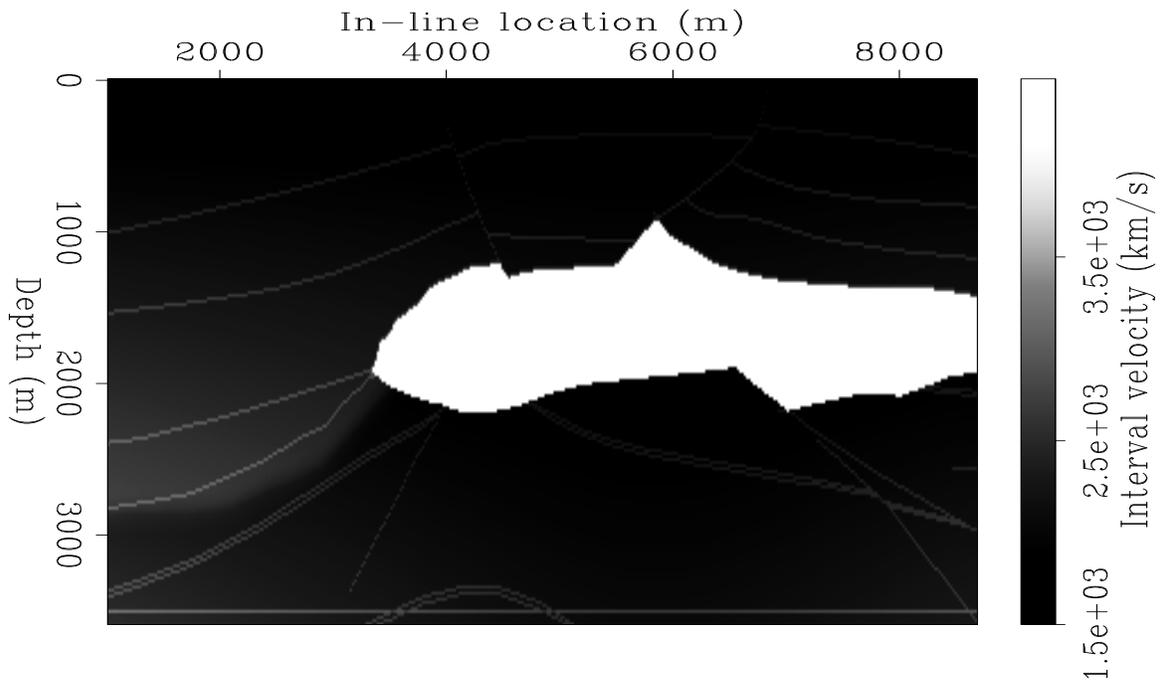


Figure 2: Velocity model at constant cross-line coordinate $y=6,850$ m.
`biondo1-Vel-salt-y6850` [CR]

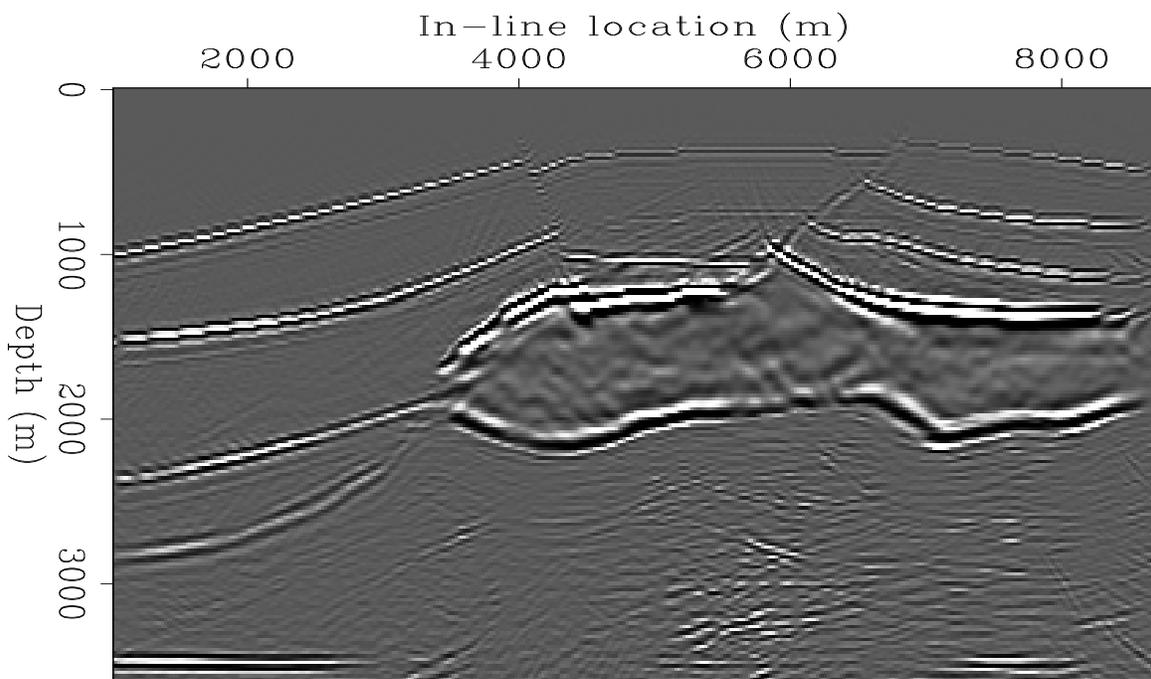


Figure 3: Common-azimuth migration at constant cross-line coordinate $y=6,850$ m.
`biondo1-Wave-salt-whole-y6850` [CR]

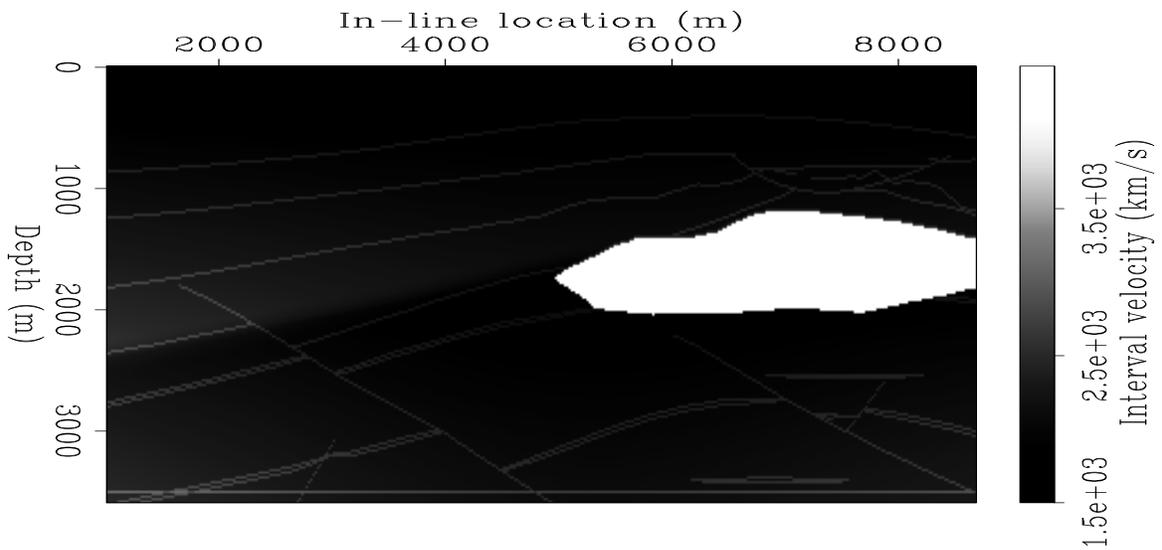


Figure 4: Velocity model at constant cross-line coordinate $y=9,820$ m. `biondo1-Vel-salt-y9820` [CR]

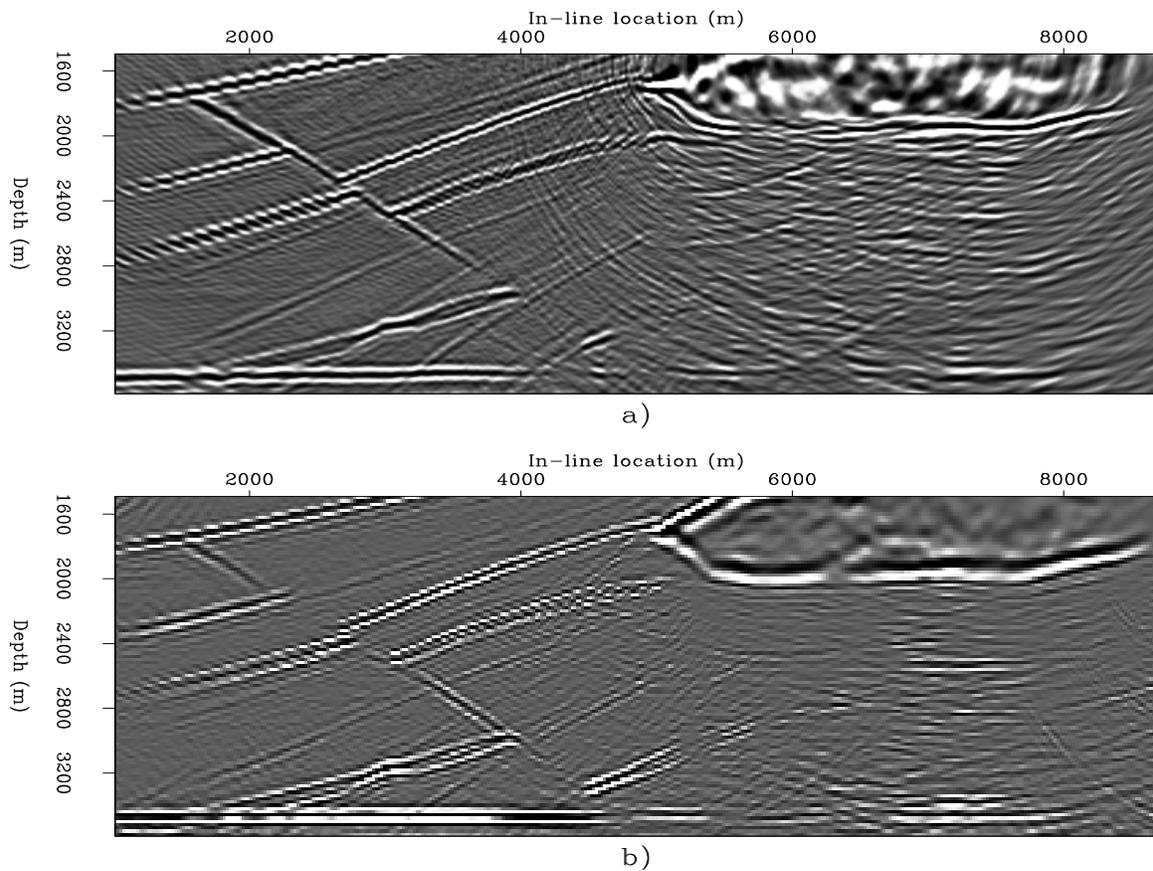


Figure 5: Kirchhoff migration (a) and common-azimuth migration (b) at constant cross-line coordinate $y=9,820$ m. Both sections are rendered using the same (98) percentile for clipping amplitudes. `biondo1-Both-salt-under-y9820` [CR]

line section shown in Figure 4. Figure 7 shows the corresponding migrated images; Kirchhoff migration on the top and common-azimuth migration on the bottom. As before, the two lenses are clearly interpretable in the common-azimuth image, whereas they are not in the Kirchhoff image. However, in this case the central portion of the salt bottom is not perfectly imaged in either of the two images. This area is right below the deep canyons in the salt body visible in Figure 6. The steep flanks of the canyons, and the large velocity contrast between the salt body and the soft sediments filling the canyons, cause a severe distortion of the reflected wavefield. The bottom of the salt and the reflectors below, including the basement, are thus poorly illuminated. In the column below the canyons, the Kirchhoff image shows strong artifacts that could be easily interpreted as reflections. The common-azimuth image is much cleaner, although without interpretable coherent events. The poor reflectors' illumination below the canyons can be analyzed further by looking at the Common Image Gathers (CIG) displayed in Figure 8. The gather on the left corresponds to a cross-line location right below the canyons; the one on the right is further toward the right. In both gathers, the images of the reflectors above the salt and the top-of-salt are well imaged and are aligned nicely along the offset ray parameter axis. In the gather on the right, the bottom of the salt, the shallower lens, the deeper lens, and the basement are also coherent and well aligned horizontally. But in the gather on the left, there is very little coherent energy below the salt.

Finally, I analyze the question of the poor imaging by common-azimuth migration of the fault shown in Figure 5. Figure 9 and Figure 10 show respectively the velocity model and the migrated images at constant in-line coordinate $x=2,560$ m. The fault under study is the fault on the right part of the sections. The deeper part of the fault is not illuminated by the data because of lack of spatial coverage, and thus is not imaged by either Kirchhoff migration or common-azimuth migration. The shallower part of the fault is well imaged by both migrations, but the middle part of the fault is well imaged by Kirchhoff migration and not by common-azimuth migration. The poor imaging seems to be correlated with the velocity inversion right above the fault visible in the velocity sections (Figure 4 and Figure 9). Both the geological dip and the local gradient of the velocity function are roughly oriented at an angle of 45 degrees with respect to the shooting direction. Therefore, they have a large component in the cross-line direction, creating the conditions under which the approximations inherent in common-azimuth migration are the worst (Biondi and Palacharla, 1996). On the other hand, the problem may be simply caused by the fact that I used too few reference velocities (three) when I downward continued the wavefield with an extended split-step method. This issue deserves more studies and to investigate it further. I am now developing a better common-azimuth continuation method based on Ristow's Fourier finite-difference methodology (Ristow and Ruhl, 1994).

Both Kirchhoff migration and common-azimuth migration have trouble to image the two deeper flattish reflectors at cross-line location of about 7,000 meters. The culprit seems to be again the sharp velocity contrast above the fault. However, the problem may be caused by the salt edge above the reflectors, not visible in these sections.

Figure 11 shows the CIG gathers taken at both problematic locations. The gather on the left shows some coherent energy for the poorly imaged reflections, though the energy is not perfectly aligned along the ray-parameter axis. The gather on the right shows no coherent energy corresponding to the poorly imaged fault, suggesting a worse imaging problem for this

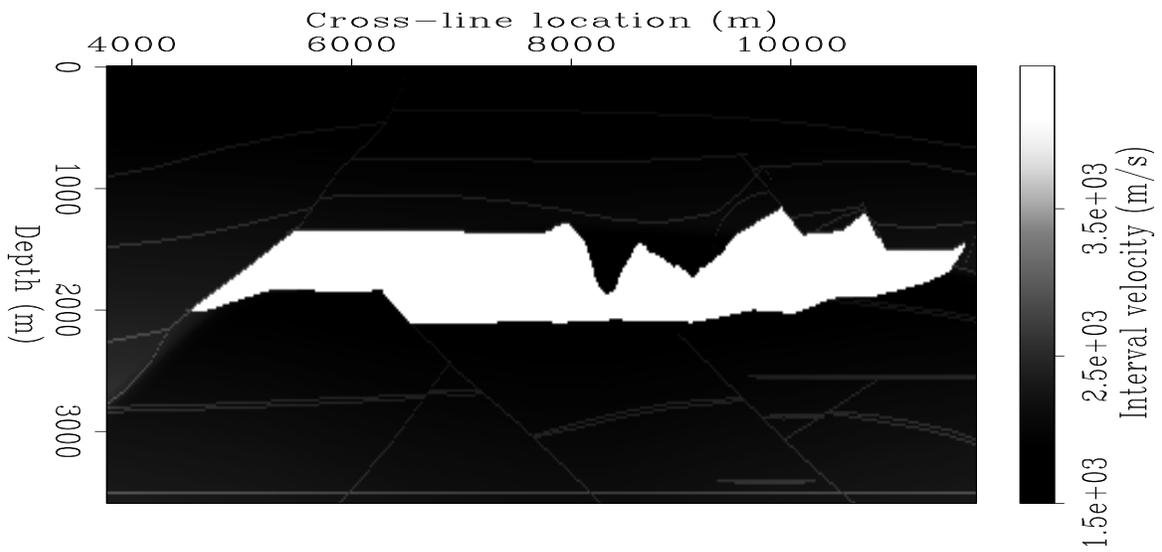


Figure 6: Velocity model at constant in-line coordinate $x=7,440$ m. `biondo1-Vel-salt-x7440` [CR]

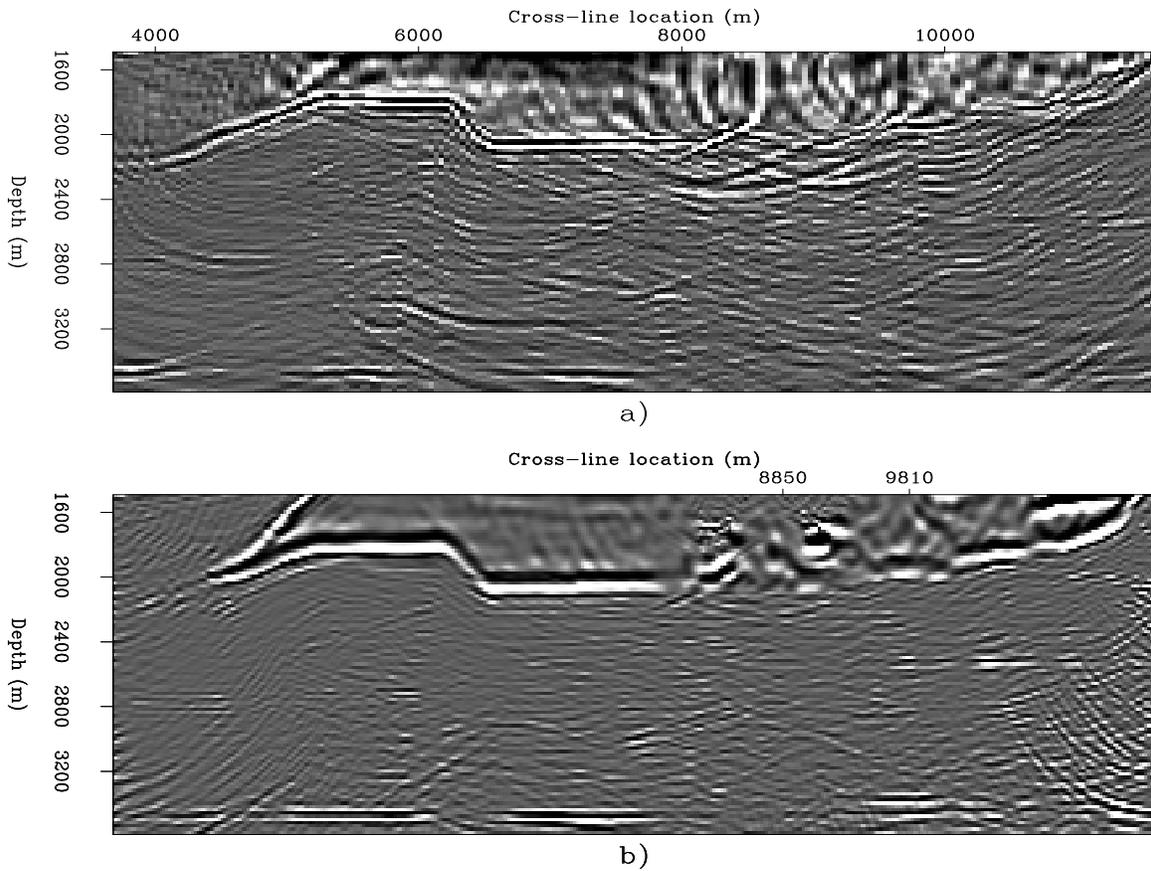


Figure 7: Kirchhoff migration (a) and common-azimuth migration (b) at constant in-line coordinate $x=7,440$ m. Both sections are rendered using the same (98) percentile for clipping amplitudes. `biondo1-Both-salt-under-x7440` [CR]

case.

CONCLUSIONS

Common-azimuth migration produced better results in the subsalt than a single-arrival Kirchhoff migration. The subsalt reflectors are much more interpretable in the common-azimuth images than in the Kirchhoff images, both because the images are devoid of the typical subsalt Kirchhoff artifacts and because the reflectors themselves are better imaged.

Although superior to Kirchhoff images, the common-azimuth images in the subsalt are far from perfect. Sub-optimal images are probably caused by a combination of poor reflectors' illumination and inaccuracies in the migration procedure. The shortcomings in the migration procedure are of two types: common-azimuth approximations and numerical approximations. We plan to address both types. We are developing a narrow-azimuth extension to common-azimuth migration (Vaillant and Biondi, 1999). And we plan to apply to common-azimuth downward continuation more accurate numerical methods that are based on a combination of the helix (Rickett et al., 1998) and Ristow's Fourier finite-difference methods (Ristow and Ruhl, 1994).

REFERENCES

- Aminzadeh, F., Burkhard, N., Long, J., Kunz, T., and Duclos, P., 1996, Three dimensional SEG/EAGE models - an update: *The Leading Edge*, **2**, 131–134.
- Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: *Geophysics*, **61**, 1822–1832.
- Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: *SEP-100*, 11–34.
- Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: *Geophysics*, **63**, no. 2, 574–588.
- Biondi, B., 1997, Azimuth moveout + common-azimuth migration: Cost-effective prestack depth imaging of marine data: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1375–1378.
- Biondi, B., 1998, Azimuth moveout vs. dip moveout in inhomogeneous media: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1740–1743.
- Etgen, J. T., 1998, $V(z)$ F-K prestack migration of common-offset common-azimuth data volumes, part I: theory: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1835–1838.
- Mosher, C. C., Foster, D. J., and Hassanzadeh, S., 1997, Common angle imaging with offset plane waves: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1379–1382.

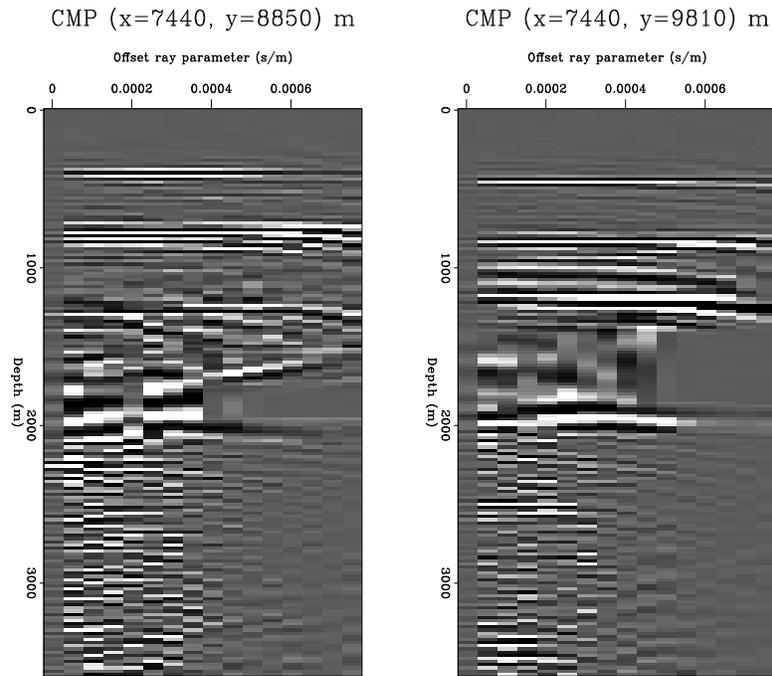


Figure 8: Angle-domain Common Image Gathers obtained by common-azimuth migration. The locations of these CIGs are marked on the in-line axis of the common-azimuth image shown in Figure 7 `biondo1-AVO-salt-8850-9810-x7440` [CR]

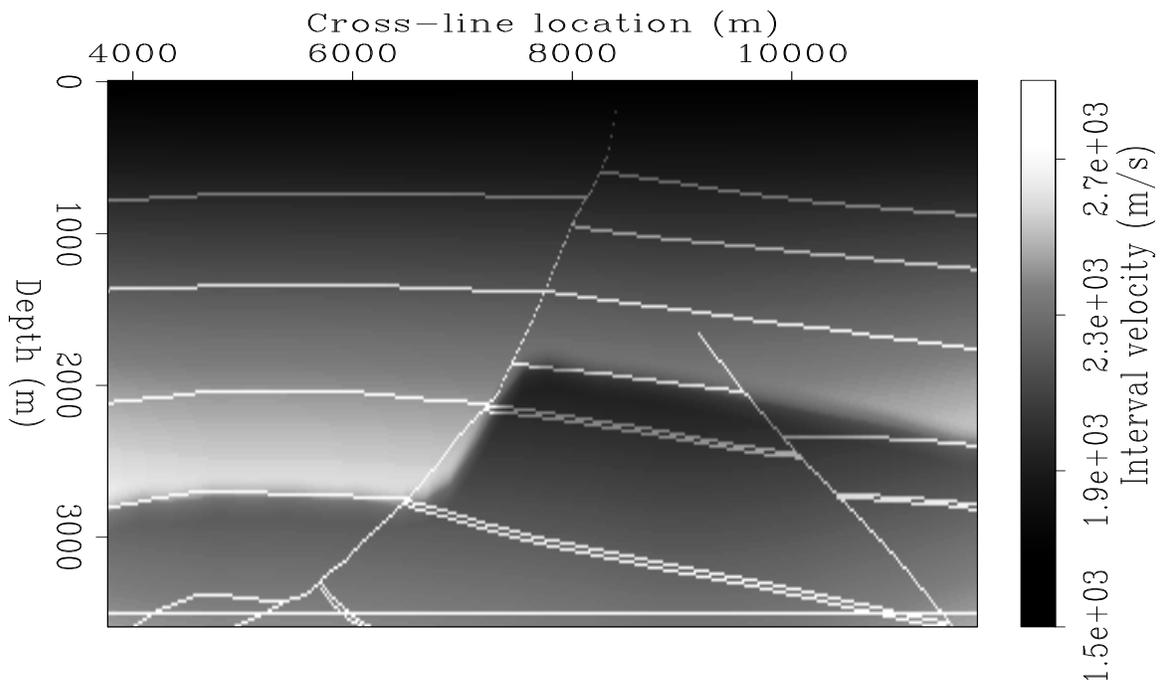


Figure 9: Velocity model at constant in-line coordinate $x=2,560$ m. `biondo1-Vel-salt-x2560` [CR]

- Nichols, D. E., Farmer, P., and Palacharla, G., 1998, Improving prestack imaging by using a new ray selection method: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1546–1549.
- Ober, C., and Oldfield, R., 1999, 3D Seismic Imaging of Complex Geologies: <http://www.cs.sandia.gov/ccober/seismic/Salvo.html>.
- Ottolini, R., and Claerbout, J. F., 1984, The migration of common-midpoint slant stacks: *Geophysics*, **49**, no. 03, 237–249.
- Prucha, M. L., Biondi, B. L., and Symes, W. W., 1999, Angle-domain common image gathers by wave-equation migration: *SEP-100*, 101–112.
- Rickett, J., Claerbout, J., and Fomel, S., 1998, Implicit 3-D depth migration by wavefield extrapolation with helical boundary conditions: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1124–1127.
- Rietveld, W. E. A., Marfurt, K. J., and Kommedal, J. H., 1997, The effect of 3-D prestack seismic migration on seismic attribute analysis: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1367–1370.
- Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration: *Geophysics*, **59**, no. 12, 1882–1893.
- SEG-EAGE, 1997, Salt Model Narrow-Azimuth Classic dataset (C3-NA): <http://archive.llnl.gov/SSD/classic/classicSalt.html#salt-c>.
- Sinha, M., and Biondi, B., 1999, Comparing Kirchhoff with wave equation migration on a hydrate region: *SEP-100*, 135–140.
- Vaillant, L., and Biondi, B., 1999, Extending common-azimuth migration: *SEP-100*, 125–134.

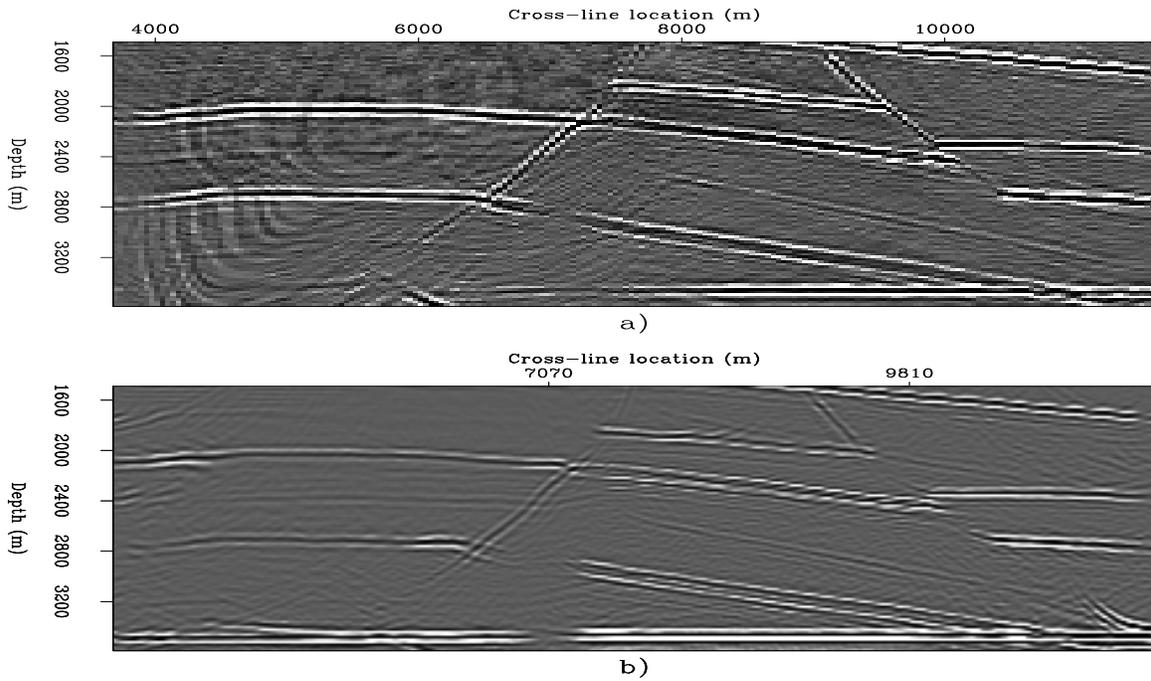


Figure 10: Kirchhoff migration (a) and common-azimuth migration (b) at constant in-line coordinate $x=2,560$ m. Both sections are rendered using the same (98) percentile for clipping amplitudes. `biondo1-Both-salt-under-x2560` [CR]

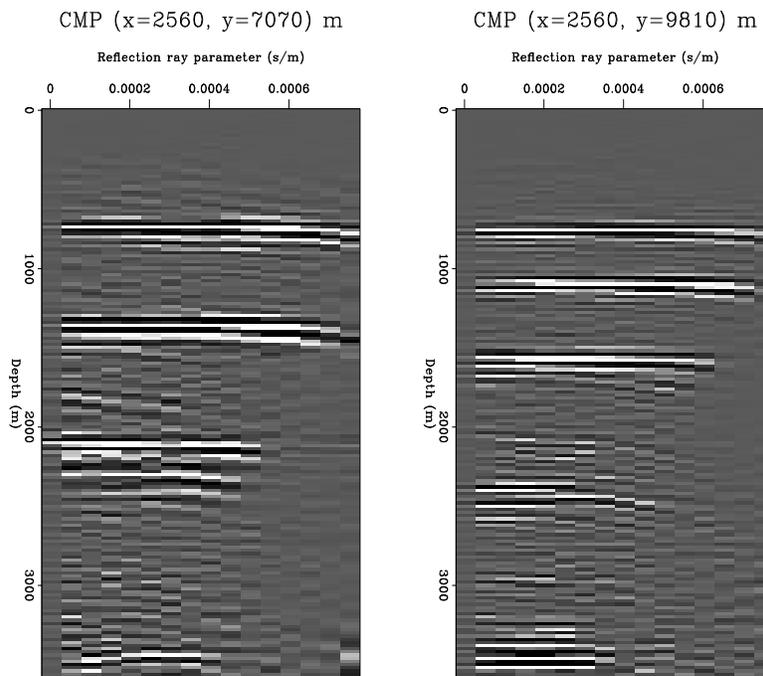


Figure 11: Angle-domain Common Image Gathers obtained by common-azimuth migration. The location of these CIGs are marked on the in-line axis of the common-azimuth image shown in Figure 10 `biondo1-AVO-salt-7070-9810-x2560` [CR]

Extending common-azimuth migration

Louis Vaillant and Biondo Biondi¹

ABSTRACT

We present a review of common-azimuth prestack depth migration theory and propose a new extension to the original method. In common-azimuth migration theory, source and receiver raypaths are constrained to lie on the same plane at each depth level. By using data with a broader range of cross-line offsets, we increase the number of raypaths examined and consider more information. Consequently, our extended common-azimuth migration is theoretically better able to model lateral velocity variations due to real 3-D structures and is more compatible with the standard marine acquisition geometry, in which cross-line offsets are concentrated in a narrow band. We first discuss the theory of the process, and then introduce computational issues leading to future implementation.

INTRODUCTION

Common-azimuth migration was first introduced by Biondi and Palacharla (1994). It is a 3-D depth imaging method based on a recursive downward continuation of prestack data, and allows a robust and accurate depth migration since it is derived directly from the full wave equation. As opposed to Kirchhoff methods, common-azimuth migration is not derived from asymptotic approximations, and thus it represents a potentially robust alternative to Kirchhoff methods for 3-D prestack migration.

The implementation of wave-equation methods presents several difficulties. Whereas Kirchhoff methods handle irregular geometries relatively easily, the downward-continuation process needs data with a regular geometry in order to correctly propagate the wavefield. Additionally, full volume 3-D wave equation migration still has a high computational cost. The challenge for such a recursive algorithm is to extrapolate the wavefield in a 5-D space: time, in-line and cross-line midpoint coordinates, and in-line and cross-line offsets, (t, m_x, m_y, h_x, h_y) . This computation is still beyond the reach of current computer technology. We can address this problem by only downward continuing common-azimuth data for which $h_y = 0$. In this way, the data space is reduced to 4-D. Common-azimuth data are obtained through azimuth moveout (AMO), which rotates and modifies the offset of 3-D prestack data (Biondi et al., 1998). AMO is used as a preprocessing step to regularize the 3-D data acquisition, and organize it in sets of constant cross-line offset. The first section discusses the original method, and the second one presents its new extension.

¹email: louis@sep.stanford.edu, biondo@sep.stanford.edu

COMMON-AZIMUTH PHASE-SHIFT MIGRATION

One way to reduce the computational cost of 3-D prestack depth migration is to use common-azimuth data, which is only a 4-D dataset. Common-azimuth migration is then derived as the solution of the one-way wave equation through a recursive downward-continuation operator. In the frequency-wavenumber domain, this operator can be expressed as a simple phase-shift applied to the wavefield. We use the notation indicated below. The vertical wavenumber k_z is given by a 3-D dispersion relation called the Double Square Root (DSR) equation (Claerbout, 1984):

$$DSR(\omega, \mathbf{k}_m, \mathbf{k}_h, z) = \sqrt{\frac{\omega^2}{v(\mathbf{r}, z)^2} - \frac{1}{4} [(k_{mx} + k_{hx})^2 + (k_{my} + k_{hy})^2]} + \sqrt{\frac{\omega^2}{v(\mathbf{s}, z)^2} - \frac{1}{4} [(k_{mx} - k_{hx})^2 + (k_{my} - k_{hy})^2]} \quad (1)$$

where \mathbf{k}_m is the midpoint wavenumber ($\mathbf{k}_m = k_{mx}\mathbf{x}_m + k_{my}\mathbf{y}_m$), \mathbf{k}_h the offset wavenumber ($\mathbf{k}_h = k_{hx}\mathbf{x}_h + k_{hy}\mathbf{y}_h$) and k_z the vertical wavenumber ($k_z = DSR(\omega, \mathbf{k}_m, \mathbf{k}_h, z)$), with \mathbf{x}_m , \mathbf{y}_m , \mathbf{x}_h and \mathbf{y}_h unit vectors of the four midpoint and offset axes. The propagation velocities $v(\mathbf{r}, z)$ and $v(\mathbf{s}, z)$ correspond respectively to the receiver and source locations.

The first square root in equation (1) downward-continues the receiver wavefield, whereas the second one downward-continues the source wavefield. In the algorithm developed by Biondi and Palacharla (1996), the data at the new depth level D_{z+dz} are obtained from common-azimuth data D_z by the following integration:

$$\begin{aligned} D_{z+dz}(\omega, \mathbf{k}_m, k_{hx}, h_y = 0) &= \int_{-\infty}^{+\infty} D_z(\omega, \mathbf{k}_m, k_{hx}, h_y = 0) e^{-ik_z dz} dk_{hy} \\ &= D_z(\omega, \mathbf{k}_m, k_{hx}, h_y = 0) \int_{-\infty}^{+\infty} e^{-ik_z dz} dk_{hy} \\ &= D_z(\omega, \mathbf{k}_m, k_{hx}, h_y = 0) \times \text{Down}(\omega, \mathbf{k}_m, k_{hx}, z) \end{aligned} \quad (2)$$

In practice, we use the stationary-phase approximation to compute this integral (see Appendix A). Popovici (1995) finds a very similar expression for the kernel of migration to zero-offset (MZO) in 2-D and asserts that the stationary-phase approximation avoids the very high computational cost of a numerical evaluation of the integral.

In our case, the stationary phase approximation of the common-azimuth downward-continuation operator for data where $h_y = 0$ can be written as

$$\text{Down}(\omega, \mathbf{k}_m, k_{hx}, z) = \sqrt{\frac{2\pi}{\phi''(\hat{k}_{hy}^0)}} e^{-i\phi(\hat{k}_{hy}^0) + i\frac{\pi}{4}} \quad (3)$$

with the phase $\phi = DSR(\omega, \mathbf{k}_m, \mathbf{k}_h, z)dz$.

For data evaluated only at the origin of the cross-line offset axis ($h_y = 0$), Biondi and

Palacharla (1996) derived an analytical expression for the stationary point:

$$\hat{k}_{hy}^0(z) = k_{my} \frac{\sqrt{\frac{1}{v(\mathbf{r},z)^2} - \frac{1}{4\omega^2}(k_{mx} + k_{hx})^2} - \sqrt{\frac{1}{v(\mathbf{s},z)^2} - \frac{1}{4\omega^2}(k_{mx} - k_{hx})^2}}{\sqrt{\frac{1}{v(\mathbf{r},z)^2} - \frac{1}{4\omega^2}(k_{mx} + k_{hx})^2} + \sqrt{\frac{1}{v(\mathbf{s},z)^2} - \frac{1}{4\omega^2}(k_{mx} - k_{hx})^2}} \quad (4)$$

We have implemented the kernel of downward-continuation and phase-shift migration with the preceding theory. Simulations performed with the standard common-azimuth migration code have already proven its efficiency in imaging complex media (Biondi, 1999). As shown in Figures 1 and 2, the common-azimuth technique provides accurate subsalt images. Still, on the left side of Figure 1(b), imaging the major fault has obviously met with some difficulties, which can be attributed to rapid lateral variations of the velocity model. As seen in Figures 2(a) and 2(b), the same fault, now on the right side, is also dipping in the cross-line direction. Extending the common-azimuth technique may help improve the imaging of these lateral velocity variations.

EXTENDING CROSS-LINE OFFSET RANGE

Marine data usually contain in-line offsets up to 3-5km, and only show a narrow range of offsets in the cross-line direction. For standard common-azimuth migration, AMO is used to rotate the data to zero cross-line offset. What we do now is use AMO to regularize the acquisition but organize data along several near-zero cross-line offsets. Strictly speaking, the data are not “common-azimuth”, but “quasi-common-azimuth” or “narrow-azimuth” since we have only a narrow range of cross-line offsets.

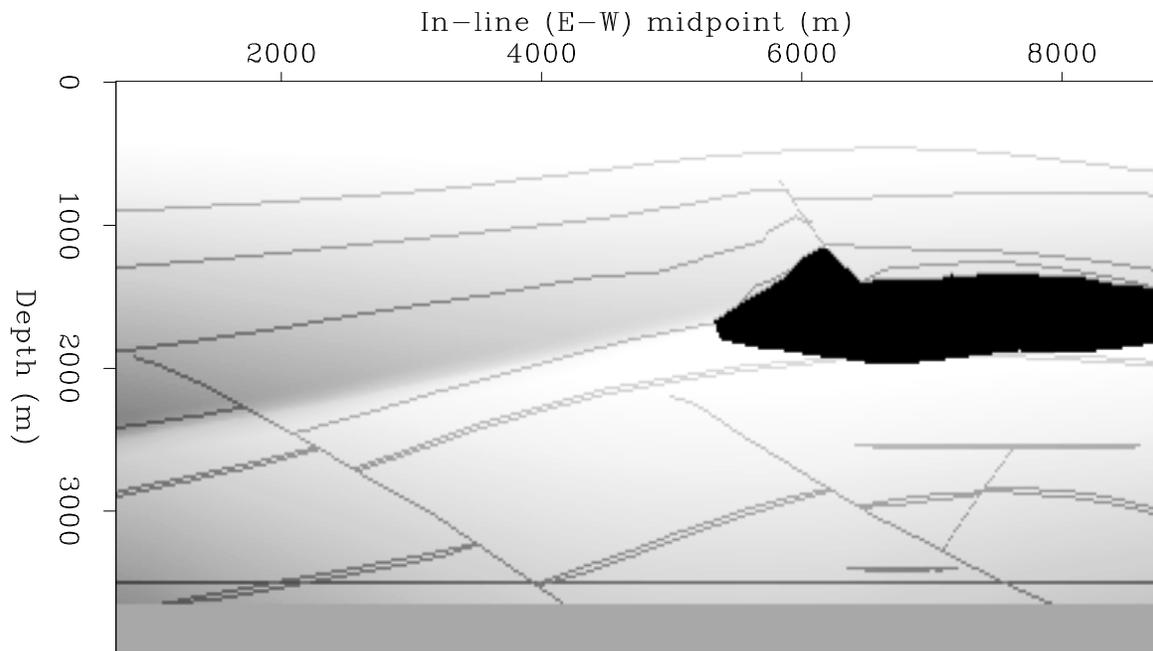
These narrow-azimuth data $D_z(\omega, \mathbf{k}_m, k_{hx}, h_y)$ are downward-continued as before, with an important difference in the phase of the integrand:

$$\begin{aligned} D_{z+dz}(\omega, \mathbf{k}_m, k_{hx}, h_y) &= \int_{-\infty}^{+\infty} D_z(\omega, \mathbf{k}_m, k_{hx}, h_y) e^{-ik_z dz} e^{-ik_{hy} h_y} dk_{hy} \\ &= D_z(\omega, \mathbf{k}_m, k_{hx}, h_y) \int_{-\infty}^{+\infty} e^{-ik_z dz} e^{-ik_{hy} h_y} dk_{hy} \\ &= D_z(\omega, \mathbf{k}_m, k_{hx}, h_y) \times \text{Down}(\omega, \mathbf{k}_m, k_{hx}, z) \end{aligned} \quad (5)$$

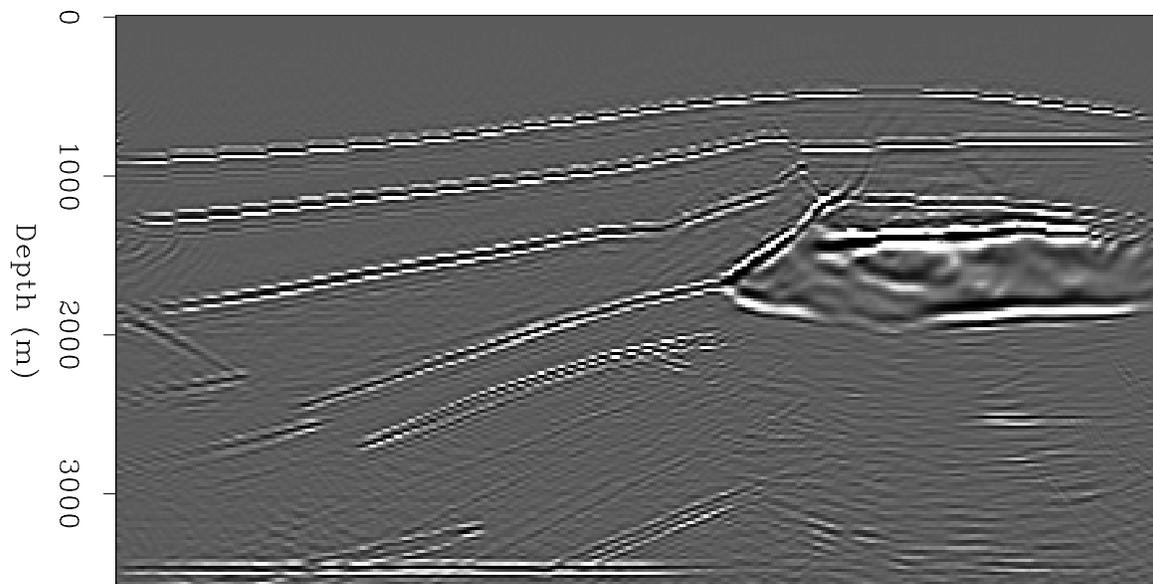
with

$$\begin{aligned} \text{Down}(\omega, \mathbf{k}_m, k_{hx}, z) &= \sqrt{\frac{2\pi}{\phi''(\hat{k}_{hy})}} e^{-i\phi(\hat{k}_{hy}) + i\frac{\pi}{4}} \\ \phi &= DSR(\omega, \mathbf{k}_m, \mathbf{k}_h, z) dz + k_{hy} h_y \end{aligned} \quad (6)$$

In 2-D, for the case of a phase-shift migration in which the equations are similar to ours, Alkhalifah (1997) found that solving for the minimum of the phase ϕ involves a sixth-order polynomial, making it difficult to find an analytical expression of the stationary path. Thus, since we are unable to use such an analytical expression for non-zero h_y , we instead use only

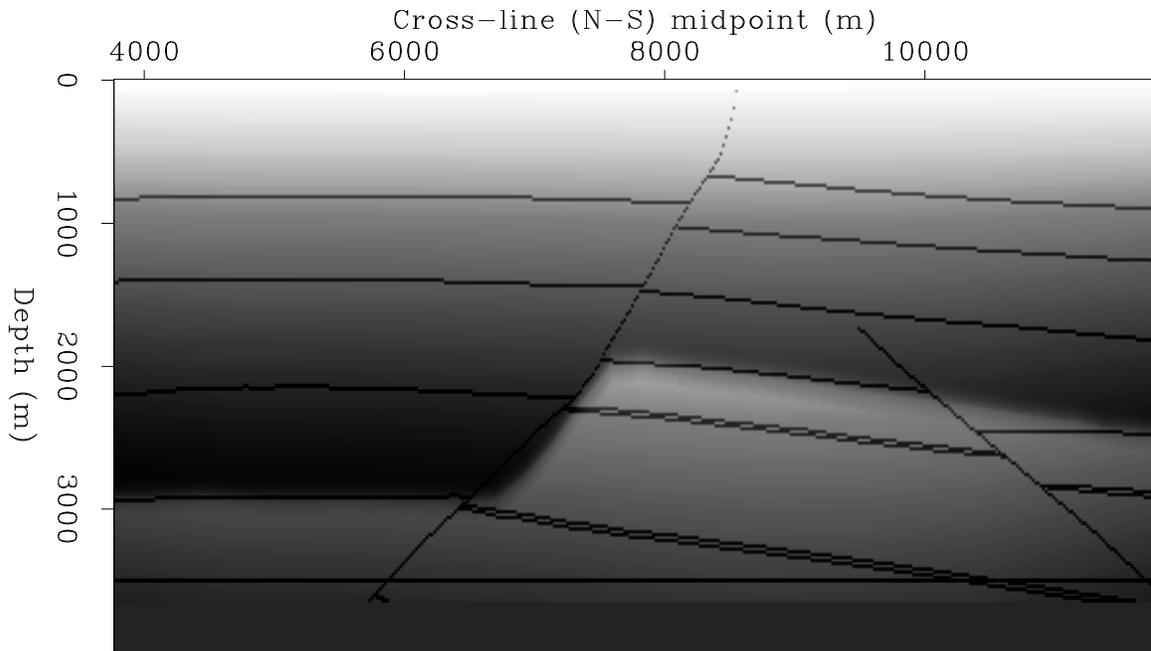


(a)

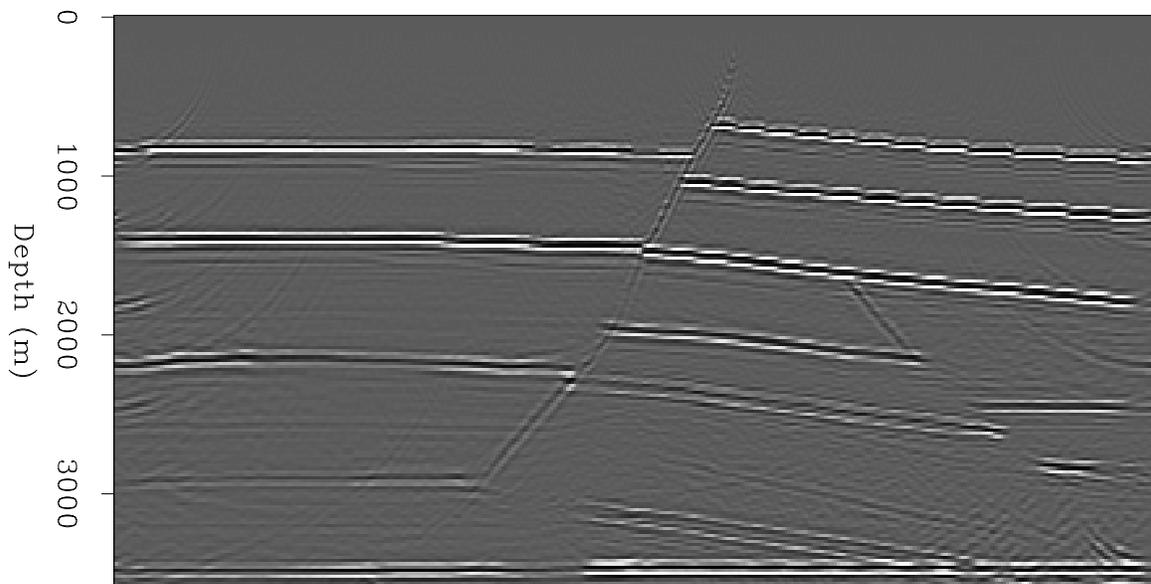


(b)

Figure 1: Common-azimuth prestack depth migration of data from the SEG-EAGE salt model. In-line section at CMPy=10400 (m). The top plot (a) represents the exact velocity model and the bottom plot (b) represents the migration result. [louis1-SEG-EAGE-in](#) [CR]



(a)



(b)

Figure 2: Common-azimuth prestack depth migration of data from the SEG-EAGE salt model. Cross-line section at $CMP_x=2100$ (m). The top plot (a) represents the exact velocity model and the bottom plot (b) represents the migration result. [louis1-SEG-EAGE-cross](#) [CR]

a range of values of \hat{k}_{hy} centered around the stationary path \hat{k}_{hy}^0 , which is given by equation (4), at zero cross-line offset. At each \hat{k}_{hy} , a different phase shift is applied: the data $D_z(\omega, \mathbf{k}_m, k_{hx}, h_y)$ are transformed by modulation and FFT along the cross-line offset axis into $\tilde{D}_z(\omega, \mathbf{k}_m, k_{hx}, k_{hy})$, in which the range of the coordinate k_{hy} is centered around \hat{k}_{hy}^0 . Then we apply the phase shift, $\phi(\hat{k}_{hy})$, for each value of the set $\{\hat{k}_{hy} = \hat{k}_{hy}^0 \pm i_{k_{hy}} dk_{hy}\}$.

The reformulation of the problem has some interesting geometrical consequences. In the case of standard common-azimuth migration, the stationary path of phase ϕ derived in equation (4) is equivalent to the following relationship between the ray parameters (Biondi, 1998):

$$\frac{p_{sy}}{p_{sz}} = \frac{p_{ry}}{p_{rz}} \quad (7)$$

The subscript, s , refers to the source ray, and, r , to the receiver ray:

$$p_{sx} = \frac{k_{mx} - k_{hx}}{2\omega} \quad p_{rx} = \frac{k_{mx} + k_{hx}}{2\omega} \quad p_{sy} = \frac{k_{my} - k_{hy}}{2\omega} \quad p_{ry} = \frac{k_{my} + k_{hy}}{2\omega}$$

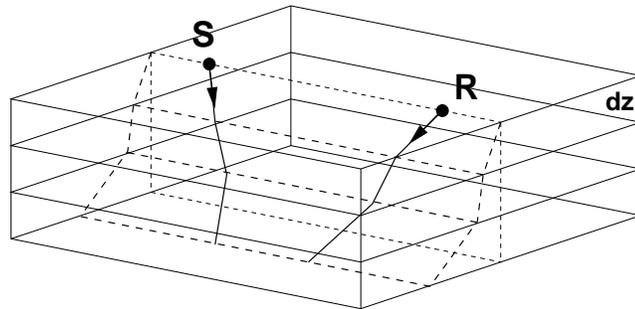
The ray parameters are linked by the relationship:

$$p_{sx}^2 + p_{sy}^2 + p_{sz}^2 = \frac{1}{v(\mathbf{s}, z)^2} \quad p_{rx}^2 + p_{ry}^2 + p_{rz}^2 = \frac{1}{v(\mathbf{r}, z)^2}$$

Equation (7) constrains the source ray and the receiver ray to lie on the same plane between

Figure 3: Schematic showing the ray geometry for common-azimuth downward continuation.

`louis1-ray-comaz` [NR]



depths z and $z + dz$, so that the sources and receivers at the new depth are aligned along the same azimuth as at the preceding depth (Figure 3). Assuming this constraint, common-azimuth migration is not strictly correct even in the simple case of a $v(z)$ medium (Figure 4). Since the projections of the source and receiver rays on a cross-line plane do not coincide unless the angles α and β are equal, azimuth is not conserved at each depth step of downward continuation. As a general consequence, local variations in velocity, which cause ray bending phenomena, introduce an error in the computation of the common-azimuth downward continuation. Biondi and Palacharla (1996) put forward that this error is small, but by using several discrete values of k_{hy} in a neighborhood of \hat{k}_{hy}^0 , we hope to improve the result.

The narrow-azimuth approach may improve the imaging process by better taking velocity variations into account (Figure 5). Lateral velocity variations, for example, can make rays bend in the cross-line direction, and even make rays start in different planes before they converge to a deeper reflection point. The bending due to real 3-D structures is implicitly eliminated when

Figure 4: Schematic showing the ray geometry for common-azimuth downward continuation in a simple $v(z)$ velocity model. The projections of both rays on a cross-line plane do not coincide due to non-conservation of azimuth at each depth level. [louis1-ray-vofz] [NR]

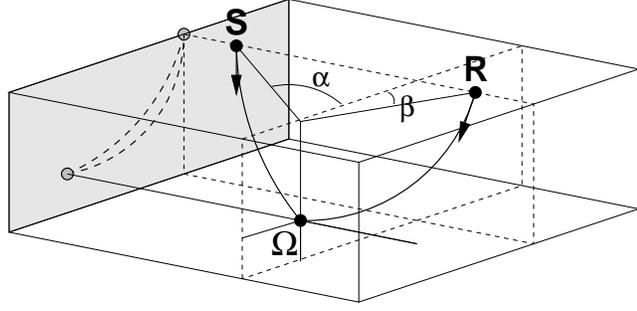
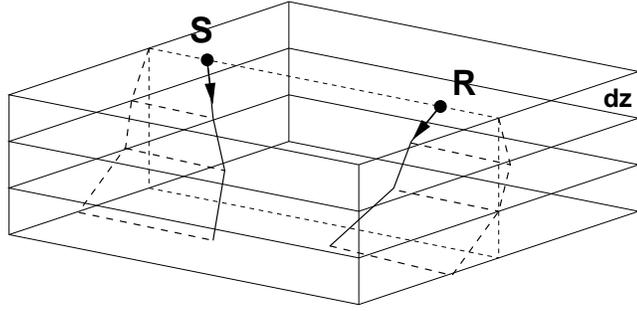


Figure 5: Schematic showing the ray geometry for narrow-azimuth downward continuation. A larger range of cross-line offsets will allow the modeling of lateral variations in the velocity model by taking into consideration rays not lying in the same slanted plane. [louis1-ray-extended] [NR]



strictly zero-azimuth data are downward continued, whereas we would like to keep it for an extension of common-azimuth migration.

Our approximation, using a neighborhood of \hat{k}_{hy}^0 , is valid if the stationary path, \hat{k}_{hy} , varies slowly with h_y around the analytical solution, \hat{k}_{hy}^0 . To estimate these variations, we can use the Implicit Function theorem: the equation $\frac{\partial \phi}{\partial k_{hy}} = 0$, which gives the stationary path \hat{k}_{hy} , also defines a relation of the general form $F(k_{hy}, h_y) = 0$. Thus, in a neighborhood of the solution \hat{k}_{hy} , we use a function $k_{hy}(h_y)$, whose variations are determined by the inverse of the phase second derivative:

$$\left(\frac{dk_{hy}}{dh_y} \right)_{\hat{k}_{hy}} = - \frac{\frac{\partial^2 \phi}{\partial h_y \partial k_{hy}}}{\frac{\partial^2 \phi}{\partial k_{hy}^2}} = - \left(\frac{\partial^2 \phi}{\partial k_{hy}^2} \right)^{-1} \quad (8)$$

We now need to examine the variations of the phase with k_{hy} . We introduced the ray parameter notation in equation (4). Therefore the phase has the expression shown in equation (9). We can calculate its first and second partial derivatives with respect to k_{hy} (see Appendix B):

$$\phi = \omega dz \left\{ \sqrt{\frac{1}{v(\mathbf{r}, z)^2} - (p_{rx}^2 + p_{ry}^2)} + \sqrt{\frac{1}{v(\mathbf{s}, z)^2} - (p_{sx}^2 + p_{sy}^2)} \right\} + k_{hy} h_y \quad (9)$$

$$\frac{\partial \phi}{\partial k_{hy}} = dz \left\{ \frac{-p_{ry}}{2\sqrt{\frac{1}{v(\mathbf{r}, z)^2} - (p_{rx}^2 + p_{ry}^2)}} + \frac{p_{sy}}{2\sqrt{\frac{1}{v(\mathbf{s}, z)^2} - (p_{sx}^2 + p_{sy}^2)}} \right\} + h_y \quad (10)$$

$$\frac{\partial^2 \phi}{\partial k_{hy}^2} = -\frac{dz}{4\omega} \left\{ \frac{\left(\frac{1}{v(\mathbf{r},z)^2} - p_{rx}^2 \right)}{\left(\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2) \right)^{\frac{3}{2}}} + \frac{\left(\frac{1}{v(\mathbf{s},z)^2} - p_{sx}^2 \right)}{\left(\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2) \right)^{\frac{3}{2}}} \right\} \leq 0 \quad (11)$$

According to equation (6), the computation process will blow up if the second derivative becomes close to zero. The failure happens for (and only for) $p_{rx} = \frac{1}{v}$, corresponding to the quasi-horizontal propagation of evanescent waves. We also see that the second derivative is independent from the value of the cross-line offset, h_y . This means that the variations of k_{hy} with h_y , given by equation (8), will be the same around each stationary point.

Ideally, we would like to have small variations of the stationary path with h_y , in order to use correctly values of k_{hy} around the known solution \hat{k}_{hy}^0 for the stationary points \hat{k}_{hy} , which we cannot determine analytically. Small variations mean a high value of the phase second derivative in equation (11). When the derivative is close to zero, problems in the computation may arise, but this will never happen for waves propagating downward. The “limiting” case is horizontal propagation. It means that narrow-azimuth migration may not completely overcome the steep-dip limitations of common-azimuth migration.

CONCLUSION AND FUTURE WORK

We have presented a review of common-azimuth prestack depth migration theory and proposed an extension to improve the imaging where there are strong lateral variations of the velocity model. Since the extended code is still “work in progress,” we are not, as yet, able to present concrete results. We intend to test the program in the case of a simple $V(z)$ medium and with more complex synthetics, before applying both common-azimuth and narrow-azimuth on a real dataset adequately preprocessed with AMO.

ACKNOWLEDGEMENTS

We would like to thank Elf Aquitaine for providing the data and supporting this study. We thank in particular Henri Calandra for his help and fruitful discussions.

REFERENCES

- Alkhalifah, T., 1997, Prestack time migration for anisotropic media: SEP-94, 263–298.
- Biondi, B., and Palacharla, G., 1994, 3-D prestack migration of common-azimuth data: SEP-80, 109–124.
- Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: Geophysics, 61, no. 6, 1822–1832.

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: *Geophysics*, **63**, no. 2, 574–588.

Biondi, B. 3-D Seismic Imaging: <http://sepwww.stanford.edu/sep/biondo/Lectures>, 1998.

Biondi, B., 1999, Subsalt imaging by common-azimuth migration: *SEP-100*, 113–124.

Bleistein, N. *Mathematical methods for wave phenomena*: Academic Press, 1984.

Claerbout, J. F., 1984, *Imaging the Earth's Interior*: *SEP-40*.

Popovici, A. M., 1995, *Migration to zero offset in variable velocity media*: Ph.D. thesis, Stanford University.

APPENDIX A

The stationary phase theorem announces the following result (Bleistein, 1984):

Integrals of the form

$$I(k) = \int_{-\infty}^{+\infty} e^{ik\phi(t)} f(t) dt \quad (\text{A-1})$$

can be approximated asymptotically when $k \rightarrow \infty$ by:

$$I(k) \approx \sqrt{\frac{2\pi}{k|\phi''(t_0)|}} f(t_0) e^{ik\phi(t_0) + \text{sgn}(\phi''(t_0)) \frac{i\pi}{4}} \quad (\text{A-2})$$

where t_0 is the stationary point at which the phase derivative $\phi'(t) = 0$, $f(t)$ is a complex function, and the phase $\phi(t)$ is real.

The method is based on a high-frequency approximation, the general idea being that the integral has most of its area near the stationary point t_0 . The approximate value of the integral in the neighborhood of the stationary point is then obtained analytically by expanding $\phi(t)$ and $f(t)$ in a Taylor series around t_0 .

APPENDIX B

We develop here the calculation of the second derivative of the phase from its expression in narrow-azimuth migration:

$$\begin{aligned} \phi &= DSR(\omega, \mathbf{k}_m, \mathbf{k}_h, z) dz + k_{hy} h_y \\ &= \omega dz \left\{ \sqrt{\frac{1}{v(\mathbf{r}, z)^2} - (p_{rx}^2 + p_{ry}^2)} + \sqrt{\frac{1}{v(\mathbf{s}, z)^2} - (p_{sx}^2 + p_{sy}^2)} \right\} + k_{hy} h_y \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \phi}{\partial k_{hy}^2} &= \frac{dz}{4\omega} \frac{-\sqrt{\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2)} - p_{ry} \frac{p_{ry}}{\sqrt{\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2)}}}{\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2)} \\
&\quad + \frac{dz}{4\omega} \frac{-\sqrt{\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2)} - p_{sy} \frac{p_{sy}}{\sqrt{\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2)}}}{\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2)} \\
&= \frac{dz}{4\omega} \left\{ \frac{-\frac{1}{v(\mathbf{r},z)^2} + (p_{rx}^2 + p_{ry}^2) - p_{ry}^2}{\left(\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2)\right)^{\frac{3}{2}}} + \frac{-\frac{1}{v(\mathbf{s},z)^2} + (p_{sx}^2 + p_{sy}^2) - p_{sy}^2}{\left(\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2)\right)^{\frac{3}{2}}} \right\} \\
\frac{\partial^2 \phi}{\partial k_{hy}^2} &= -\frac{dz}{4\omega} \left\{ \frac{\left(\frac{1}{v(\mathbf{r},z)^2} - p_{rx}^2\right)}{\left(\frac{1}{v(\mathbf{r},z)^2} - (p_{rx}^2 + p_{ry}^2)\right)^{\frac{3}{2}}} + \frac{\left(\frac{1}{v(\mathbf{s},z)^2} - p_{sx}^2\right)}{\left(\frac{1}{v(\mathbf{s},z)^2} - (p_{sx}^2 + p_{sy}^2)\right)^{\frac{3}{2}}} \right\} \leq 0
\end{aligned}$$

We derive our conclusions concerning the validity of the stationary-phase approximation from this final formula. It also highlights intrinsic limitations of the algorithm with steep dips and evanescent waves.

Short Note

Comparing Kirchhoff with wave equation migration in a hydrate region

Mamta Sinha and Biondo Biondi¹

INTRODUCTION

Prestack migration is necessary before AVO analysis. Most of the present migration algorithms not only try to focus the reflections on the subsurface but also strive to preserve the amplitude for subsequent amplitude studies. A migration/inversion method developed by Lumley (1993) estimates the angle dependent reflectivity at each subsurface point by using least-squares Kirchhoff migration followed by a linearized Zoeppritz elastic parameter inversion for relative contrasts in compressional and shear wave impedance. Another migration algorithm is based on the wave-equation method which uses the Double Square Root equation (DSR) for downward continuation. Migration methods based on DSR operators have been applied to 2-D prestack migration for a long time (Claerbout, 1984). Only recently computationally efficient methods to continue 3-D prestack data have been presented (Biondi and Palacharla, 1995). The objective of this paper is to compare the results of the above two algorithms.

The data are taken from the gas hydrate region in Blake Outer Ridge on offshore Florida. In a previous study (Ecker, 1998), the prestack migration is performed with the Kirchhoff migration/inversion. In this paper the same prestack data is migrated by the DSR wave-equation migration as explained in Prucha et al(1999).

COMPARISON

Although the goal of both Kirchhoff and wave-equation migration is to properly image the subsurface, they differ in many ways. The first difference is that the two migrations are in separate domains. The Kirchhoff migration is in the offset domain and the wave-equation migration is in the angle domain (Prucha et al., 1999). The stacked section obtained by Kirchhoff migration is shown in Figure 3. It shows the BSR quite clearly and also images a relatively flat

¹email: mamta@sep.stanford.edu, biondo@sep.stanford.edu

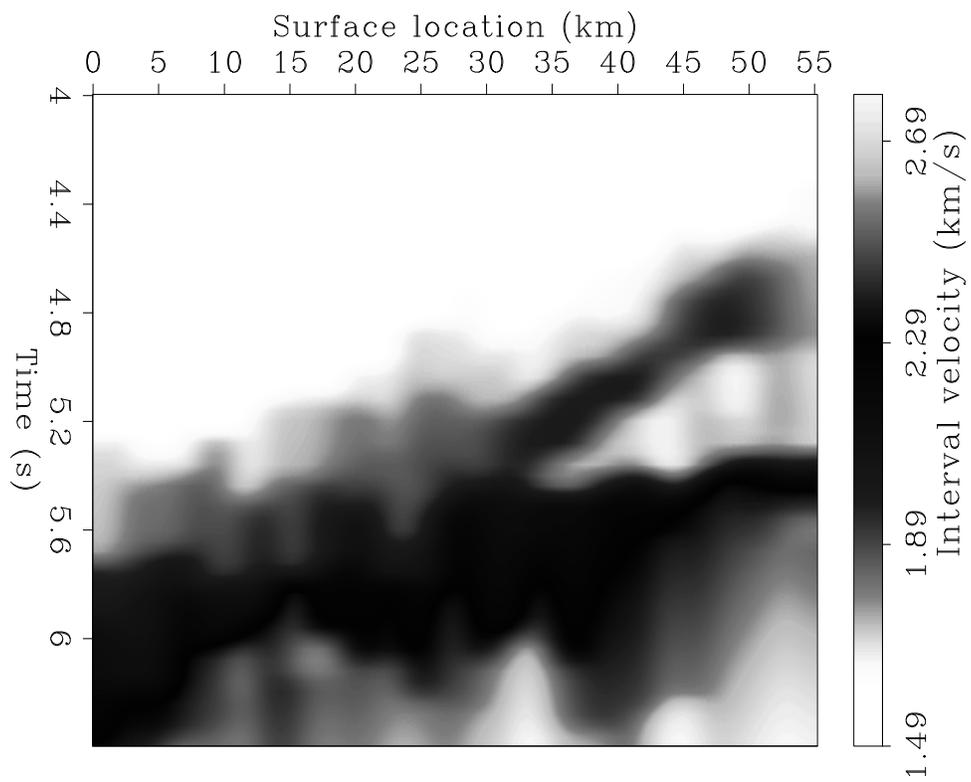


Figure 1: The velocity model used for the two migration methods.
`mamta2-Vel-hydrate-paper` [CR]

reflector beneath the BSR which seems to fade as it approaches the BSR and completely vanishes near the BSR. The same section after wave-equation migration (Figure 4) also images the BSR equally well but better resolves the reflector below the BSR. It is continuous and cuts across the BSR. Though the energy of the reflection still decreases towards the BSR it does not vanish. The latter image has a higher frequency content which shows up as the sharpness of the reflectors.

Constant image gathers are then compared to observe the accuracy of the two methods in resolving the BSR from the layer below. Figure 2 shows the two image gathers at a large distance from the crossing of the BSR with the flat layer. The results are qualitatively similar including the discontinuity in the reflectors. This discontinuity results because of the acquisition geometry, the near offset being double the far offset. It has spread out on the reflection angle gather so is faintly visible there. Figures 5 and 6 are at CMPs 37.5 km and 33.5 km respectively, where the lateral velocity variation is stronger, the wave-equation migration clearly resolves the two reflectors whereas the lower layer vanishes for the Kirchhoff migration.

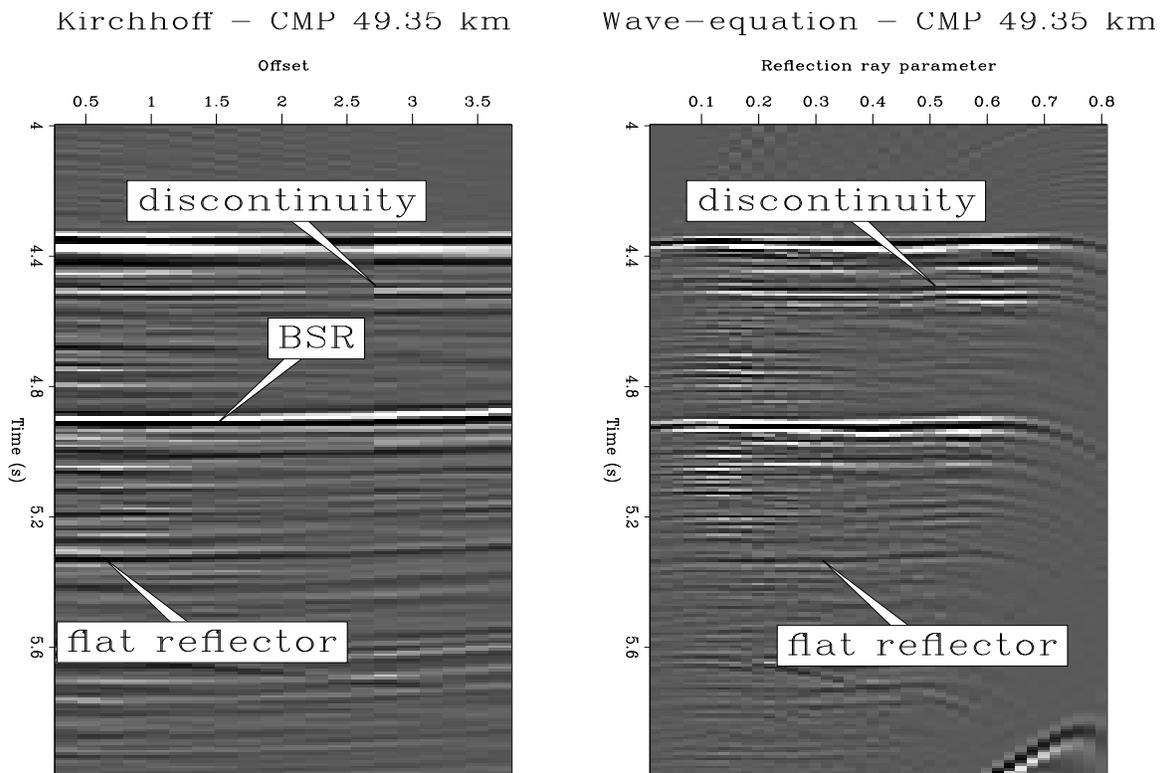


Figure 2: Common midpoint gather at 49.35 km. `mamta2-AVO-49.35-ann` [CR]

DISCUSSION AND FUTURE WORK

The above comparisons lead us to hypothesize that wave-equation migration can yield more accurate estimates of reflectors' amplitudes in places of lateral velocity variation. Since it now

images the lower layer below the BSR, it encourages us to study it in detail. The fact that it fades away as it approaches the BSR could lead to some interesting results. It would enable us to do an amplitude analysis of this particular layer.

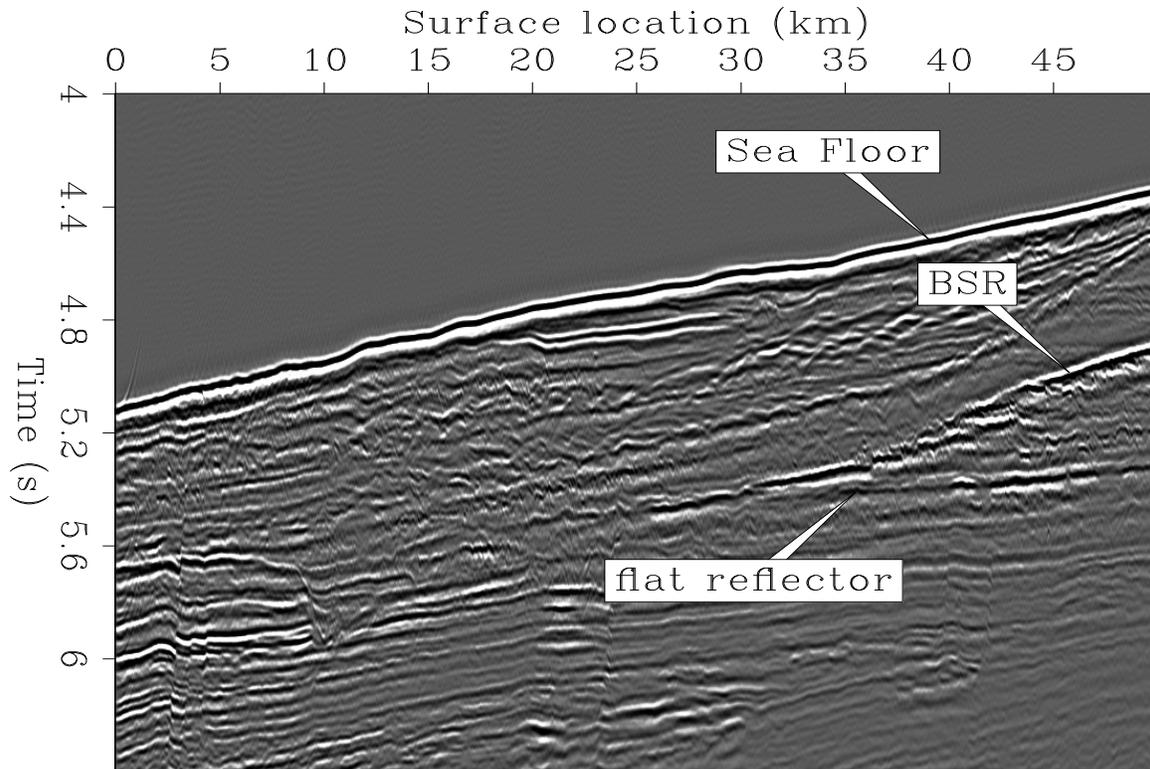


Figure 3: The stacked section after Kirchhoff migration. mamta2-Kir-hydrate-paper-ann
[CR]

REFERENCES

- Biondi, B., and Palacharla, G., 1995, 3-D prestack depth migration of common-azimuth data: SEP-84, 11–24.
- Claerbout, J. F., 1984, The craft of wavefield extrapolation: SEP-40, 230–308.
- Ecker, C., 1998, Seismic characterization of methane hydrates structures: SEP, 96.
- Lumley, D. E., 1993, Kirchhoff prestack impedance inversion: A gas reservoir pilot study: SEP-77, 211–230.
- Prucha, M. L., Biondi, B. L., and Symes, W. W., 1999, Angle-domain common image gathers by wave-equation migration: SEP-100, 101–112.

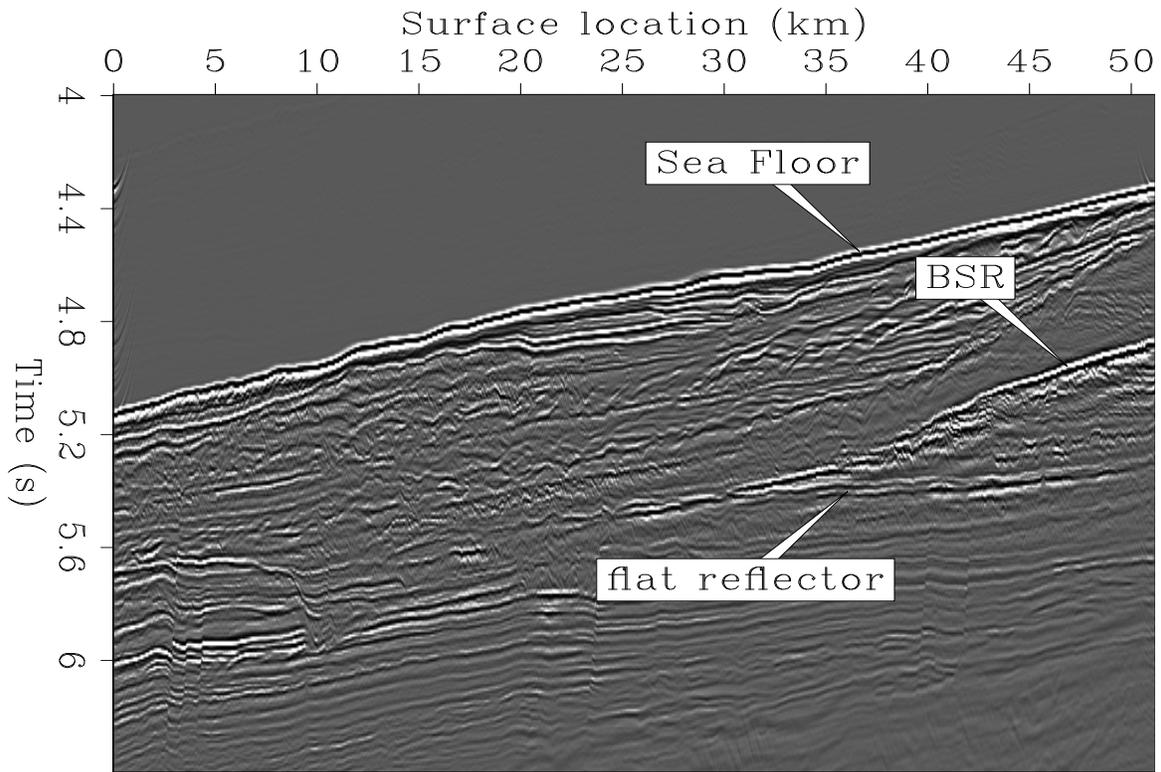


Figure 4: Wave-equation migration. mamta2-Wave-ann [CR]

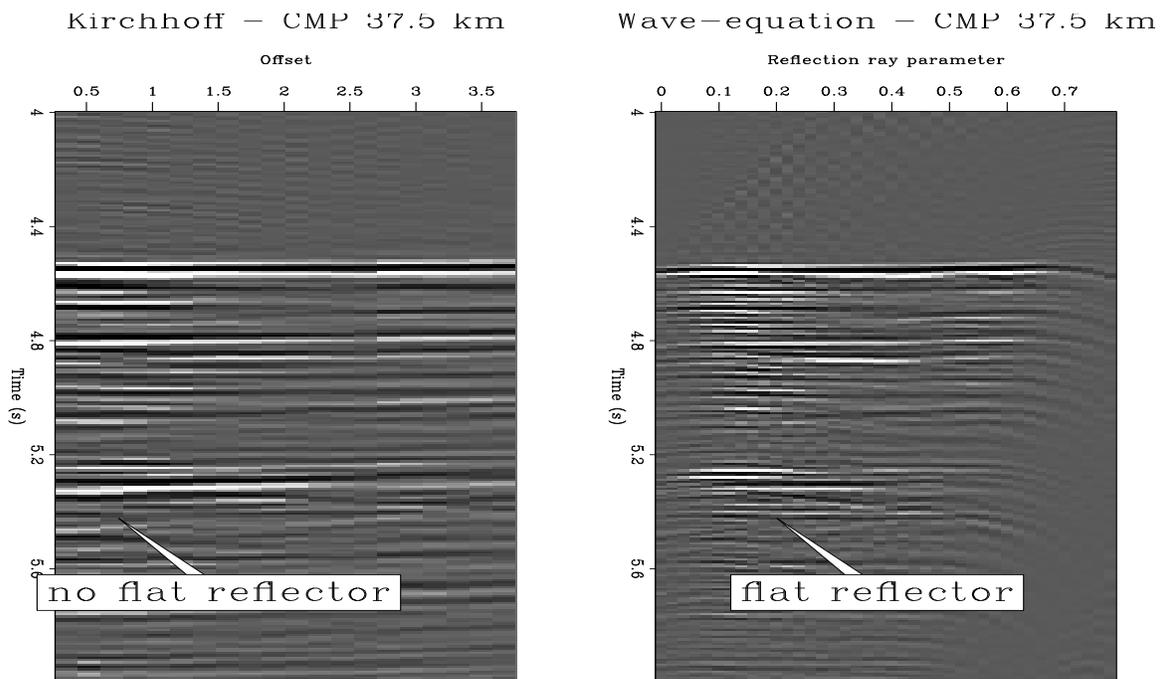


Figure 5: Common midpoint gather at 37.5 km shows the flat reflector on the right section but not on the left section. mamta2-AVO-37.5-ann [CR]

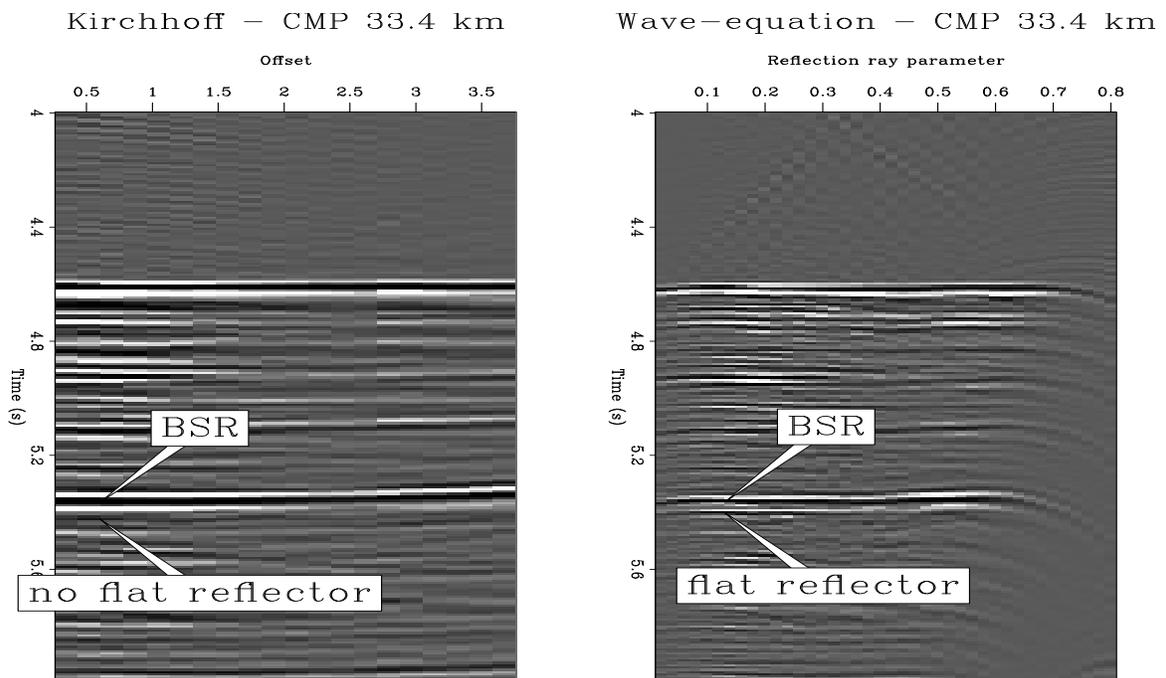


Figure 6: Common midpoint gather at 33.4 km. Right image shows resolved BSR and the flat reflector at around 5.4 sec. Not visible on the left. [mamta2-AVO-33.4-ann](#) [CR]

Angle-gather time migration

Sergey Fomel and Marie Prucha¹

ABSTRACT

Angle-gather migration creates seismic images for different reflection angles at the reflector. We formulate an angle-gather time migration algorithm and study its properties. The algorithm serves as an educational introduction to the angle gather concept. It also looks attractive as a practical alternative to conventional common-offset time migration both for velocity analysis and for AVO/AVA analysis.

INTRODUCTION

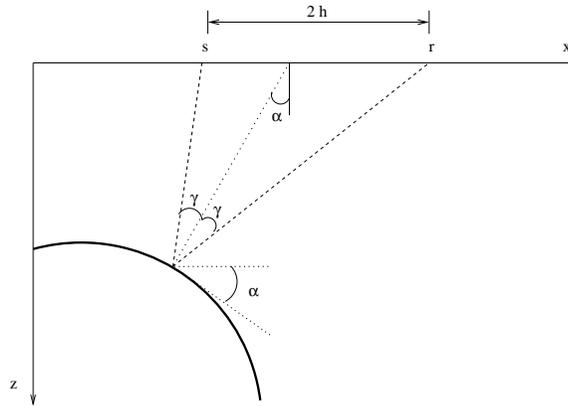
Angle-gather migration creates seismic images collected by the reflection angle at the point of reflection. Major advantages of this approach are apparent in the case of prestack depth migration. As shown by Prucha et al. (1999), the ray pattern of angle-gather migration is significantly different from that of the conventional common-offset migration. The difference can be exploited for overcoming illumination difficulties of the conventional depth migration in complex geological areas.

In this paper, we explore the angle-gather concept in the case of prestack time migration. The first goal of this study is educational. Since we can develop the complete mathematical theory of angle-gather time migration analytically, it is much easier to understand the most basic properties of the method in the time migration domain. The second goal is practical. Angle gathers present an attractive tool for post-migration AVO/AVA studies and velocity analysis, and even the most basic time migration approach can find a valuable place in the complete toolbox of seismic imaging.

We start with analyzing the traveltimes relations for the basic Kirchhoff implementation of angle-gather time migration. The analysis follows Fowler's general approach to prestack time migration methods (Fowler, 1997). Next, we derive formulas for the amplitude weighting and discuss some frequency-domain approaches to angle gathers. Finally, we present simple synthetic tests of the method and discuss further research directions.

Figure 1: Reflection rays in a constant-velocity medium: a scheme.

sergey1-rays [NR]



TRAVELTIME CONSIDERATIONS

Let us consider a simple reflection experiment in an effectively constant-velocity medium, as depicted in Figure 1. The pair of incident and reflected rays and the line between the source s and the receiver r form a triangle in space. From the trigonometry of that triangle we can derive simple relationships among all the variables of the experiment (Fomel, 1995, 1996a, 1997).

Introducing the dip angle α and the reflection angle γ , the total reflection traveltime t can be expressed from the law of sines as

$$t = \frac{2h}{v} \frac{\cos(\alpha + \gamma) + \cos(\alpha - \gamma)}{\sin 2\gamma} = \frac{2h}{v} \frac{\cos \alpha}{\sin \gamma}, \quad (1)$$

where v is the medium velocity, and h is the half-offset between the source and the receiver.

Additionally, by following simple trigonometry, we can connect the half-offset h with the depth of the reflection point z , as follows:

$$h = \frac{z}{2} \frac{\sin 2\gamma}{2 \cos(\alpha + \gamma) \cos(\alpha - \gamma)} = z \frac{\sin \gamma \cos \gamma}{\cos^2 \alpha - \sin^2 \gamma}. \quad (2)$$

Finally, the horizontal distance between the midpoint x and the reflection point ξ is

$$x - \xi = h \frac{\cos(\alpha - \gamma) \sin(\alpha + \gamma) + \cos(\alpha + \gamma) \sin(\alpha - \gamma)}{\sin 2\gamma} = h \frac{\sin \alpha \cos \alpha}{\sin \gamma \cos \gamma} \quad (3)$$

Equations (1–3) completely define the kinematics of angle-gather migration. Regrouping the terms, we can rewrite the three equations in a more symmetric form:

$$t = \frac{2z}{v} \frac{\cos \alpha \cos \gamma}{\cos^2 \alpha - \sin^2 \gamma} \quad (4)$$

¹email: sergey@sep.stanford.edu, marie@sep.stanford.edu

$$h = z \frac{\sin \gamma \cos \gamma}{\cos^2 \alpha - \sin^2 \gamma} \quad (5)$$

$$x - \xi = z \frac{\sin \alpha \cos \alpha}{\cos^2 \alpha - \sin^2 \gamma} \quad (6)$$

For completeness, here is the inverse transformation from t , h , and $x - \xi$ to z , γ , and α :

$$z^2 = \frac{[(vt/2)^2 - (x - \xi)^2] [(vt/2)^2 - h^2]}{(vt/2)^2} \quad (7)$$

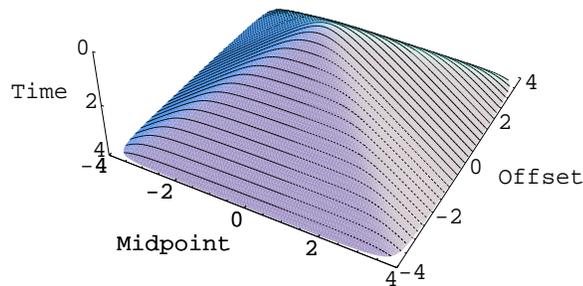
$$\sin^2 \gamma = \frac{h^2 [(vt/2)^2 - (x - \xi)^2]}{(vt/2)^4 - h^2 (x - \xi)^2} \quad (8)$$

$$\cos^2 \alpha = \frac{(vt/2)^2 [(vt/2)^2 - (x - \xi)^2]}{(vt/2)^4 - h^2 (x - \xi)^2} \quad (9)$$

The inverse transformation (7-9) can be found by formally solving system (4-6).

The lines of constant reflection angle γ and variable dip angle α for a given position of a reflection (diffraction) point $\{z, \xi\}$ have the meaning of summation curves for angle-gather Kirchhoff migration. The whole range of such curves for all possible values of γ covers the diffraction traveltime surface - ‘‘Cheops’ pyramid’’ (Claerbout, 1985) in the $\{t, x, h\}$ space of seismic reflection data. As pointed out by Fowler (1997), this condition is sufficient for proving the kinematic validity of the angle-gather approach. For comparison, Figure 2 shows the diffraction traveltime pyramid from a diffractor at 0.5 km depth. The pyramid is composed of common-offset summation curves of the conventional time migration. Figure 3 shows the same pyramid composed of constant- γ curves of the angle-gather migration.

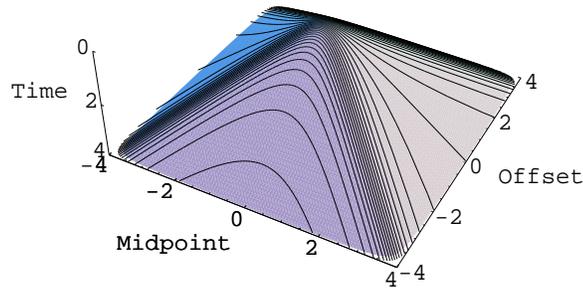
Figure 2: Traveltime pyramid, composed of common-offset summation curves. sergey1-coffset [CR]



The most straightforward Kirchhoff algorithm of angle-gather migration can be formulated as follows:

- For each reflection angle γ and each dip angle α ,
 - For each output location $\{z, \xi\}$,
 1. Find the traveltime t , half-offset h , and midpoint x from formulas (4), (5), and (6) respectively.

Figure 3: Traveltime pyramid, composed of common-reflection-angle summation curves. sergey1-cangle [CR]



2. Stack the input data values into the output.

As follows from equations (4-6), the range of possible α 's should satisfy the condition

$$\cos^2 \alpha > \sin^2 \gamma \quad \text{or} \quad |\alpha| + |\gamma| < \frac{\pi}{2}. \quad (10)$$

The described algorithm is not the most optimal in terms of the input/output organization, but it can serve as a basic implementation of the angle-gather idea. The stacking step requires an appropriate weighting. We discuss the weighting issues in the next section.

AMPLITUDE CONSIDERATIONS

One simple approach to amplitude weighting for angle-gather migration is based again on Cheops' pyramid considerations. Stacking along the pyramid in the data space is a double integration in midpoint and offset coordinates. Angle-gather migration implies the change of coordinates from $\{x, h\}$ to $\{\alpha, \gamma\}$. The change of coordinates leads to weighting the integrand by the following Jacobian transformation:

$$dx dh = \left| \det \begin{pmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial \gamma} \\ \frac{\partial h}{\partial \alpha} & \frac{\partial h}{\partial \gamma} \end{pmatrix} \right| d\alpha d\gamma \quad (11)$$

Substituting formulas (5) and (6) into equation (11) gives us the following analytical expression for the Jacobian weighting:

$$W_J = \left| \det \begin{pmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial \gamma} \\ \frac{\partial h}{\partial \alpha} & \frac{\partial h}{\partial \gamma} \end{pmatrix} \right| = \frac{z^2}{(\cos \alpha^2 - \sin \gamma^2)^2} \quad (12)$$

Weighting (12) should be applied in addition to the weighting used in common-offset migration. By analyzing formula (12), we can see that the weight increases with the reflector depth and peaks where the angles α and γ approach condition (10).

The Jacobian weighting approach, however, does not provide physically meaningful amplitudes, when migrated angle gathers are considered individually. In order to obtain a physically meaningful amplitude, we can turn to the asymptotic theory of true-amplitude migration

(Goldin, 1992; Schleicher et al., 1993; Tygel et al., 1994). The true-amplitude weighting provides an asymptotic high-frequency amplitude proportional to the reflection coefficient, with the wave propagation (geometric spreading) effects removed. The generic true-amplitude weighting formula (Fomel, 1996b) transforms in the case of 2-D angle-gather time migration to the form:

$$W_{\text{TA}} = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{L_s L_r}}{v \cos \gamma} \left| \frac{\partial^2 L_s}{\partial \xi \partial \gamma} + \frac{\partial^2 L_r}{\partial \xi \partial \gamma} \right|, \quad (13)$$

where L_s and L_r are the ray lengths from the reflector point to the source and the receiver respectively. After some heavy algebra, the true-amplitude expression takes the form

$$W_{\text{TA}} = \frac{2z \sin \alpha}{\sqrt{2\pi} v} \frac{\cos^2 \alpha + \sin^2 \gamma}{(\cos^2 \alpha - \sin^2 \gamma)^{5/2}}. \quad (14)$$

Under the constant-velocity assumption and in high-frequency asymptotic, this weighting produces an output, proportional to the reflection coefficient, when applied for creating an angle gather with the reflection angle γ . Despite the strong assumptions behind this approach, it might be useful in practice for post-migration amplitude-versus-angle studies. Unlike the conventional common-offset migration, the angle-gather approach produces the output directly in reflection angle coordinates. One can use the generic true-amplitude theory (Fomel, 1996b) for extending formula (14) to the 3-D and 2.5-D cases.

EXAMPLES

We created some simple synthetic models with constant velocity backgrounds to test our angle-gather migration method. One model is a simple dome (Figure 4). The other has a series of flat reflectors of various dips (Figure 5). Both of these figures also show the corresponding data that will be generated by Kirchhoff methods for zero and far offsets.

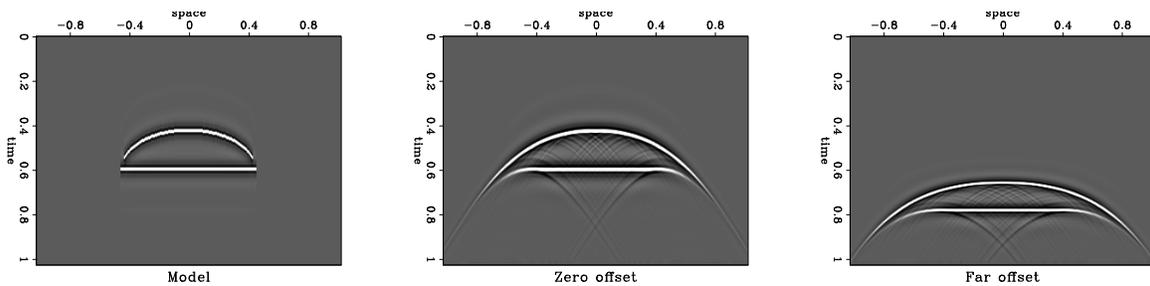


Figure 4: Left: Model. Center: Data at zero offset. Right: Data at far offset. `sergey1-data.dome` [ER]

Dome model

This model contains a wide range of geologic dips across the dome as well as having a flat reflector at the base of the dome. Figure 6 shows the resulting common offset sections from

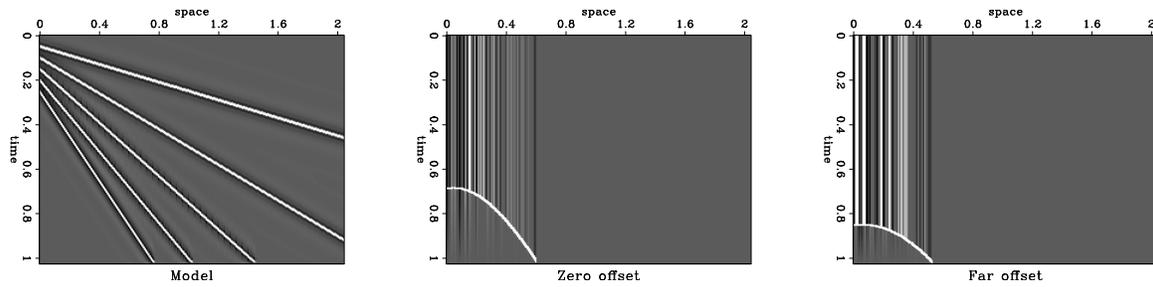


Figure 5: Left: Model. Center: Data at zero offset. Right: Data at far offset. `sergey1-data.lines` [ER]

traditional Kirchhoff migration. As is expected for such a simple model, the near and far offset sections are very similar and the stacked section is almost perfect. We are more interested in the result of the angle-gather migration. Figure 7 shows the zero and large angle sections as well as the stack for angle-gather Kirchhoff migration. The zero-angle section is weak but clearly shows the correct shape and position. The large-angle section is actually only for $\gamma = 25^\circ$. The reason for this is clear if you consider Figure 1. At greater depths, the rays associated with large reflection angles (γ) will not emerge at the surface within the model space. Therefore at angles greater than 25° (the maximum useful angle), the information at later times disappears. We expect the stacked sections for the offset method and the angle method to be identical. Although we sum over different paths for the offset-domain migration (Figure 2) and the angle-domain migration (Figure 3), the stack should sum all of the same information together for both methods. Fortunately, a comparison of the stacked sections in Figures 6 and 7 show that the results are identical as expected.

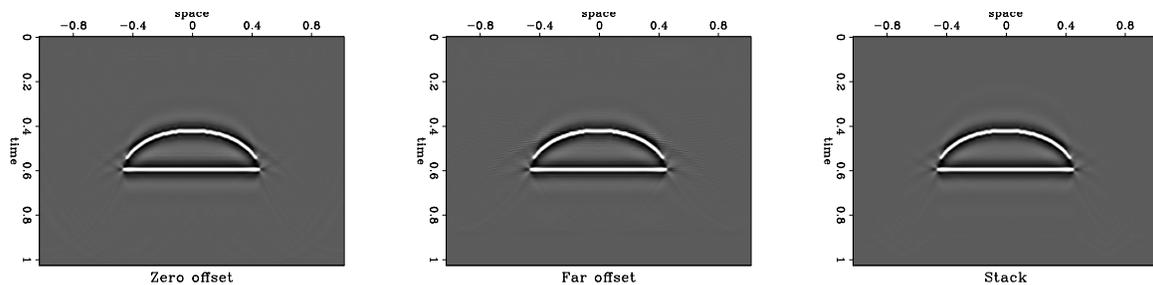


Figure 6: Left: Migrated offset section at zero offset. Center: Migrated offset section at far offset. Right: Stack. `sergey1-offset.dome` [ER]

Dipping reflectors model

This model contains fewer dips than the dome model but it allows us to see what is happening at later times. Figure 8 shows the common offset sections and stacked section from offset-domain Kirchhoff migration. Once again, they are practically perfect. The only problem is near the bottom of the section where we lose energy because the data was truncated. The

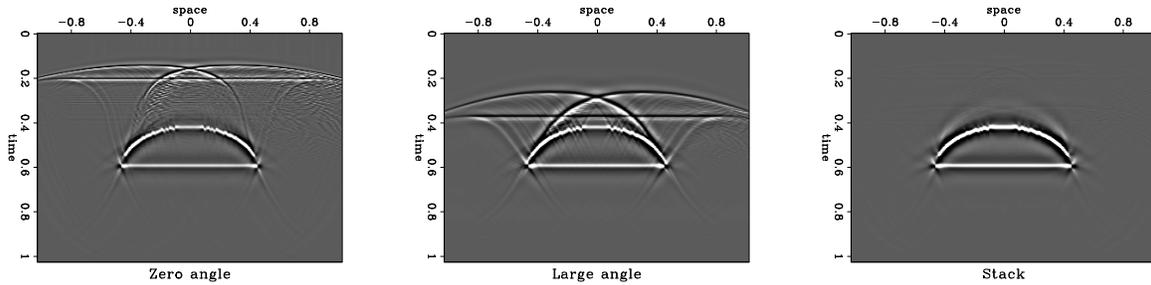


Figure 7: Left: Migrated angle section at small angle. Center: Migrated offset section at large angle. Right: Stack. `sergey1-angle-ta.dome` [ER]

zero-angle and large-angle sections from the angle-domain migration are in Figure 9, along with the stacked section. Once again, the zero angle section is very weak and the large angle section only contains information down to a time of $\approx .85$ seconds, for the same reason as explained for the dome model. Once again, we expect the stacked sections in Figures 8 and 9 to be the same. Although the angle-domain stack is slightly lower amplitude throughout the section, it is clear that this is a simple scale factor so our expectations remain intact.

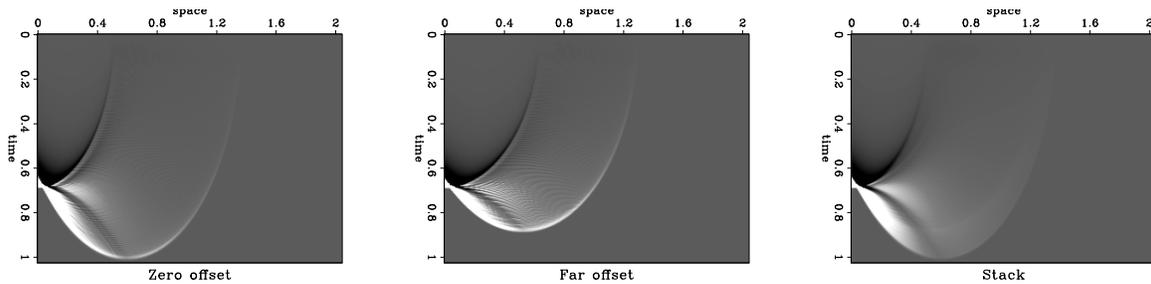


Figure 8: Left: Migrated offset section at zero offset. Center: Migrated offset section at far offset. Right: Stack. `sergey1-offset.lines` [ER]

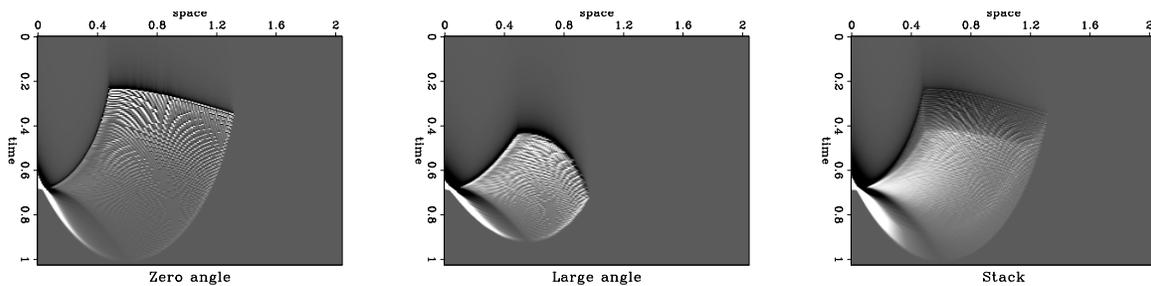


Figure 9: Left: Migrated angle section at zero angle. Center: Migrated angle section at large angle. Right: Stack. `sergey1-angle-ta.lines` [ER]

Reflectivity variation with angle

Amplitude variation with offset (AVO) would not be expected to be very interesting for the simple models just shown. Consider Figure 10 which contains an offset gather and a reflection angle gather taken from space location zero from the dome model in Figure 4. The offset gather shows exactly what we expect for such a model - no variation. The angle gather also shows no variation for angles less than the maximum useful angle (25°) as discussed in the previous two subsections. However, when the angle exceeds the maximum useful angle, the event increases in amplitude and width. This is the phenomenon seen in de Bruin et al. (1990).

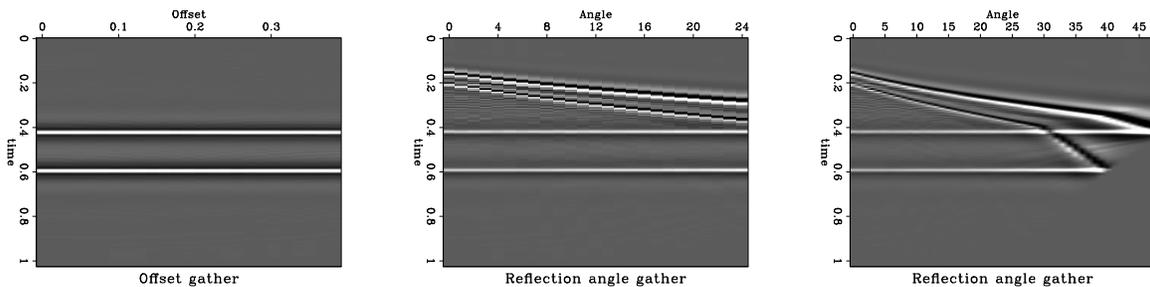


Figure 10: Gathers taken from space location zero in the dome model. Left: Offset domain. Center: Angle domain less than 25° . Right: Angle domain. [sergey1-reflect-ta.dome](#) [ER]

Velocity sensitivity

When dealing with real data we almost never know what the true velocity of the subsurface is. Therefore it is important to understand the effects of velocity on our angle-gather time migration algorithm. To do this we simply created data for the dome model in Figure 4 at a fairly high velocity (3 km/s) and migrated it using a low velocity (1.5 km/s). The results are in Figure 11. For angles less than the maximum useful angle ($\gamma = 25^\circ$), the angle-domain gather behaves exactly as the offset-domain gather does. Beyond the maximum useful angle, the events become even more curved and the amplitudes begin to change. The behavior of the angle-gather migration is very similar to that of offset-domain migration as long as the limitation of the maximum useful angle is recognized. Therefore, we can probably expect angle-gather migration to behave like offset-domain migration in $v(z)$ media also.

FREQUENCY-DOMAIN CONSIDERATIONS

As pointed out by Prucha et al. (1999), the angle gathers can be conveniently formed in the frequency domain. This conclusion follows from the simple formula (Fomel, 1996a)

$$\tan \gamma = \frac{\partial z}{\partial h}, \quad (15)$$

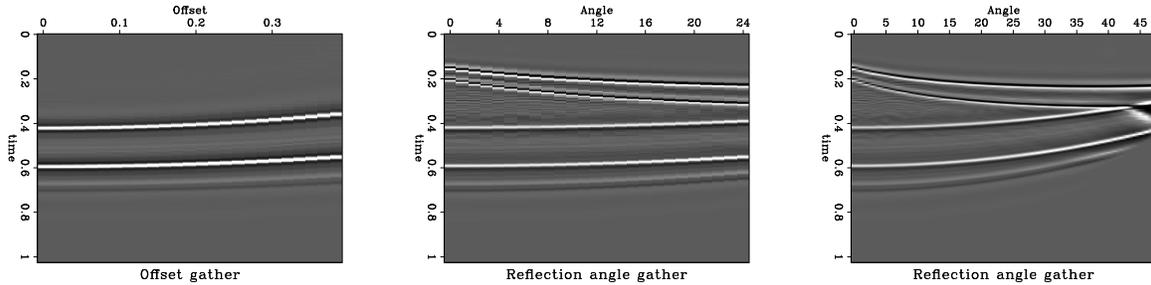


Figure 11: Gathers taken from space location zero in the dome model and migrated at too low a velocity. Left: Offset domain. Center: Angle domain less than 25° . Right: Angle domain.

`sergey1-reflect-ta.fast.dome` [ER]

where z refers to the depth coordinate of the migrated image. In the frequency-wavenumber domain, formula (15) takes the trivial form

$$\tan \gamma = \frac{k_h}{k_z}. \quad (16)$$

It indicates that angle gathers can be conveniently formed with the help of frequency-domain migration algorithms (Stolt, 1978). This interesting opportunity requires further research.

CONCLUSIONS

We have presented an approach to time migration based on angle gathers. The output of this procedure are migrated angle gathers - images for constant reflection angles. When stacked together, angle gathers can produce the same output as the conventional common-offset gathers. Looking at angle gathers individually opens new possibilities for amplitude-versus-angle studies and for velocity analysis.

Our first synthetic tests produced promising results. In the future, we plan to study the amplitude behavior of angle-gather migration and the velocity sensitivity more carefully. We also plan to investigate the frequency-domain approaches to this method. Initial results indicate that angle-gather migration is comparable to offset-domain migration for angles less than the angle at which rays exit the sides of the model, but further study will hopefully allow us to extract useful information from the larger angles as well. Although the major advantages of angle gathers lay in the depth migration domain, it is easier to analyze the time migration results because of their theoretical simplicity.

REFERENCES

- Claerbout, J. F., 1985, *Imaging the Earth's Interior*: Blackwell Scientific Publications.
- de Bruin, C., Wapenaar, C., and Berkhout, A., 1990, Angle-dependent reflectivity by means of prestack migration: *Geophysics*, **55**, no. 9, 1223–1234.

- Fomel, S., 1995, Amplitude preserving offset continuation in theory Part 1: The offset continuation equation: SEP-**84**, 179–198.
- Fomel, S., 1996a, Migration and velocity analysis by velocity continuation: SEP-**92**, 159–188.
- Fomel, S., 1996b, Stacking operators: Adjoint versus asymptotic inverse: SEP-**92**, 267–292.
- Fomel, S., 1997, Velocity continuation and the anatomy of residual prestack migration: 67th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1762–1765.
- Fowler, P., 1997, A comparative overview of prestack time migration methods: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1571–1574.
- Goldin, S. V., 1992, Estimation of reflection coefficient under migration of converted and monotype waves: Soviet Geology and Geophysics, **33**, no. 4, 76–90.
- Prucha, M. L., Biondi, B. L., and Symes, W. W., 1999, Angle-domain common image gathers by wave-equation migration: SEP-**100**, 101–112.
- Schleicher, J., Tygel, M., and Hubral, P., 1993, 3-D true-amplitude finite-offset migration: Geophysics, **58**, no. 8, 1112–1126.
- Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, no. 1, 23–48.
- Tygel, M., Schleicher, J., and Hubral, P., 1994, Kirchhoff-Helmholtz theory in modeling and migration: Journal of Seismic Exploration, **3**, 203–214.

Short Note

On Stolt prestack residual migration

Paul Sava¹

INTRODUCTION

Residual migration has proved to be a useful tool in imaging and in velocity analysis.

Rothman (1983) shows that post-stack residual migration can be successfully used to improve the focusing of the migrated sections. He also showed that migration with a given velocity v_m is equivalent to migration with a reference velocity v_0 followed by residual migration with a velocity v_r that can be expressed as a function of v_0 and v_m .

Residual migration has also been used as a tool in velocity analysis. Al-Yahya (1987) discusses a residual migration operator in the prestack domain, and shows that it can be posed as a function of a nondimensional parameter γ that is the ratio of the correct velocity and the reference velocity used for the initial migration. Etgen (1988, 1989) defines a kinematic residual migration operator as a cascade of NMO and DMO, and shows that it, again, is only a function of the nondimensional parameter γ defined by Al-Yahya. Finally, Stolt (1996) defines a prestack residual migration operator in the (f, k) domain, and shows that it depends on the reference (v_0) and the correct (v_m) migration velocities.

In this short note, I review the prestack residual Stolt migration, and show that it also can be formulated as a function of a nondimensional parameter that is the ratio of the reference (v_0) and correct (v_m) velocities. Consequently, we can use Stolt residual migration in the prestack domain to obtain a better focused image without making any assumption about the velocity. This approach has a direct application to migration velocity analysis, for instance in cases when we repeatedly do residual migration on data that have been depth-migrated with an arbitrary velocity function that cannot be approximated by a constant velocity (Biondi and Sava, 1999).

¹email: paul@sep.stanford.edu

STOLT MIGRATION

Prestack Stolt migration (PrSM) was summarized as (Claerbout, 1985)

$$p(t, y, h) \rightarrow P(\omega, k_y, k_h) \rightarrow P'(k_z, k_y, k_h) \rightarrow p'(z, y, h).$$

An important component of PrSM is the remapping from the (ω, k_y, k_h) domain to the (k_z, k_y, k_h) domain, where ω, k_z represent, respectively, the frequency and the vertical wavenumber, and k_y, k_h represent the midpoint and offset wavenumbers.

If we consider the alternative representation of the input data in shot-geophone coordinates, the mapping takes the form

$$k_z = \frac{1}{2} \left(\sqrt{\frac{\omega^2}{v^2} - k_g^2} + \sqrt{\frac{\omega^2}{v^2} - k_s^2} \right), \quad (1)$$

where k_g and k_s stand for, respectively, the geophone and the source wavenumbers. It is desirable to implement the remapping as a pull operator, to avoid numerical problems in the inverse Fourier transform. A detailed discussion on the advantages and disadvantages of the different mappings is done by Levin (1994). We can, therefore, express ω as a function of k_z from Equation (1) as:

$$\omega^2 = \frac{v^2}{16k_z^2} [4k_z^2 + (k_g - k_s)^2] [4k_z^2 + (k_g + k_s)^2] \quad (2)$$

or

$$\omega^2 = \frac{v^2}{k_z^2} [k_z^2 + k_h^2] [k_z^2 + k_y^2]. \quad (3)$$

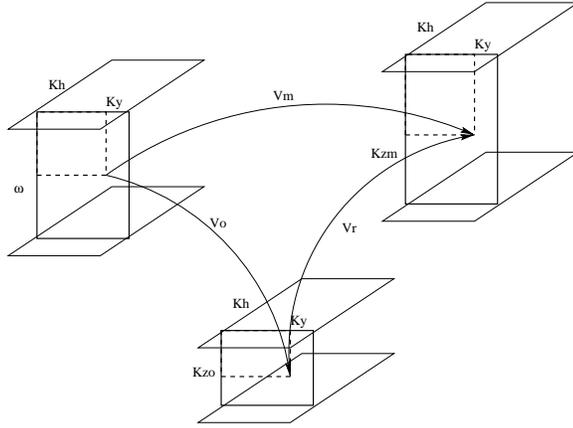
2-D RESIDUAL STOLT MIGRATION

In general, residual migration represents a method of improving the quality of the image without having to remigrate the original data, but rather only applying a transformation to the current migration image.

In residual prestack Stolt migration (RPrSM), we attempt to correct the effects of migrating with an inaccurate reference velocity by applying a transformation to the data that have been transformed to the Fourier domain (Figure 1). Supposing that the initial migration was done with the velocity v_0 , and that the correct velocity is v_m , we can then write

$$\begin{cases} k_{z_0} = \frac{1}{2} \left(\sqrt{\frac{\omega^2}{v_0^2} - k_g^2} + \sqrt{\frac{\omega^2}{v_0^2} - k_s^2} \right) \\ k_{z_m} = \frac{1}{2} \left(\sqrt{\frac{\omega^2}{v_m^2} - k_g^2} + \sqrt{\frac{\omega^2}{v_m^2} - k_s^2} \right). \end{cases} \quad (4)$$

Figure 1: A sketch of Stolt residual migration `paul1-stolt` [NR]



The goal of RPrSM is to obtain k_{z_m} from k_{z_0} . If we use the first equation of (4) to substitute ω in the second equation of (4), we obtain

$$k_{z_m} = \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (k_g - k_s)^2][4k_{z_0}^2 + (k_g + k_s)^2]}{16k_{z_0}^2}} - k_g^2 + \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (k_g - k_s)^2][4k_{z_0}^2 + (k_g + k_s)^2]}{16k_{z_0}^2}} - k_s^2 \quad (5)$$

or

$$k_{z_m} = \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[k_{z_0}^2 + k_h^2][k_{z_0}^2 + k_y^2]}{k_{z_0}^2}} - (k_y + k_h)^2 + \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[k_{z_0}^2 + k_h^2][k_{z_0}^2 + k_y^2]}{k_{z_0}^2}} - (k_y - k_h)^2. \quad (6)$$

Equation (6) represents the RPrSM equation in two dimensions. For post-stack data, the same equation takes the familiar form

$$k_{z_m} = \sqrt{\frac{v_0^2}{v_m^2} [k_{z_0}^2 + k_y^2]} - k_y^2. \quad (7)$$

3-D RESIDUAL STOLT MIGRATION

The RPrSM equations in three dimensions can be written similarly to those in two dimensions if we consider the relationship between source-geophone and midpoint-offset coordinates as follows:

$$\begin{aligned} \vec{k}_g &= \vec{k}_y + \vec{k}_h \\ \vec{k}_s &= \vec{k}_y - \vec{k}_h. \end{aligned}$$

The equivalent prestack residual migration equation in source-geophone coordinates thus becomes

$$\begin{aligned} k_{z_m} &= \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (|\vec{k}_g| - |\vec{k}_s|)^2][4k_{z_0}^2 + (|\vec{k}_g| + |\vec{k}_s|)^2]}{16k_{z_0}^2}} - |\vec{k}_g|^2 \\ &+ \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (|\vec{k}_g| - |\vec{k}_s|)^2][4k_{z_0}^2 + (|\vec{k}_g| + |\vec{k}_s|)^2]}{16k_{z_0}^2}} - |\vec{k}_s|^2. \end{aligned} \quad (8)$$

In midpoint-offset coordinates, the same equation becomes

$$k_{z_m} = \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (|\vec{k}_y + \vec{k}_h| - |\vec{k}_y - \vec{k}_h|)^2][4k_{z_0}^2 + (|\vec{k}_y + \vec{k}_h| + |\vec{k}_y - \vec{k}_h|)^2]}{16k_{z_0}^2}} - |\vec{k}_y + \vec{k}_h|^2} + \frac{1}{2} \sqrt{\frac{v_0^2}{v_m^2} \frac{[4k_{z_0}^2 + (|\vec{k}_y + \vec{k}_h| - |\vec{k}_y - \vec{k}_h|)^2][4k_{z_0}^2 + (|\vec{k}_y + \vec{k}_h| + |\vec{k}_y - \vec{k}_h|)^2]}{16k_{z_0}^2}} - |\vec{k}_y - \vec{k}_h|^2}, \quad (9)$$

which, for the post-stack case, takes the form

$$k_{z_m} = \sqrt{\frac{v_0^2}{v_m^2} [k_{z_0}^2 + |\vec{k}_y|^2]} - |\vec{k}_y|^2. \quad (10)$$

EXAMPLES

In this section, I present two 2-D post-stack synthetic examples, shown in Figures 2 and 3, to prove the applicability of the equations derived in the preceding sections.

In the first example, the input is a set of three spikes. Initially, I do forward modeling with a velocity of $v_m = 3.0$ km/s. Next, I do Stolt migration with a velocity of $v_0 = 3.6$ km/s, and residual Stolt migration with a ratio $v_0/v_m = 1.2$ (Figure 2). I then take the same input and perform Stolt migration with a velocity of $v_0 = 2.4$ km/s, followed by residual Stolt migration with a ratio $v_0/v_m = 0.8$ (Figure 3). In both cases, the data are correctly collapsed at the location of the original spikes. Residual migration can be done without knowing the absolute values of the velocity.

In the next example, shown in Figure 4, I apply the same methodology to a real dataset (Ecker, 1998). All three panels are the result of residual migration as described in the preceding theory sections. The corresponding ratios are 0.96 for the top panel, 0.98 for middle panel, and 1.00 for the bottom panel. Residual migration with ratio=1.0 is equivalent to no residual migration at all. It is apparent that different images are focused better in one region or another, although the image corresponding to the ratio 0.98 seems to have the highest overall energy. We can use such an observation to obtain an image that has the best focusing in all the regions. This can be achieved by doing residual migration for a range of velocity ratios, and then interpolating the image that is best focused everywhere. One possible application is in wave-equation migration velocity analysis, where we can improve the focusing of a depth-migrated dataset by residual migration, without making any assumption about the original velocity distribution (Biondi and Sava, 1999).

CONCLUSIONS

Equations (5) and (6) show that residual prestack Stolt migration can be done without actual knowledge of the reference and correct velocities, but only by knowing or assuming their ratio. This understanding has important practical consequences, for example, in applications that use

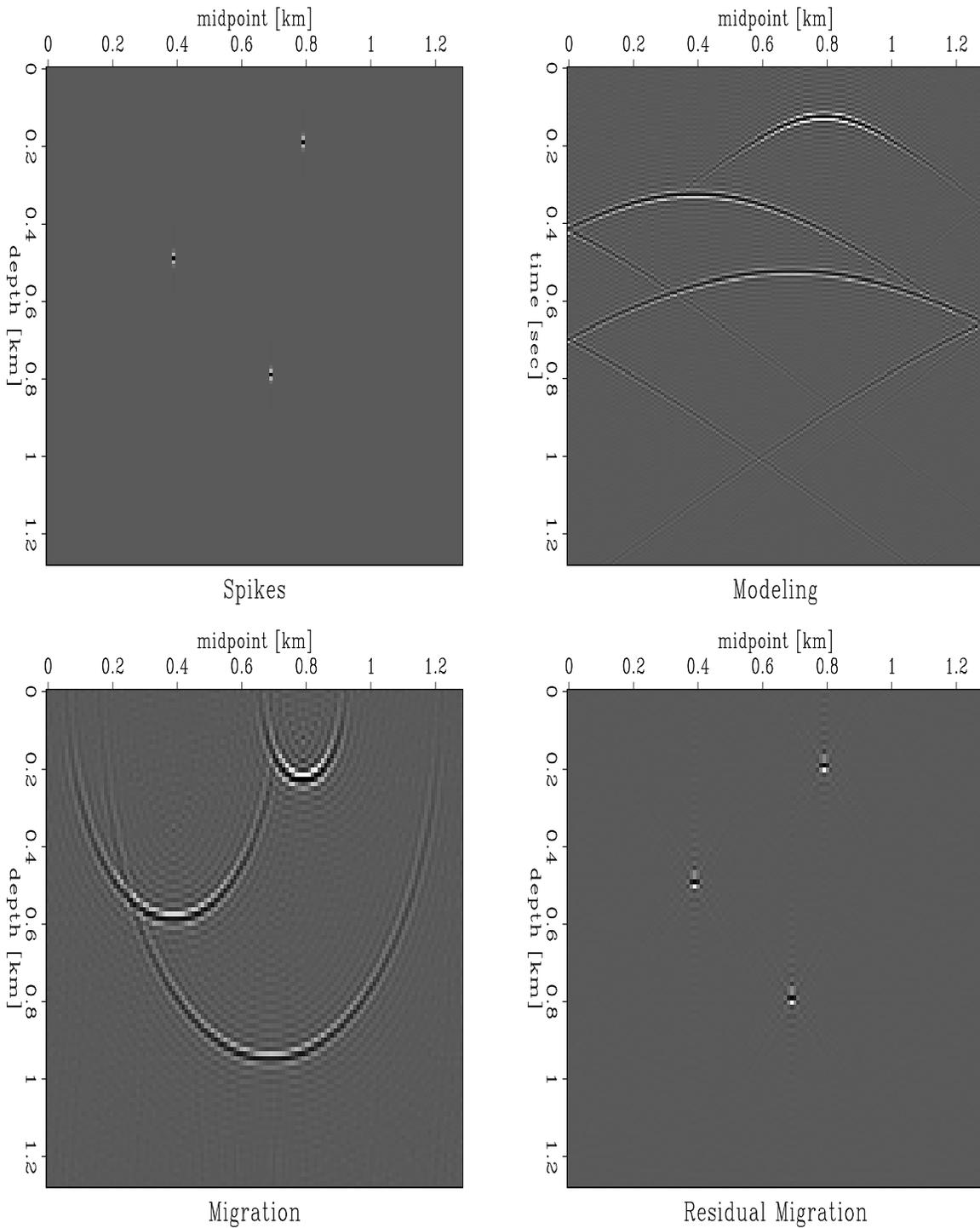


Figure 2: Top left: input data. Top right: Stolt forward modeling with the correct velocity $v_m = 3.0$ km/s. Bottom left: Stolt migration with an incorrect velocity of $v_0 = 3.6$ km/s. Bottom right: Stolt residual migration with the ratio $v_0/v_m = 1.2$. `paul1-posp` [ER]

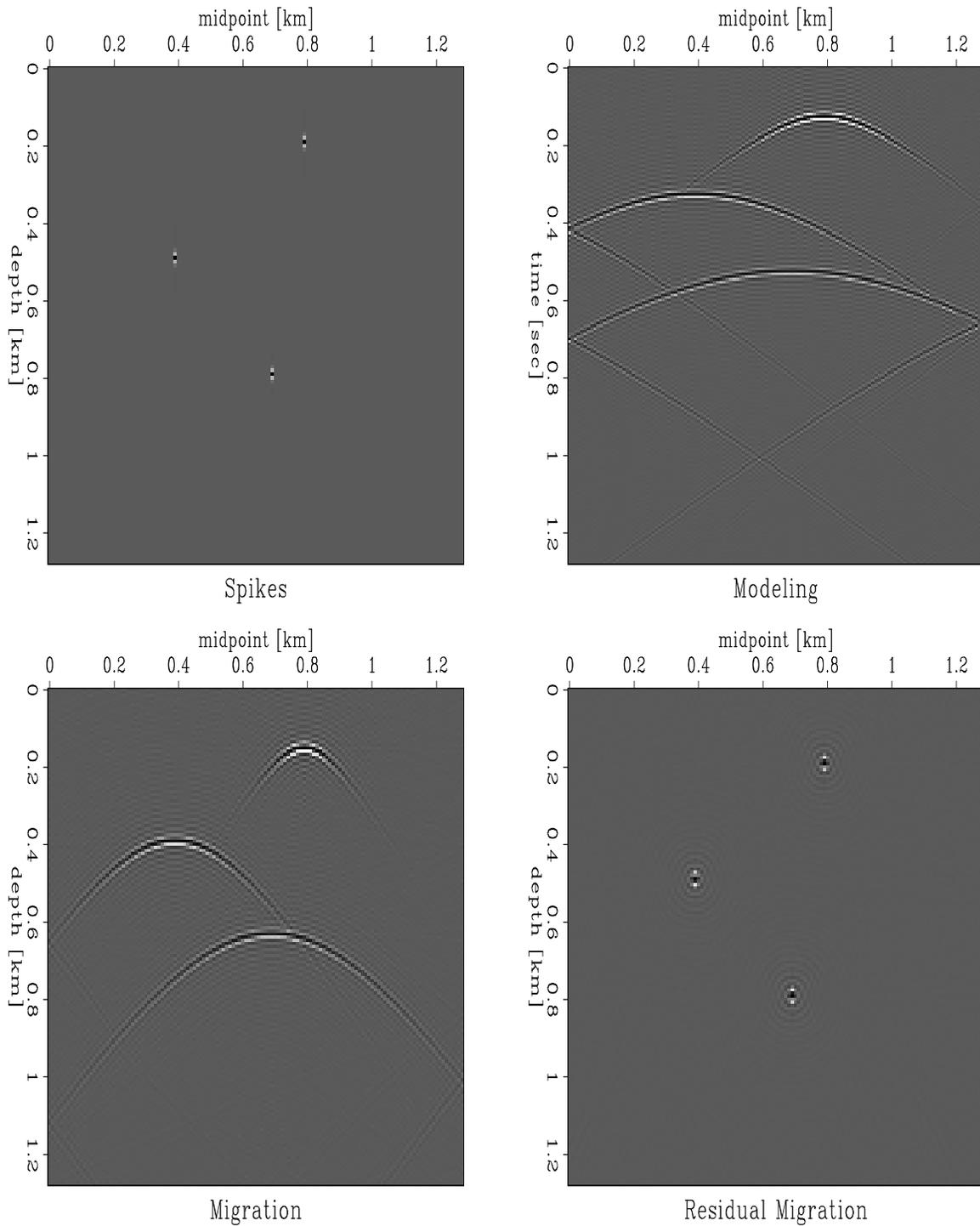


Figure 3: Top left: input data. Top right: Stolt forward modeling with the correct velocity $v_m = 3.0$ km/s. Bottom left: Stolt migration with an incorrect velocity of $v_0 = 2.4$ km/s. Bottom right: Stolt residual migration with the ratio $v_0/v_m = 0.8$. `paul1-posm` [ER]

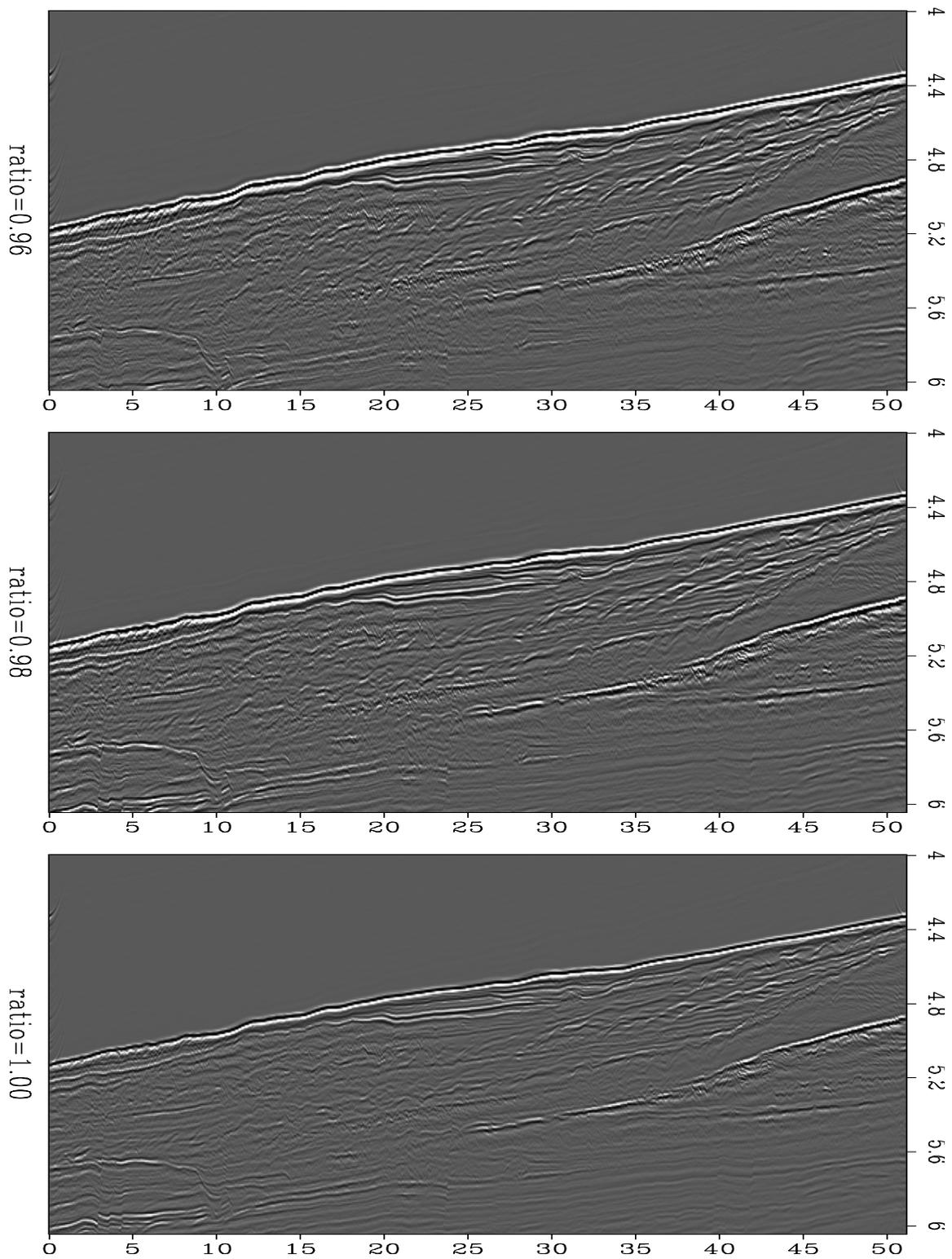


Figure 4: Images obtained after residual migration. Ratio = 1.00 means no residual migration. Different images are focused better in different regions. The image corresponding to ratio = 0.98 seems to have the best focusing. `paul1-hydrates` [CR]

RPrSM as a method of improving the focusing of an image that has been depth-migrated with an arbitrary velocity function.

ACKNOWLEDGMENTS

I would like to thank Sergey Fomel for allowing me to use his interpolation code in the Stolt programs.

REFERENCES

- Al-Yahya, K., 1987, Velocity analysis by iterative profile migration: Ph.D. thesis, Stanford University.
- Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: SEP-100, 11–34.
- Claerbout, J. F., 1985, Imaging the Earth's Interior: Blackwell Scientific Publications.
- Ecker, C., 1998, Seismic characterization of gas hydrates structures: Ph.D. thesis, Stanford University.
- Etgen, J., 1988, Velocity analysis by prestack depth migration: Linear theory: SEP-57, 77–98.
- Etgen, J., 1989, Kinematic residual prestack migration: SEP-61, 79–102.
- Levin, S. A., 1994, Stolt without artifacts? — dropping the Jacobian: SEP-80, 513–532.
- Rothman, D. H., Levin, S. A., and Rocca, F., 1983, Residual migration: SEP-35, 153–174.
- Stolt, R. H., 1996, Short note—a prestack residual time migration operator: Geophysics, 61, no. 02, 605–607.

Anti-aliasing multiple prediction beyond two dimensions

Yalei Sun¹

ABSTRACT

Theoretically, the Delft method of surface-related multiple elimination can be applied in three dimensions, as long as the source and receiver coverage is dense enough. In reality, such a dense coverage is still far from reach, using the available multi-streamer acquisition system. One way to fill the gap is to massively interpolate the missing sources and receivers in the survey, which requires a huge computational cost. In this paper, I propose a more practical approach for the multi-streamer system. Instead of using large-volume missing-streamer interpolation, my method finds the most reasonable proxy from the collected dataset for each missing trace needed in the multiple prediction. Although this approach avoids missing-streamer interpolation, another problem pops up in the multi-streamer case, the aliasing noise caused by the sparse sampling in the cross-line direction. To solve this problem, I introduce a new concept, the partially-stacked multiple contribution gather (PSMCG). Using multi-scale prediction-error filter (MSPEF) theory, this approach interpolates the PSMCG in the cross-line direction to remove the aliasing noise.

INTRODUCTION

The Delft approach to surface-related multiple elimination (Berkhout and Verschuur, 1997; Verschuur and Berkhout, 1997) formulated the demultiple process as a two-step inversion problem based on the Huygens principle, that is, first predicting the multiple and then subtracting it from the original dataset.

The multiple prediction step, crucial for the success of the whole algorithm, involves one important assumption about the data acquisition geometry, namely, a source/receiver pair is needed wherever a multiple reflects. The Delft approach is quite successful in solving 2-D problems (Verschuur and Prein, 1999), in which the assumption is relatively easily satisfied. However, in many 3-D surveys, there is a large gap between this assumption and the reality (Dragoset and Jeričević, 1998).

Two different directions have been taken to solve the problem. One is to interpolate the trace at the missing source and receiver positions massively to attain a dense coverage of the surface (van Dedem and Verschuur, 1998). However, the computational cost of this method is huge. The other approach is to predict the multiple based on the 2-D theory and then extend

¹email: yalei@sep.stanford.edu

the subtraction step to handle incorrectly predicted multiples (Ross, 1997; Ross et al., 1997). The success of this approach is restricted to simple 3-D cases.

This paper proposes a method designed for multi-streamer geometry. Two distinctive features make it more practical. First, this approach finds the most reasonable proxy from the collected dataset for any missing trace. There is no need to interpolate missing streamers and shotlines. Second, I introduce a concept, the partially-stacked multiple contribution gather (PSMCG). Using the multi-scale prediction-error filter (MSPEF) theory (Claerbout, 1992), the proposed approach interpolates the PSMCG in the cross-line direction. This gives us a densely sampled multiple contribution from all possible Huygens secondary sources. The following summation step can remove aliasing noise better.

Two numerical examples in this paper demonstrate how the approach works.

MULTIPLE PREDICTION BEYOND TWO DIMENSIONS

One of key issues in 3-D demultiple is that there are many missing traces. Instead of interpolating the missing traces, the critical points of my approach are:

1. This approach tries to find proxies for the missing traces.
2. To be qualified as a proxy, a trace must have the same offset and either a similar CMP location or a similar azimuth angle.
3. Those proxies for the missing traces are used to predict multiples with first-order accuracy in multi-streamer geometry.

In order to better understand the approach, let's assume that we have a multi-streamer acquisition system, as shown in Figure 1, with one shotline and seven streamers. Supposing that we want to predict the multiple from source S_0 to receiver R_4 , we need to consider the contributions from all the possible multiple reflection points between S_0 and R_4 by cross-convolution. For instance, we need to collect all the traces with sources located at S_i ($i = 1, \dots, 7$) and a receiver located at R_4 . In Figure 1, the thin solid line represents the corresponding trace collected in the survey, and the thin dashed line stands for a missing trace in the survey. The challenge is to find appropriate proxies for such missing traces.

There is one well-known geophysical concept that can help us meet the challenge, the common-midpoint (CMP), which assumes that traces with the same CMP location and the same offset contain the same information about one location in the earth. Although the common-midpoint assumption is a first-order approximation when the structure is not strictly flat, I will demonstrate that it is useful in our search for the substitute traces.

For example, for the virtual trace $\overrightarrow{S_1 R_4}$, the real trace $\overrightarrow{S_4 R_1}$ shares the same CMP location and has the same offset as well. The only difference is the azimuth angle. Therefore, trace $\overrightarrow{S_4 R_1}$ is a proxy for trace $\overrightarrow{S_1 R_4}$ in the multiple prediction, with first-order accuracy. Similarly, we can find substitutes for other virtual traces.

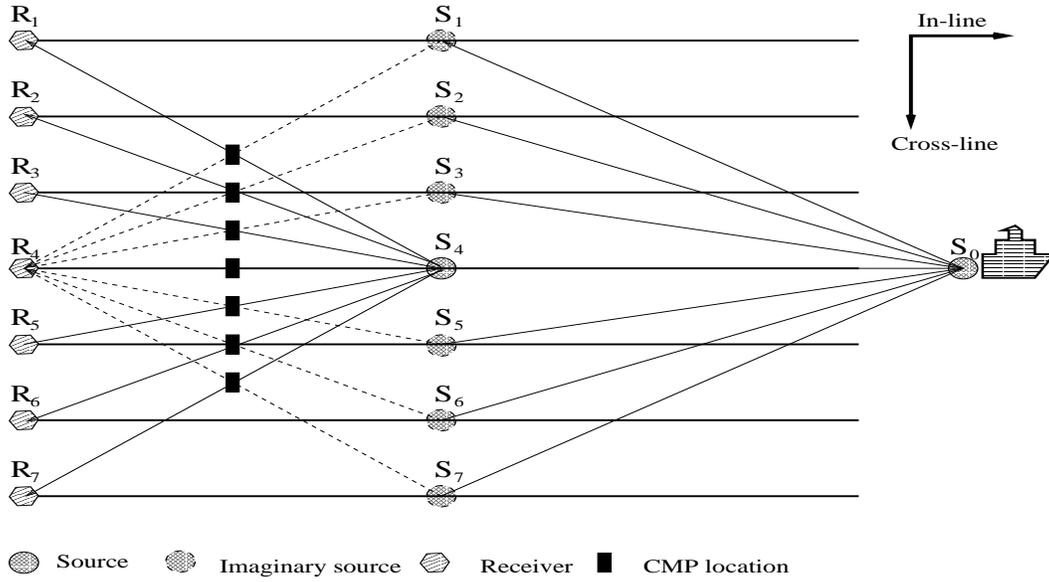


Figure 1: A multi-streamer geometry. The $\overrightarrow{S_4 R_1}$ path is a proxy for the absent out-of-plane shot $\overrightarrow{S_1 R_4}$, since the two paths share the same midpoint and have the same offset. Similarly, $\overrightarrow{S_2 R_4}$ is replaced by $\overrightarrow{S_4 R_2}$, $\overrightarrow{S_3 R_4}$ by $\overrightarrow{S_4 R_3}$, $\overrightarrow{S_5 R_4}$ by $\overrightarrow{S_4 R_5}$, $\overrightarrow{S_6 R_4}$ by $\overrightarrow{S_4 R_6}$, and $\overrightarrow{S_7 R_4}$ by $\overrightarrow{S_4 R_7}$.
yalei1-multi-streamer [NR]

The central streamer in Figure 1 is a special case, in which we can always find substitute traces for the virtual ones with the same CMP location and offset. When we try to predict other streamers' multiple reflections, though, as in Figure 2, we are not so lucky to find proxies with the same CMP location and offset. However, we can relax the definition of a substitute trace by giving up the requirement that the proxy share the same CMP location. Then we can find another group of proxies for the missing traces, as Figure 2 illustrates. Since the cross-line spreading aperture is usually smaller than the in-line aperture, this extension may be acceptable in many real applications.

There are some limitations to the method discussed in this paper. Before addressing those limitations, I would first classify the surface multiple reflections into two categories. Figure 3 depicts two types of geometries for the surface multiples, $S_0 \overrightarrow{M_1} R_1$ and $S_0 \overrightarrow{M_2} R_2$. The difference between these two categories is that, source S_0 , surface multiple reflection position M_1 , and receiver R_1 are aligned together on the surface; whereas S_0 , M_2 , and R_2 can not be aligned together. The embedded physical reason is that, multiple $S_0 \overrightarrow{M_1} R_1$ is mainly caused by 1-D earth's structures or in-line dip reflectors (2.5-D), and multiple $S_0 \overrightarrow{M_2} R_2$ by cross-line dips or scattering reflectors.

The definition of proxies in my proposal guarantees that the method in this paper is fully applicable to the multiples like $S_0 \overrightarrow{M_1} R_1$ without kinematic approximations. The approximation errors occur only when we deal with the multiples like $S_0 \overrightarrow{M_2} R_2$. In other words, when

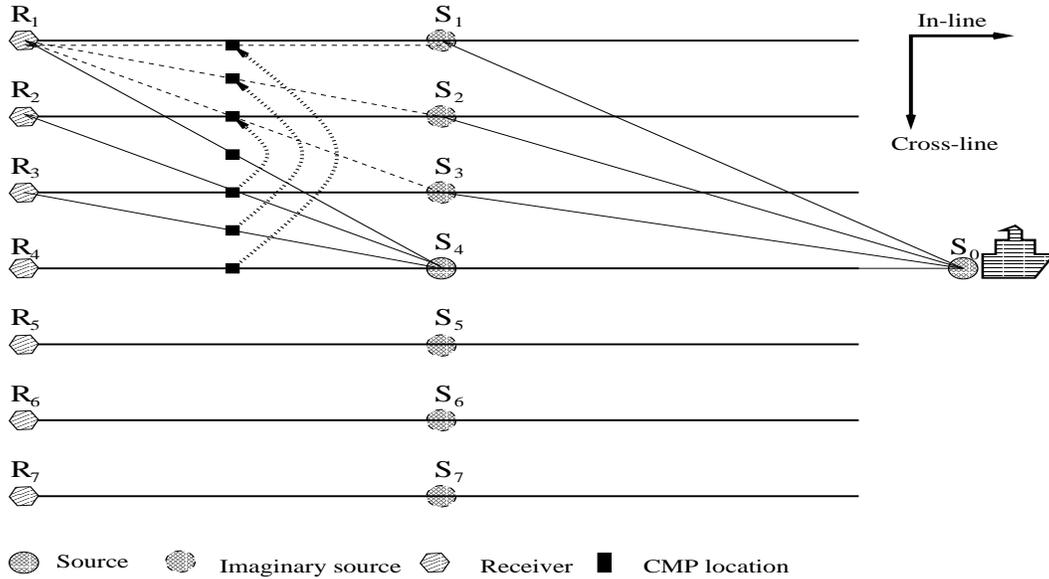
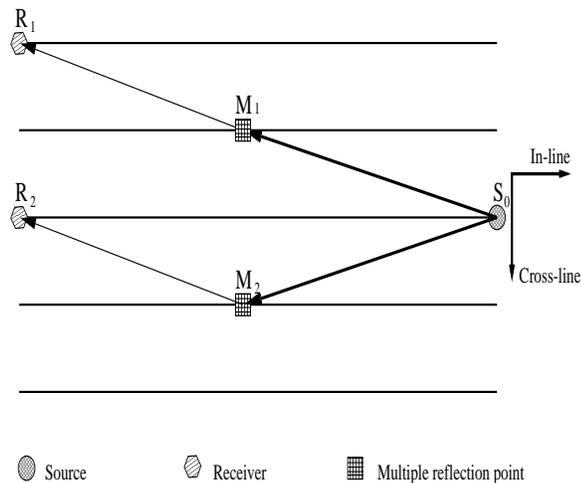


Figure 2: While working with streamers not in the middle, we have to give up the requirement that the proxy share the same CMP location, and find a group of proxies with reasonable accuracy. $\vec{S_1 R_1}$ can be replaced by $\vec{S_4 R_4}$, $\vec{S_2 R_1}$ by $\vec{S_4 R_3}$, and $\vec{S_3 R_1}$ by $\vec{S_4 R_2}$. Each pair shares the same offset and azimuth angle. `yalei1-multi-streamer-1` [NR]

there are strong cross-line dips or scattering reflectors, my approach will introduce the approximation errors inevitably.

Figure 3: Two types of surface multiples. $\vec{S_0 M_2 R_2}$, in which the source S_0 , the multiple reflection point M_2 , and the receiver R_2 can be aligned, is more likely caused by cross-line dips or scattering reflectors. $\vec{S_0 M_1 R_1}$, in which S_0 , M_1 , and R_1 are aligned together, occurs when the earth's structures are approximately 1-D or there are only in-line dip reflectors. `yalei1-multiple-type` [NR]



A MULTIPLE CONTRIBUTION GATHER

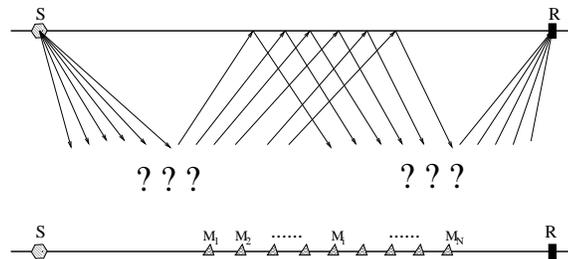
The multiple prediction can be further divided into two sub-steps: trace cross-convolution and multiple contribution summation, which, in practice, people usually collapse into a single

procedure. In order to gain more insight into the method, however, I consider them as two separate steps.

Figure 4 schematically demonstrates 2-D multiple prediction. In order to predict the multiple from source S to receiver R , we need to cross-convolute all the possible contributing traces marked by M_i in the middle, since we do not know where exactly the multiple reflection occurs on the surface. Then the summation step locates the exact multiple reflection position as long as the contributing traces are densely sampled on the surface.

Figure 4: 2-D multiple prediction. All the contributing traces form a 2-D MCG prior to summation. The question marks in the plot indicate that the underground structures are unknown.

yalei1-noah2d [NR]



Prior to summation, if we lay out the cross-convoluted traces from left to right, we get an intermediate result, a multiple contribution gather (MCG). Figure 5 shows a 2-D multiple contribution gather and the corresponding summation result. The MCG is a section in 2-D and a cube in 3-D. The restriction of the multi-streamer geometry makes the 3-D MCG cube densely sampled in the in-line direction and coarsely sampled in the cross-line direction. Therefore, we can safely apply the summation in the in-line direction first and get a partially-stacked MCG (PSMCG).

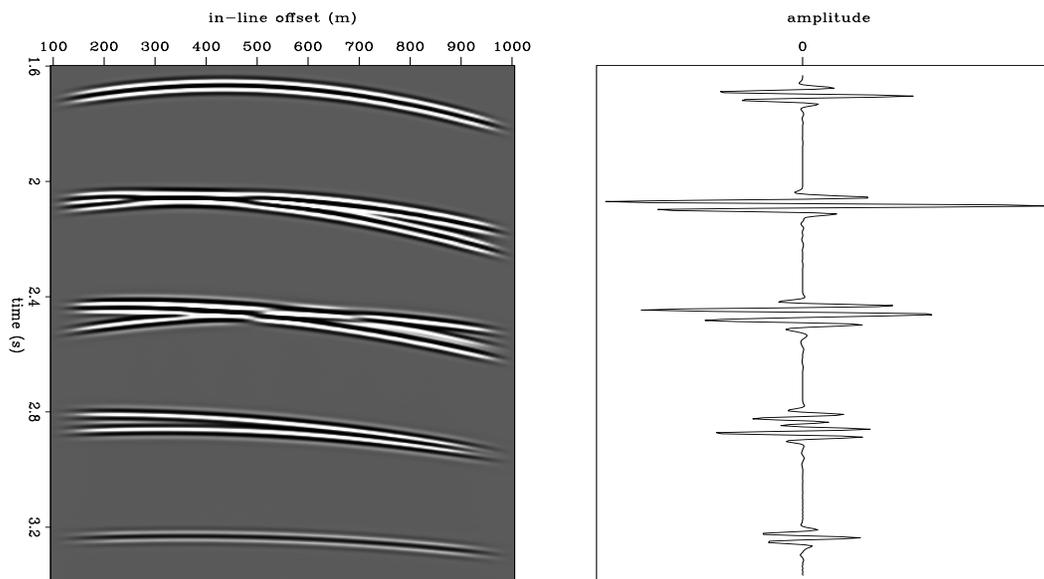


Figure 5: Left: a 2-D MCG. The amplitude has been tapered off to avoid edge effects. Crossing events imply that the multiples are split into different branches. Right: the stacking of the MCG. The locations of predicted multiples correspond to the tops of the curved events.

yalei1-mcg-2d [ER]

Figure 6 displays two PSMCGs with different sampling intervals and the corresponding stacked multiples. Unfortunately, a brutal summation in the sparsely sampled cross-line direction introduces a large amount of aliasing noise into the predicted multiple. The next section describes a method of avoiding such noise.

ANTI-ALIASING IN THE MULTIPLE PREDICTION

The multiple prediction proposal discussed in the preceding section suggests that we can estimate 3-D multiples without trace interpolation. However, as Figure 6 shows, the other problem—aliasing noise—has to be dealt with carefully if there is no missing-streamer interpolation. Like any other Kirchhoff-style operation, anti-aliasing is an important issue in multiple prediction. This issue deserves even more attention in three dimensions, since the cross-line sampling is more sparse than the in-line sampling.

The 3-D estimation of a multiple trace is achieved by stacking a 3-D MCG. As discussed in the preceding section, we can safely stack the 3-D MCG into a 2-D PSMCG along the in-line direction. In the cross-line direction, we must sample the PSMCG more densely to avoid the aliasing noise. Therefore, I propose interpolating the PSMCG directly and then stack it into a multiple trace.

We can interpolate the aliased data in either the F-X (Spitz, 1991) or the T-X domain (Claerbout, 1992). I have chosen the time-space domain multi-scale prediction-error filter (MSPEF) theory discussed in Section 8.4 of Claerbout (1992) to interpolate the PSMCG. The basic idea of the theory is that large objects often resemble small objects. Supposing that we have input data with alternate missing traces, we can estimate a PEF with the following shape:

$$\begin{array}{cccccc} a & \cdot & b & \cdot & c & \cdot & d & \cdot & e \\ \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \end{array} \quad (1)$$

Then we can make the filter smaller by throwing away the zeros (represented by dots) in filter (1) to get

$$\begin{array}{ccccc} a & b & c & d & e \\ \cdot & \cdot & 1 & \cdot & \cdot \end{array} \quad (2)$$

which has the same dip characteristics as filter (1).

Figure 7 shows two PSMCGs containing crossing events, before and after interpolating the alternative missing traces and the corresponding stacking results. The aliasing noise has decreased significantly after trace interpolation.

NUMERICAL EXAMPLES

Combining the multiple prediction and the PSMCG interpolation, we get a new, practical multiple prediction scheme beyond two dimensions. In this section, two 3-D synthetic datasets with similar acquisition geometries are used to evaluate the new scheme. The corresponding model and acquisition parameters appear in Tables 1 and 2.

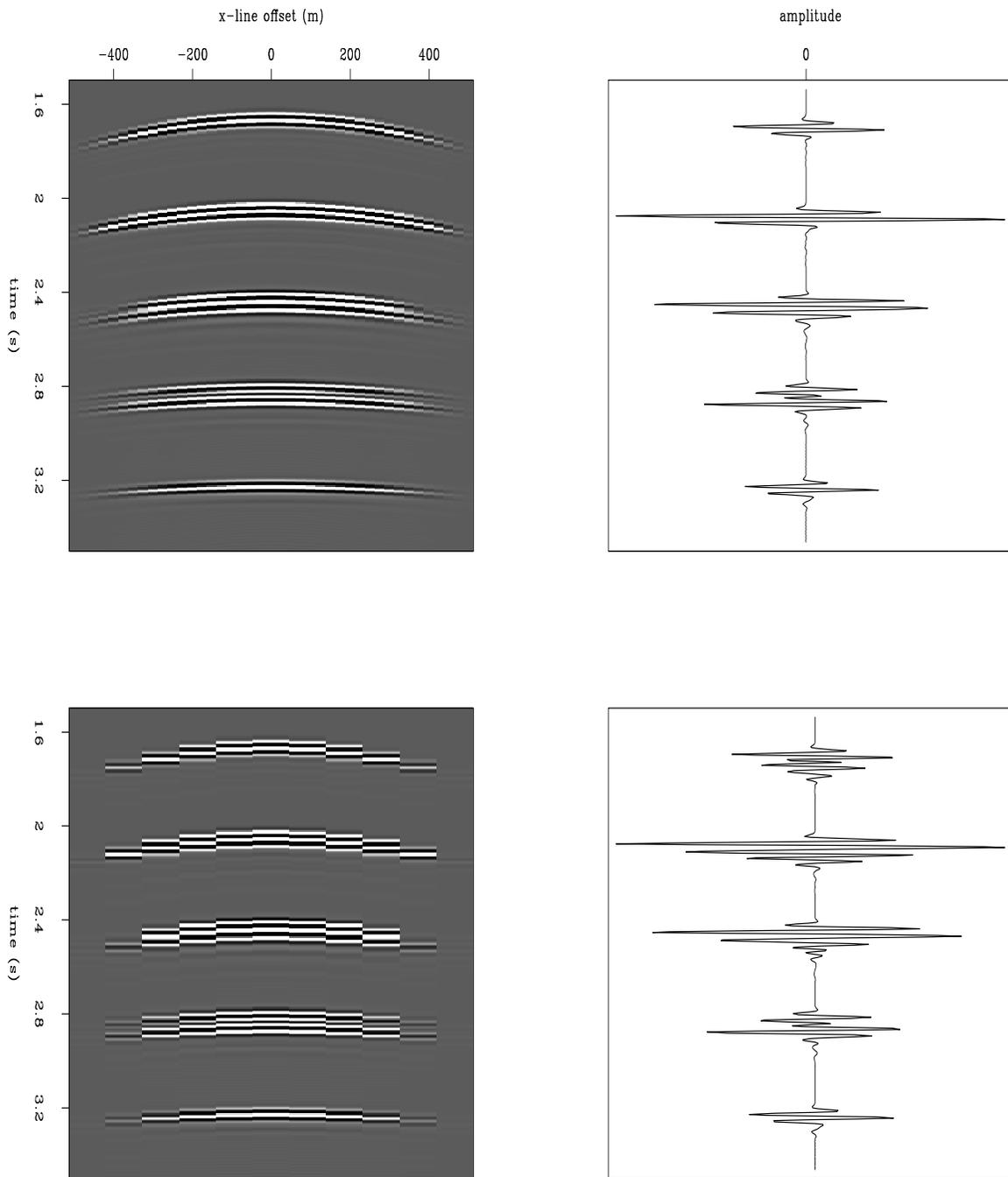


Figure 6: Top: a densely sampled ($\Delta_{\text{streamer}}=25\text{m}$) PSMCG and its stacking result. Bottom: a sparsely sampled ($\Delta_{\text{streamer}}=100\text{m}$) PSMCG and its stacking result. Stacking of the bottom PSMCG introduces aliasing noise to the multiple trace, especially to the top two wavelets in the plot. `yalei1-mcg-ps` [CR]

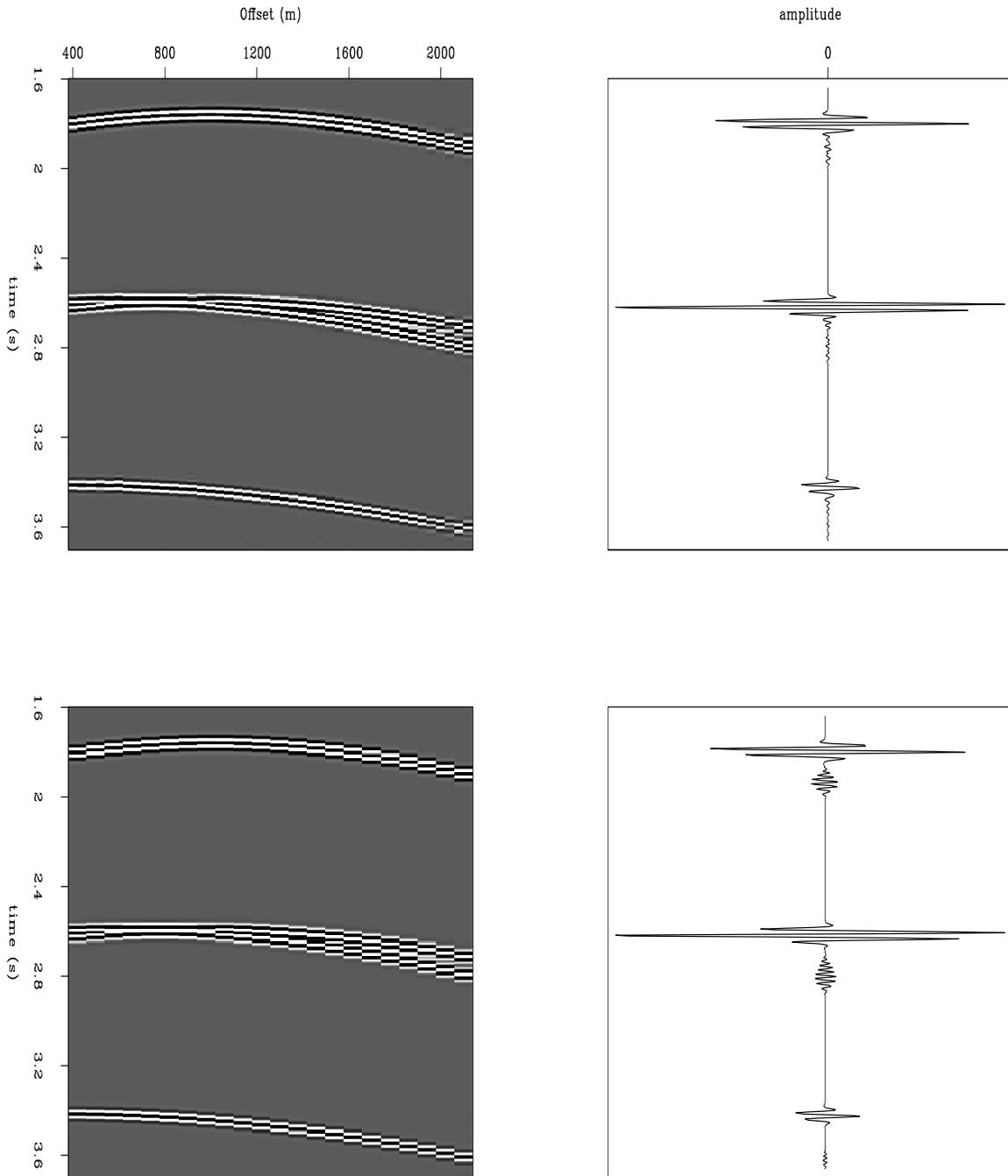


Figure 7: Top: a densely sampled ($\Delta_{\text{streamer}}=25m$) PSMCG and its stacking result. Bottom: a sparsely sampled ($\Delta_{\text{streamer}}=50m$) PSMCG and its stacking result. The aliasing noise has been greatly decreased after the trace interval is halved from $50m$ to $25m$. `yalei1-mcg-interp` [ER]

Model	In-line dip ($^{\circ}$)	X-line dip ($^{\circ}$)	Layer velocity (km/s)
A	(0.0,5.0,10.0)	(0.0, 0.0, 0.0)	(1.5, 2.0, 2.5)
B	(0.0,5.0,10.0)	(0.0, -5.0, 5.0)	(1.5, 2.0, 2.5)

Table 1: Model parameters of Models A and B

Model	Δ_{source} (m)	Δ_{receiver} (m)	Δ_{streamer} (m)
A	20	20	100
B	20	20	50

Table 2: Acquisition parameters of Models A and B

Model A is designed so that the shotline and streamers are deployed along the in-line dip direction. Therefore, there is no approximation error in our approach. With 11 streamers covering from $-500m$ to $+500m$ and a $100m$ streamer interval in the cross-line direction, this is a wide azimuth survey.

For a given shot location, Figure 8 shows the ideal multiple gather and the predicted one. Each 3-D MCG cube in this example is first stacked along the in-line direction into a 2-D PSMCG containing at most 11 traces sampled at intervals of $100m$. The 2-D PSMCG is then interpolated in the cross-line direction to be sampled at $25m$ intervals. The interpolated PSMCG is further stacked into a trace.

Model B is a relatively narrow azimuth survey that still contains 11 streamers covering from $-250m$ to $+250m$ at $50m$ streamer intervals in the cross-line direction. The bottom two reflectors in the model have the opposite cross-line angles, which inevitably introduce approximation error into the estimation. However, as shown in Figure 9, as long as the cross-line dips have no dominant direction, the error can possibly be compensated for in the subtraction step.

CONCLUSIONS AND FUTURE WORK

This paper has proposed a method of predicting 3-D multiples in multi-streamer geometry that does not require massive missing-streamer interpolation. Two numerical examples suggest that this approach can be used in the multi-streamer survey if the cross-line dip is mild. I am now working on the subtraction step. My future work will focus on evaluating and improving this approach with real data.

REFERENCES

- Berkhout, A. J., and Verschuur, D. J., 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations and examples: *Geophysics*, **62**, no. 5, 1586–1611.

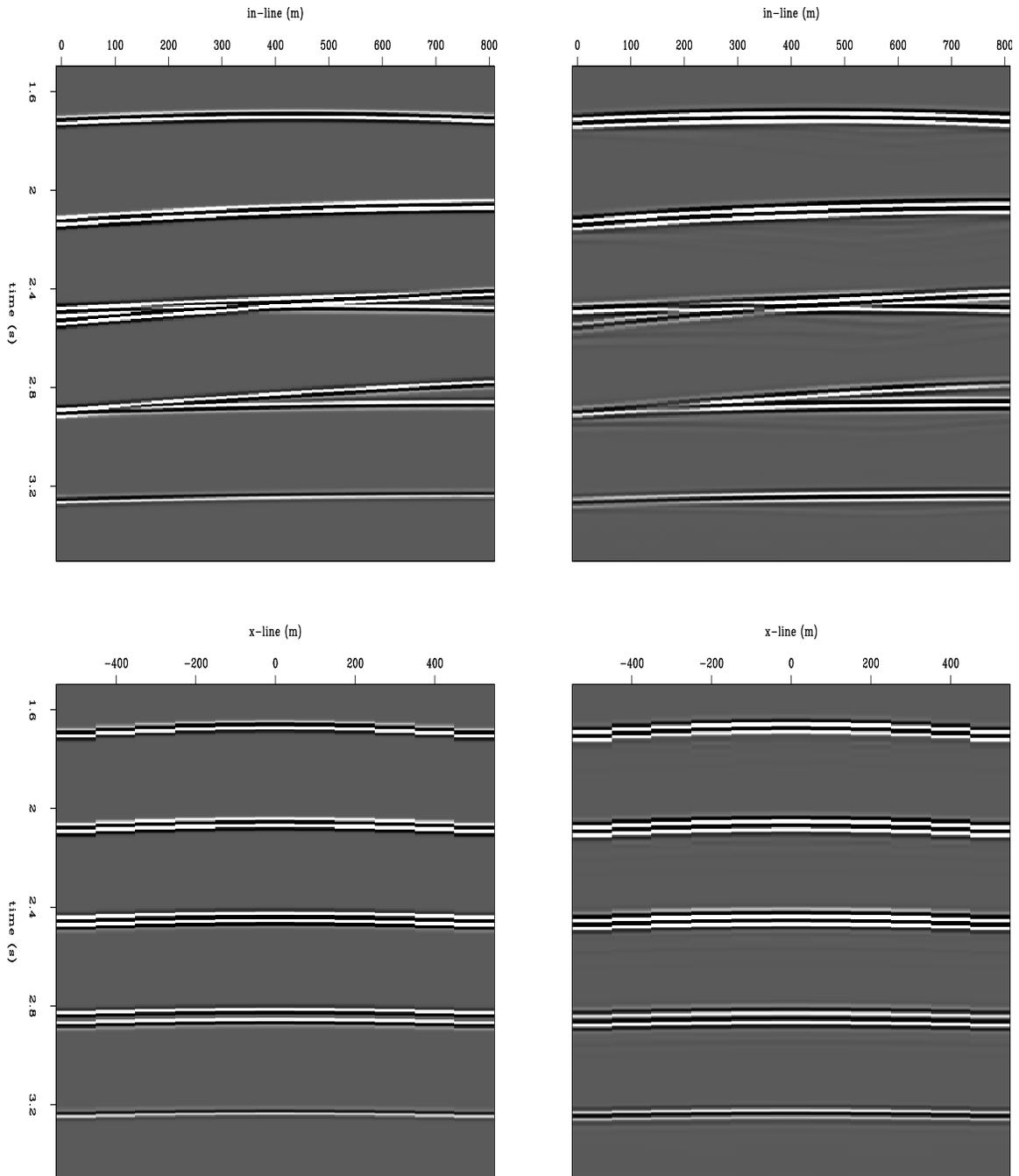


Figure 8: An in-line dip model. Left: the ideal multiple reflection. Right: the estimated multiple reflection. `yale1-wa2d-mult` [CR]

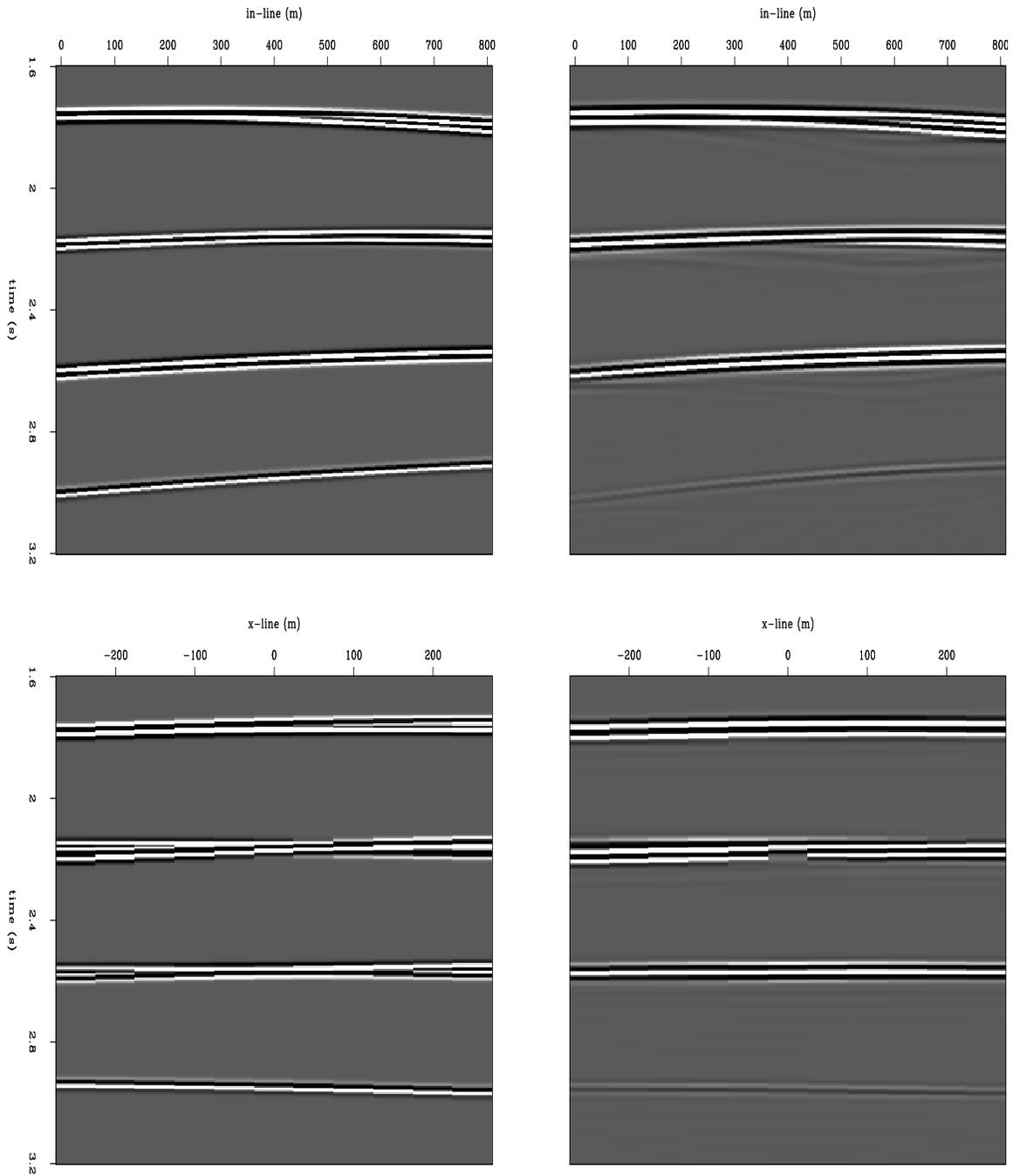


Figure 9: A mixing dip model. Left: the ideal multiple reflection. Right: the estimated multiple reflection. `yale1-na3d-mult` [CR]

- Claerbout, J. F., 1992, *Earth Soundings Analysis: Processing Versus Inversion*: Blackwell Scientific Publications.
- Dragoset, W. H., and Jeričević, Ž., 1998, Some remarks on surface multiple attenuation: *Geophysics*, **63**, no. 2.
- Ross, W. S., Yu, Y., and Gasparotto, F. A., 1997, Multiple suppression: Beyond 2-D. Part II: Application to subsalt multiples: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1391–1394.
- Ross, W. S., 1997, Multiple suppression: Beyond 2-D. Part I: Theory: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1387–1390.
- Spitz, S., 1991, Seismic trace interpolation in the f-x domain: *Geophysics*, **56**, no. 6, 785–794.
- van Dedem, E. J., and Verschuur, D. J., 1998, 3-D surface-related multiple elimination and interpolation: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1321–1324.
- Verschuur, D. J., and Prein, R. J., 1999, Multiple removal results from Delft University: *The Leading Edge*, **18**, no. 1, 86–91.
- Verschuur, D. J., and Berkhout, A. J., 1997, Estimation of multiple scattering by iterative inversion, Part II: Practical aspects and examples: *Geophysics*, **62**, no. 5, 1596–1611.

Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring

James Rickett and Jon Claerbout¹

ABSTRACT

The acoustic time history of the sun's surface is a stochastic (t, x, y) -cube of information. Helioseismologists cross-correlate these noise traces to produce impulse response seismograms, providing the proof of concept for a long-standing geophysical conjecture. We pack the (x, y) -mesh of time series into a single super-long one-dimensional time series. We apply Kolmogoroff spectral factorization to the super-trace, unpack, and find the multidimensional acoustic impulse response of the sun. State-of-the-art seismic exploration recording equipment offers tens of thousands of channels, and permanent recording installations are becoming economically realistic. Helioseismology, therefore, provides a conceptual prototype for using natural noises for continuous reservoir monitoring.

INTRODUCTION

The earth and the sun both have a noisy surface. This acoustic noise generates sonic waves that dive into the sphere and emerge at all distances. A process that we call "acoustic daylight imaging²" enables us to form various pictures of the interior.

Acoustic daylight imaging: previous work

At the Stanford Exploration Project, we began our interest in acoustic daylight imaging many years ago when theory (Claerbout, 1976) showed that under certain idealized conditions, the autocorrelation of an earthquake seismogram should mimic an echo sounding like those made with explosives for petroleum prospecting.

Early attempts to verify this theory in practice quickly failed and we came to realize that the essential physical feature is real-world three dimensionality while both our data and our mathematical theory were merely one dimensional. We didn't have a three-dimensional theory, but we did have a conjecture:

By cross-correlating noise traces recorded at two locations on the surface, we

¹email: james@sep.stanford.edu, claerbout@stanford.edu

²The term "acoustic daylight imaging" is known in oceanography where measurements are not made at the free surface. This difference leads ocean acoustics away from seismology.

can construct the wavefield that would be recorded at one of the locations if there was a source at the other.

Cole (1995, 1988) initially tried to verify this conjecture on a passive 3-D survey recorded with an array of 4056 geophones covering more than a half kilometer square on the Stanford campus. Unfortunately, however, only twenty minutes of passive seismic data was recorded, and beam steering showed ambient noise was predominately incident from only one direction. His cross-correlation results were not conclusive. The proximity of the San Andreas fault makes the Stanford area difficult to analyze, and we were also troubled by poor coupling between the geophones and the dry summer soil.

Nobody expected the geophones to record plunging waves from great distances but that is exactly what happened. We saw seismic waves apparently coming from the American Midwest. Earthquake seismologists were surprised to learn that we could receive seismic waves from so far at such high frequency (10 Hz) because with their small numbers of seismometers they cannot. Unfortunately, we were not able to observe what we sought, the much smaller scale reflected waves that we would crosscorrelate within our array. Such waves would illuminate the area within drilling distance so proof of concept would interest our sponsors.

Modeling studies (Rickett and Claerbout, 1996) showed that longer time-series, and a white spatial distribution of random noise events would be necessary for the conjecture to work in practice.

Helioseismology

In 1995, solar physicists developed a new instrument for studying the sun. The Michelson Doppler Imager (MDI) instrument measures the Doppler shift of solar absorption lines formed in the lower part of the solar atmosphere. This provides line-of-sight velocity measurements for points on the sun's surface that can be used to study solar oscillations. This amounts to having a million (1024×1024) seismometers uniformly distributed on the surface of the sun. Furthermore, the solar seismologists are able to zoom their lens to reposition their million virtual seismometers to give them a magnified view anywhere they choose.

Most helioseismology (e.g. Kosovichev, 1999) has been done in the frequency domain with spherical harmonic functions. Spherical harmonics provide an excellent tool for studying the whole sun at one time. However, small-scale events are only described by harmonic modes of very high-order. Spherical harmonic functions are therefore inefficient for studying small, localized areas of the sun's surface.

Solar seismologists (Duvall et al., 1993) had also come up with the idea of creating 'time-distance' seismograms by crosscorrelating surface noise observations to mimic impulsive sources on the solar surface. They were successful with real data in three dimensions on the sun, before we could do it on earth.

Convective flow in the outer third of the sun leads to a breakdown in reciprocity of time-distance seismograms derived by cross-correlation. Helioseismologists have used this break-

down in reciprocity to estimate the three-dimensional flow velocity structure in the outer third of the sun.

Permanent recording installations

Ebrom et al. (1998) discuss the economic justification for installing permanent recording installations in marine environments over producing hydrocarbon fields. The incremental cost of additional 3-D seismic surveys is much less for permanent monitoring systems than for conventional repeated 3-D. The quality of ocean bottom cable data has also increased dramatically in recent years, and so permanent monitoring systems also provide a higher level of repeatability.

For these reasons, permanent installations are already installed in many basins worldwide. In the future, explorationists may take advantage of continual recordings and acoustic daylight imaging to continually monitor reservoir production.

RAW DATA

The Solar Oscillations Investigation project (<http://soi.stanford.edu>) provided us with a cube of data recorded by their MDI instrument. They transformed the coordinate system to Cartesian coordinates by projecting high-resolution data from an area approximately 18° square onto a tangent plane.

Figure 1 shows a time slice through a cube of raw velocity data from the MDI. The object in the center of the panel is a sun-spot. Figure 2 shows a time-distance section through the same cube. The sampling spacing is 1 minute on the time-axis and approximately 68 km on the two spatial axes.

Events on Figure 2 fall into two distinct spectral windows. The low temporal frequency events (<1.25 mHz) are related to solar convection (super-granulation), while the higher frequency events are related to acoustic wave propagation. We were interested in studying acoustic wave phenomena; so as a preprocessing step, we removed the lower frequency spectral window by applying a $\frac{1-Z}{1-\rho Z}$ low-cut filter to the data.

3-D KOLMOGOROFF SPECTRAL FACTORIZATION

A simple linear model for the observed solar oscillations consists of a convolution of a source function with the impulse response of the sun's surface. The source function is stochastic in nature, and may be characterized as being spectrally white in time and space with random phase. The impulse response contains the spectral color, and is commonly a minimum-phase function.

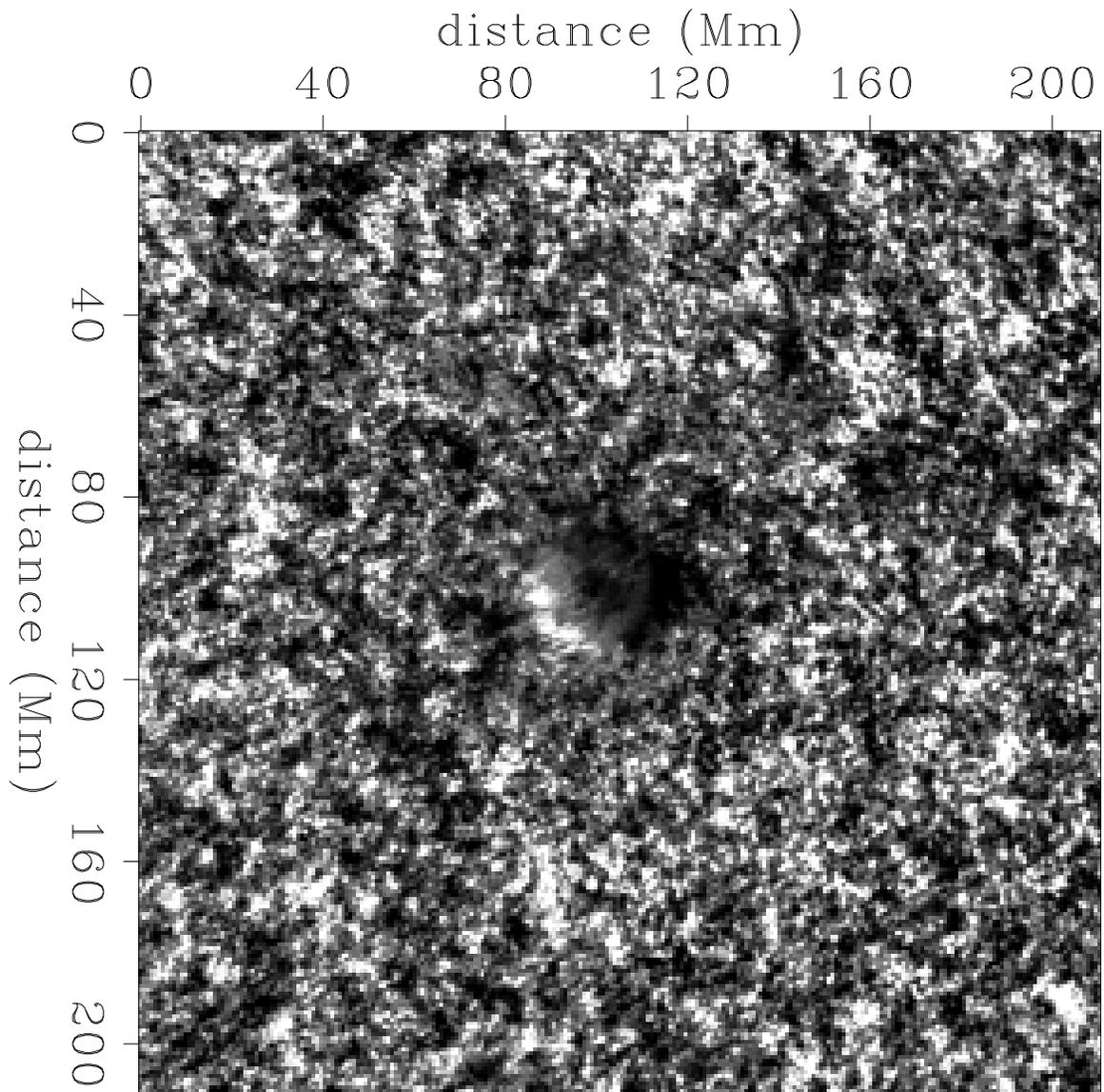


Figure 1: Time-slice through the raw data. The object in the center is a sun-spot. Distances are in Megameters (thousands of kilometers), and time units are kiloseconds. `james1-timedata` [CR]

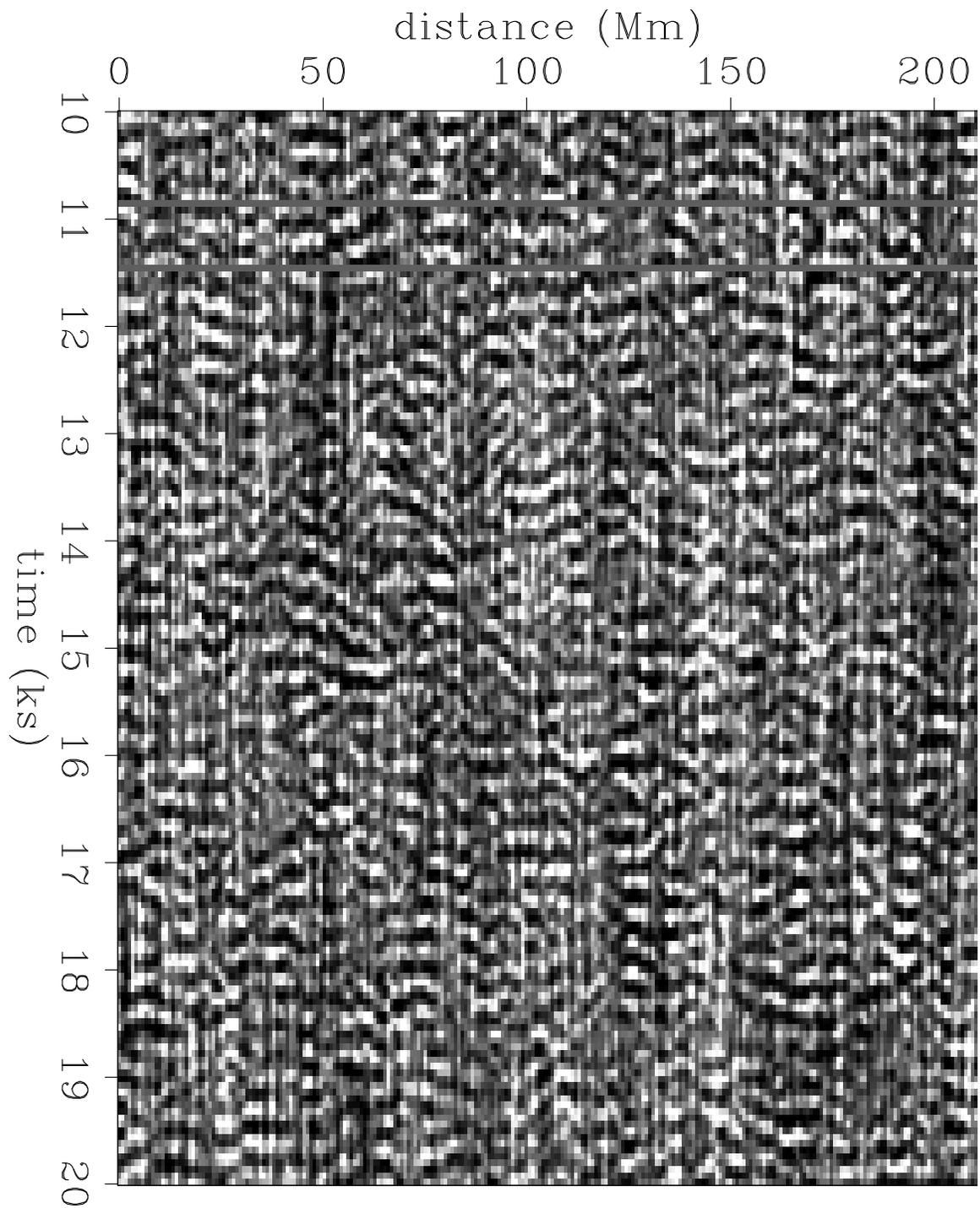


Figure 2: Time-space section through a subset of the raw data. `james1-txdata` [CR]

If this model holds true, then estimating the source function reduces to estimating a minimum-phase function with the same (w, k_x, k_y) spectrum as the original data: multi-dimensional spectral factorization.

Helical boundary conditions (Claerbout, 1998) provide a framework for converting a multi-dimensional problem into an equivalent problem in only one dimension, and allow us to solve the three-dimensional spectral factorization problem efficiently.

We perform the spectral factorization rapidly in the frequency domain in three steps. Firstly, we transform the multi-dimensional signal to an equivalent one-dimensional signal using helical boundary conditions. Secondly, we perform one-dimensional spectral factorization with Kolmogoroff's (1939) algorithm. Finally, we remap the impulse response back to three-dimensional space. We reduce wrap-around effects by padding the spatial axes.

Figure 3 shows the impulse response derived from Kolmogoroff spectral factorization as a function of radial distance from the impulse. It looks very similar to the cross-correlation time-distance seismogram shown in Figure 4, and those displayed by (Kosovichev, 1999). However, for the dataset described above, this operation was approximately twenty times faster than cross-correlating every trace in either (w, \mathbf{x}) or (t, \mathbf{x}) . The speed-up becomes apparent when you consider that cross-correlating every trace with every other trace requires $O(N_x^2 N_y^2 N_t)$ operations, whereas one-dimensional spectral factorization requires only $O(N \log N)$ operations where $N = N_x N_y N_t$.

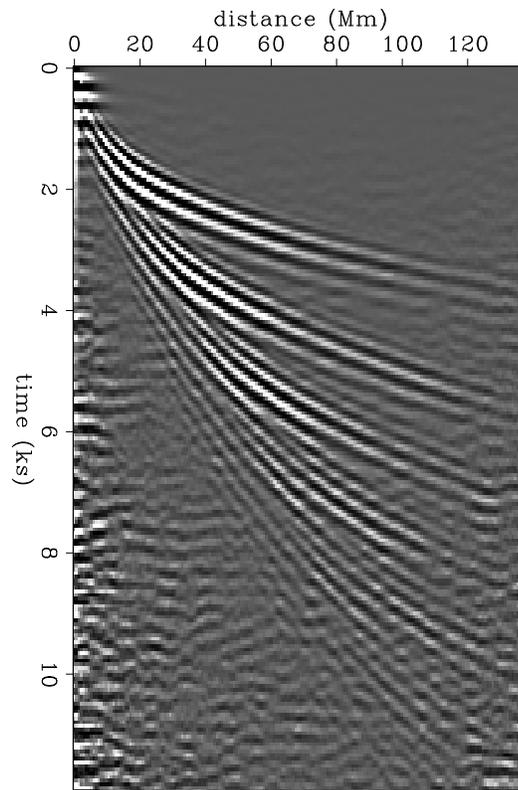
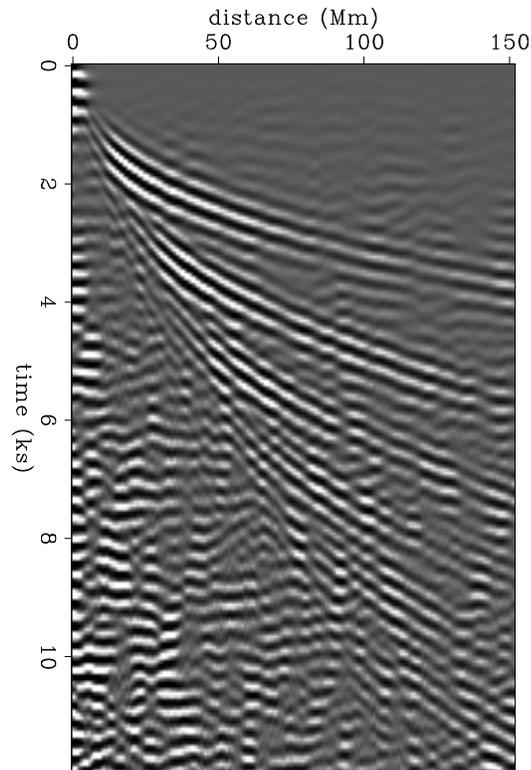


Figure 3: Impulse response derived by Kolmogoroff spectral factorization binned as a function of radial distance from the impulse. james1-kolstack
[CR]

Figure 4: Impulse response derived by cross-correlation binned as a function of radial distance from the impulse. More noise is present in this Figure compared to Figure 3 because less data was used in the calculation.

`james1-xcorr` [CR]



INTERPRETATION

The lack of sharp velocity contrasts mean there are no clearly observable reflections in the impulse response. The visible events are diving waves (refractions) of increasing order. The first arrival is the the direct wave that would be characterized as the ‘P’ arrival in terrestrial geophysics. The second arrival is equivalent to a ‘PP’ event (or first order multiple) having bounced once on the solar surface. About six distinct arrivals are visible in Figure 3, corresponding to multiples up to fifth-order.

Since there are no distinct velocity interfaces, only a smooth velocity gradient, the head waves continually curve upward. The contrast between this behavior and head waves on terrestrial exploration seismograms is illustrated in Figure 5.

Radial trace analysis

The radial trace domain provides a natural way to view the seismograms. Since multiples arrive at integer multiples of primary traveltime and distance. The radial trace domain becomes periodic. Rather than preserving traveltimes in the radial trace transform, and stacking all azimuths together, we preserve x and y coordinates.

Figure 6 shows a constant velocity slice through the Kolmogoroff impulse response. Multiple bounces produce the series of concentric circles. Breakdown in reciprocity due to con-

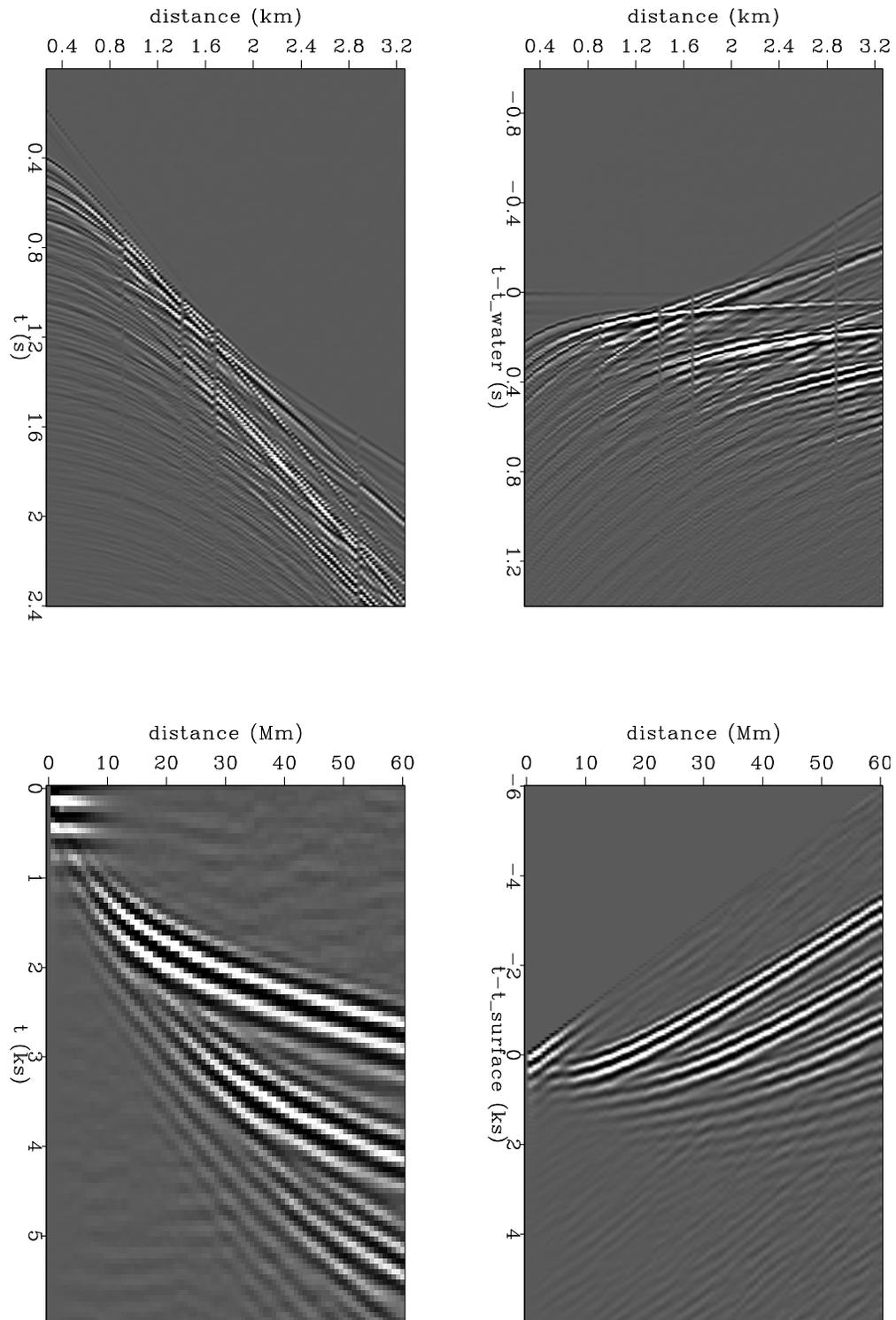
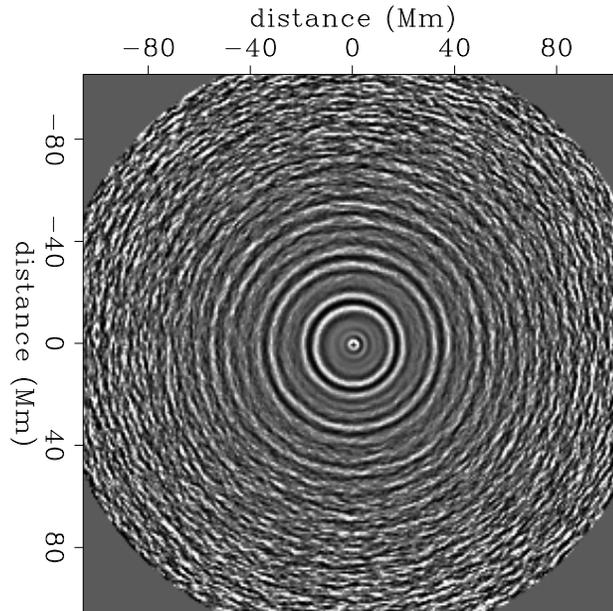


Figure 5: The top two panels show a shot gather from the Gulf of Mexico before (left) and after (right) linear moveout with the water velocity. The lower two panels show the solar impulse response before (left) and after (right) linear moveout at 10 km/s. james1-lmo [CR]

vective flow in the sun, will manifest itself as azimuthal anisotropy in this domain.

Figure 6: Constant velocity slice through radial transform.
james1-radialcube [CR]



CONCLUSIONS

Helioseismology validates a long-standing geophysical conjecture that the cross-correlation of noise traces may provide impulse response seismograms.

Additionally, we have shown that combining one-dimensional Kolmogoroff spectral factorization algorithms with helical boundary conditions allows us to calculate the acoustic impulse response an order of magnitude faster than cross-correlation in either (t, x, y) or (ω, x, y) .

This is all very interesting, but our sponsors trust that we are looking towards matters that can have significant practical and financial implications. Not long ago that would have seemed far fetched. But electronics and communications revolutions have been revolutionizing seismology for decades. Many oil fields are already instrumented with multiple hundreds of geophones. We are now prepared to evaluate the prospects for the multiple tens of thousands of geophones that are already practical.

REFERENCES

- Claerbout, J., 1976, Fundamentals of geophysical data processing: Blackwell.
- Claerbout, 1998, Multidimensional recursive filters via a helix: *Geophysics*, **63**, 1532–1541.
- Cole, S., 1988, Examination of a passive seismic dataset using beam steering: SEP-57, 417–426.

- Cole, S., 1995, Passive seismic and drill-bit experiments using 2-D arrays: Ph.D. thesis, Stanford University.
- Duvall, T. L., Jefferies, S. M., Harvey, J. W., and Pomerantz, M. A., 1993, Time-distance helioseismology: *Nature*, **362**, 430–432.
- Ebrom, D., Krail, P., Ridyard, D., and Scott, L., 1998, 4-C/4-D at Teal South: The Leading Edge, **17**, no. 10, 1450–1453.
- Kolmogorov, A. N., 1939, Sur l'interpolation et l'extrapolation des suites stationnaires: *C.R. Acad.Sci.*, **208**, 2043–2045.
- Kosovichev, A., 1999, Inversion methods in helioseismology and solar tomography: submitted to Elsevier Preprint.
- Rickett, J., and Claerbout, J., 1996, Passive seismic imaging applied to synthetic data: *SEP-92*, 83–90.

Interpolation with smoothly nonstationary prediction-error filters

Sean Crawley¹

ABSTRACT

Building on the notions of time-variable filtering and the helix coordinate system, I develop software for filters that are smoothly variable in multiple dimensions, but that are quantized into large enough regions to be efficient. Multiscale prediction-error filters (PEFs) can estimate dips from recorded data and use the dip information to fill in unrecorded shot or receiver gathers. The data are typically divided into patches with approximately constant dips, with the requirement that the patches contain enough data samples to provide a sufficient number of fitting equations to determine all the coefficients of the filter. Each patch of data represents an independent estimation problem. Instead, I estimate a set of smoothly varying filters in much smaller patches, as small as one data sample. They are more work to estimate, but the smoothly varying filters do give more accurate interpolation results than PEFs in independent patches, particularly on complicated data. To control the smoothness of the filters. I use filters like directional derivatives that Clapp et al. (1998) call “steering filters”. They destroy dips in easily adjusted directions. I use them in residual space to encourage dips in the specified directions. I describe the notion of “radial-steering filters” (Clapp et al., 1999), i.e., steering filters oriented in the radial direction (lines of constant x/t in (t, x) space). Break a common-midpoint gather into pie shaped regions bounded by various values of x/t . Such a pie-shaped region tends to have constant dip spectrum throughout the region so it is a natural region for smoothing estimates of dip spectra or of gathering statistics (via 2-D PEFs). In this paper I use smoothly variable PEFs to interpolate missing traces, though they may have many other uses. Finally, since noisy data can produce poor interpolation results, I deal with the separation of signal and noise along with missing data.

INTRODUCTION

Claerbout (1998b) describes a helical coordinate to cast multi-dimensional filtering as one dimensional, enabling the use of some well-developed signal processing theory in applications including missing data interpolation (Fomel et al., 1997) and low-cut filtering (Claerbout, 1998a). To account for nonstationarity in the data, missing data interpolation with PEFs is typically done in patches or gates where dips are assumed to be approximately stationary (Spitz, 1991). Each patch constitutes an independent problem, though they may overlap. The smaller

¹email: sean@sep.stanford.edu

the patch, the more stationary the data is likely to be within the patch; however, there is a lower limit on the patch size, because a patch must contain enough data to provide fitting equations for all the filter coefficients. Claerbout (1997) describes a method for estimating smoothly time-varying PEFs without patching. The helix extends the idea of smooth time-variable PEF estimation to smooth time- and space-variable PEF estimation. The smoothly-variable PEFs can perform better at interpolating missing data than PEFs estimated in independent patches.

Clapp et al. (1998) show how to use space-variable inverse steering filters to smooth in adjustable directions, and they show how to solve empty-bin problems filling in missing data along the directions of the steering. I use space-variable steering filters to control the direction of smoothness between PEFs. I orient the steering filters radially in a CMP gather to encourage PEFs to have the same dip information along lines of constant x/t , where data tends to have constant dip spectra. In this paper I review the theory for estimating smoothly varying PEFs, and show examples of their application to missing data interpolation. I describe an improvement to filter estimation for CMP gathers using “radial-steering filters.” Finally, I add the notion of signal and noise separation for interpolating noisy data.

TIME- AND SPACE-VARYING PEFs

The time dip of seismic data changes rapidly along many axes, so a single PEF can only represent a small amount of data. Often we divide the data into patches, where it is assumed the data have constant dips. Because seismic data have curvature and may not be well represented by piecewise-constant dips, it is appealing to extend the idea of time-variable filtering to include spatial dimensions as well, and have smoothly varying PEFs to represent curved events.

I decrease the patch size, to as small as a single data sample, changing the problem from overdetermined to very underdetermined. It is possible to estimate all these filter coefficients by the usual formulation, supplemented with some damping equations, say

$$\begin{aligned} \mathbf{0} &\approx \mathbf{YKa} + \mathbf{r}_0 \\ \mathbf{0} &\approx \epsilon \mathbf{Ra} \end{aligned} \tag{1}$$

where \mathbf{R} is a roughening operator, \mathbf{Y} is convolution with the data, and \mathbf{K} is a known filter coefficient mask.

When the roughening operator \mathbf{R} is a differential operator, the number of iterations can be large. To speed the calculation immensely and make the equations somewhat neater, we can “precondition” the problem. Define a new variable \mathbf{p} by $\mathbf{a} = \mathbf{Sp}$ and insert it into (1) to get

$$\mathbf{0} \approx \mathbf{YKSp} + \mathbf{r}_0 \tag{2}$$

$$\mathbf{0} \approx \epsilon \mathbf{RSp} \tag{3}$$

Now, because the smoothing and roughening operators are somewhat arbitrary, we may as well replace \mathbf{RS} by \mathbf{I} and get

$$\mathbf{0} \approx \mathbf{YKSp} + \mathbf{r}_0 \tag{4}$$

$$\mathbf{0} \approx \epsilon \mathbf{Ip} \tag{5}$$

We solve for \mathbf{p} using conjugate gradients. To see \mathbf{a} , just use $\mathbf{a} = \mathbf{S}\mathbf{p}$. To simplify things, one could just drop the damping (5) and keep only (4); then to control the null space, start from a zero solution and limit the number of iterations. For \mathbf{S} we can use polynomial division by a Laplacian or by filters with a preferred direction. If the data are CMP gathers, it is attractive to use radial filters, which are explained later.

INTERPOLATING MISSING TRACES

We estimate missing data in two steps of linear least squares (Claerbout, 1992). The first step is estimation of PEFs. After the PEFs have been estimated they are used to fill in the empty trace bins. This is the second step of least squares. We want the recorded and estimated data to have the same dips. Since the dip information is now carried in the PEFs, this is once again specifying that the convolution of the filter and data should give the minimum output, except that now the filters are known and the data is unknown. We constrain the data by specifying that the originally recorded data cannot change. To separate the known and unknown data we have a known data selector \mathbf{K} and an unknown data selector \mathbf{U} , with $\mathbf{U} + \mathbf{K} = \mathbf{I}$. These multiply by 1 or 0 depending on whether the data were originally recorded or not. With \mathbf{A} signaling convolution with the PEF and \mathbf{y} the vector of data, the regression is $0 \approx \mathbf{A}(\mathbf{U} + \mathbf{K})\mathbf{y}$, or $\mathbf{A}\mathbf{U}\mathbf{y} \approx -\mathbf{A}\mathbf{K}\mathbf{y}$.

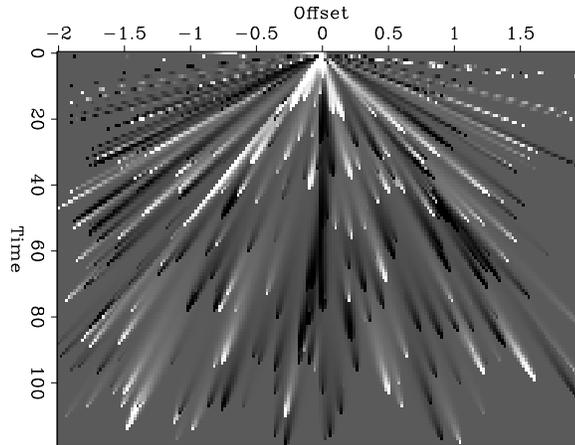
While a PEF at every sample works well for destroying the data, it is not the best choice for reconstructing it; interpolation with PEFs estimated at every data point gives poor results and requires extravagant memory allocation. One answer is just that zero is not the correct value of ϵ in (5); but we can greatly improve the results and decrease the memory usage without adding equations, by using very small patches, such as $2 \times 2 \times 2$; small enough that the assumption of stationarity within a patch is reasonable. This is similar to putting an extra roughener in the damping equation, in that it is essentially an infinite penalty on variations of \mathbf{p} between small groups of samples, and it has the important economizing effect of reducing the memory allocation. In the method where the patches are independent (Crawley, 1998), the number of filter coefficients puts a lower bound on patch size; the problem has to stay well overdetermined to produce a useful PEF. Using smoothly varying filters effectively reduces the minimum patch size, so that the filter estimation problem can be underdetermined, and still produce useful PEFs.

Radial Smoothing

Clapp et al. (1998) show how to control the direction of smoothing. In certain cases, it may make sense to specify some preferred direction of filter smoothing. For instance, CMP gathers tend to have approximately constant dip spectra at constant values of x/t , which correspond to radial lines. So it makes sense to arrange patches and smooth filter coefficients in the radial direction on a CMP gather, to accelerate convergence and get good results with as few coefficients as possible. Very small patches are desirable where the data has the most curvature, which tends to be at smaller times and offsets. At larger times and offsets, however, events in

CMP gathers tend to be near their asymptotes, and much more linear. Smoothing and patching in radial coordinates has the pleasing result that the largest patches fall at far offsets and late times, where they are most appropriate. Figure 1, which comes from Clapp (1999), shows randomly scattered points smoothed with radial-steering filters.

Figure 1: Radial smoothing. Panel shows result of smoothing random scattering of dots with the adjoint radial steering operator. The forward operator points out from the origin. Figure borrowed from Clapp (1999).
sean1-bob [NR]



Data Example

Some data from a 3-D marine survey are shown in Figure 2. The top left panel shows part of a shot gather, and the top right panel shows a time slice through several shots from the same source and the same streamer. Marine 3-D data may be broken up into a set of 2-D surveys or treated as a single 3-D survey; more on that later. In this particular case, the data are fairly routine, but they can serve to illustrate the method. The bottom half shows the output of removing half the shots and then interpolating them.

More than two sources

In the previous example I interpolated data collected with two alternating sources. Typical marine geometries have two alternating sources, both placed near the center of the crossline spread, which gives the survey efficient crossline midpoint coverage. It is possible to imagine desiring more than two sources, distributed more or less evenly across the crossline spread. If the geology is complex (salt bodies), and there are genuinely 3-D multiples, then even the best 2-D multiple suppression method will leave multiple events behind. Sources only near the center of the crossline spread prevents the reciprocity necessary for a method like the Delft SRME (van Borselen et al., 1991) method, or multiple prediction by upward continuation, to predict the multiple energy moving in the crossline direction. Using more sources, spread more widely across the crossline, means the recorded data would contain more crossline information that could be used to model 3-D multiple reflections. On the other hand, cycling through more than two sources naturally creates larger gaps in the acquisition and more opportunities for the data to be aliased. In the previous example, I could throw away lots of data and restore it pretty faithfully, because the data is very predictable. In time slice view, it is almost

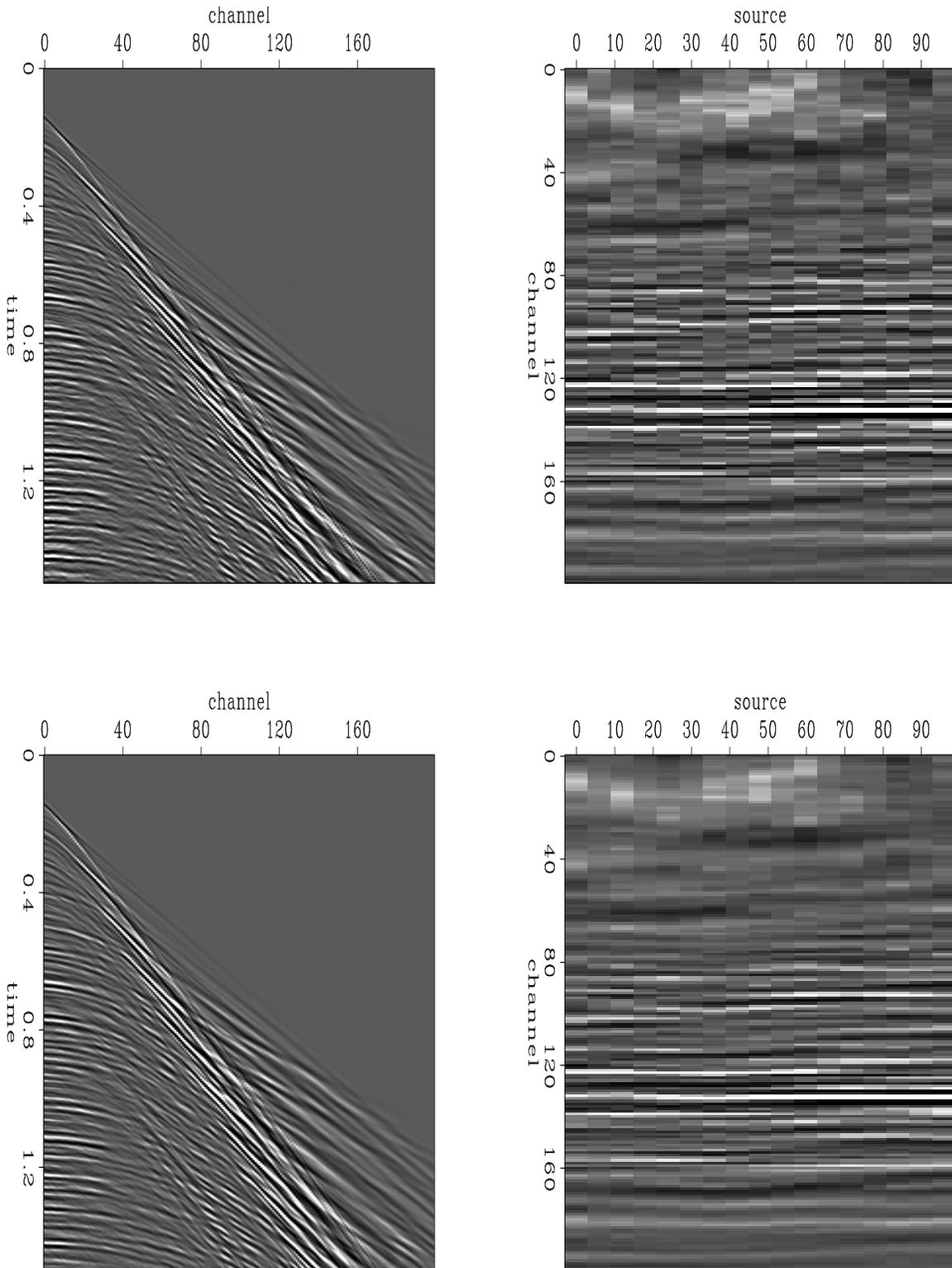


Figure 2: Example input data. The top left panel is part of a shot gather from a 3-D survey, the top right panel is a time slice through several shots from the same source into the same streamer. Bottom panels are the same, after zeroing every other gather and interpolating. `sean1-curtFig1` [CR]

entirely straight lines. The geology in that example is flat, so it is probably not necessary to think about multiples in three dimensions. The next example is from a marine survey over much more complex geology, where multiple energy moving in three dimensions could be a genuine concern. It is just a 2-D survey; I simulate one inline from a survey with four sources by throwing away three out of four shot gathers. I then attempt to reconstruct the data. Some results are shown in Figures 3, 4, and 5. In this example, the data are reduced to every fourth shot gather and interpolated back to every second shot gather. If that works, then getting from every second shot to every shot is fairly certain. There is no conceptual reason not to go directly from every fourth shot to every shot in a single step, it just means scaling the axes by four rather than two. There is a practical issue that arises, however, in that filter coefficients are smoothed by a variant of leaky integration, and there tends to be a little too much leakage. Figures 3 and 4 show shot gathers that were removed in the left panels, and the interpolated versions of those shot gathers in the right panels. The results are not bad, but not perfect, either. Surprisingly little energy dipping in towards zero offset is lost; given the particular filter smoothing strategy I use, I expected those events to be lost. However, several events close to zero offset are poorly reconstructed. Figure 5 shows a time slice from the correct (recorded) data in the left panel, the input to the interpolation in the middle, and the interpolation on the right. Again the output is not bad, but not flawless either. The spatial frequencies are too low near zero offset, for instance.

2-D or 3-D

For marine data, there is very little difference between interpolating shots in 2-D survey and interpolating shots in a 3-D survey, except that 3-D data provides some interesting choices. We can treat a single source and a single streamer as a 2-D survey (giving us several 3-D input cubes), or separate the sources but leave the streamers together (4-D input), or just use the whole 5-D input. Results get better with more dimensions, because there are more directions for events to be predictable in; but they only get marginally better when we add the crossline directions in marine data, because there are only a handful of crossline points. The cost is large because a few points worth of zero padding are necessary. Padding the inline offset axis by a few points is a small increase in the data volume, but padding the crossline offset axis by a few points may double the data volume. Newer boats tow more streamers, so it may be worth using the extra dimensions on newer data.

NOISY DATA

It is appealing to think of interpolating land data, because land data can be so expensive to acquire. Land data is more difficult to interpolate than marine data, however, because it tends to be much more noisy. We can guess that the PEFs will find the predictable part of the data, and that it will be the signal. However, even assuming we do estimate the correct PEF (one that captures the dips of the signal rather than the noise), energy from the noise will be carried along those dips to nearby interpolated traces. As an alternative, we can attempt to separate the signal from the noise while interpolating the missing traces. Taking the theory from Claerbout

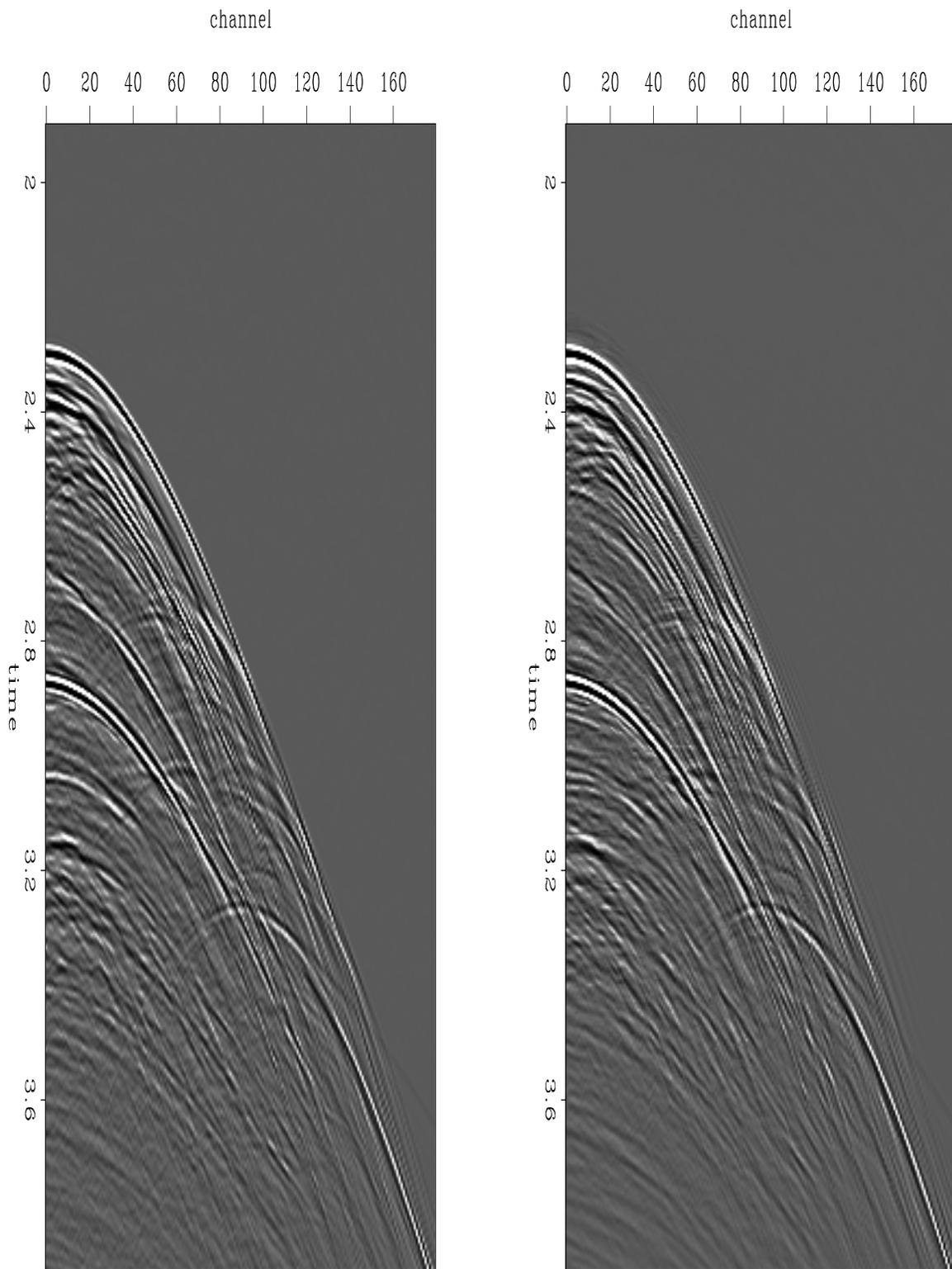


Figure 3: Interpolation test. Left panel is a shot gather that was removed from the input data. Right panel is the interpolated version of that shot gather. [sean1-355Fig2b](#) [CR]

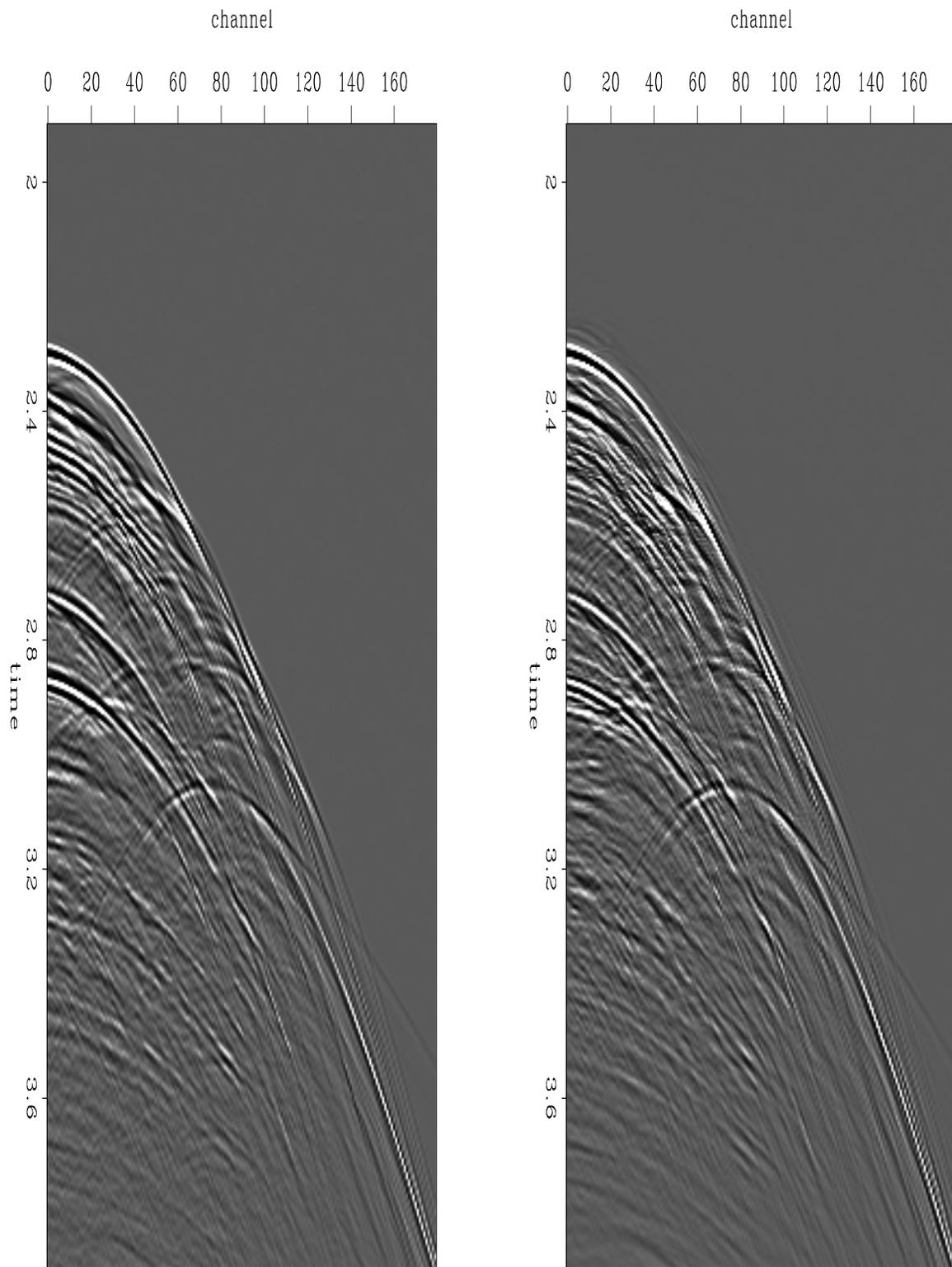


Figure 4: Interpolation test. Left panel is a shot gather that was removed from the input data. Right panel is the interpolated version of that shot gather. [sean1-355Fig2](#) [CR]

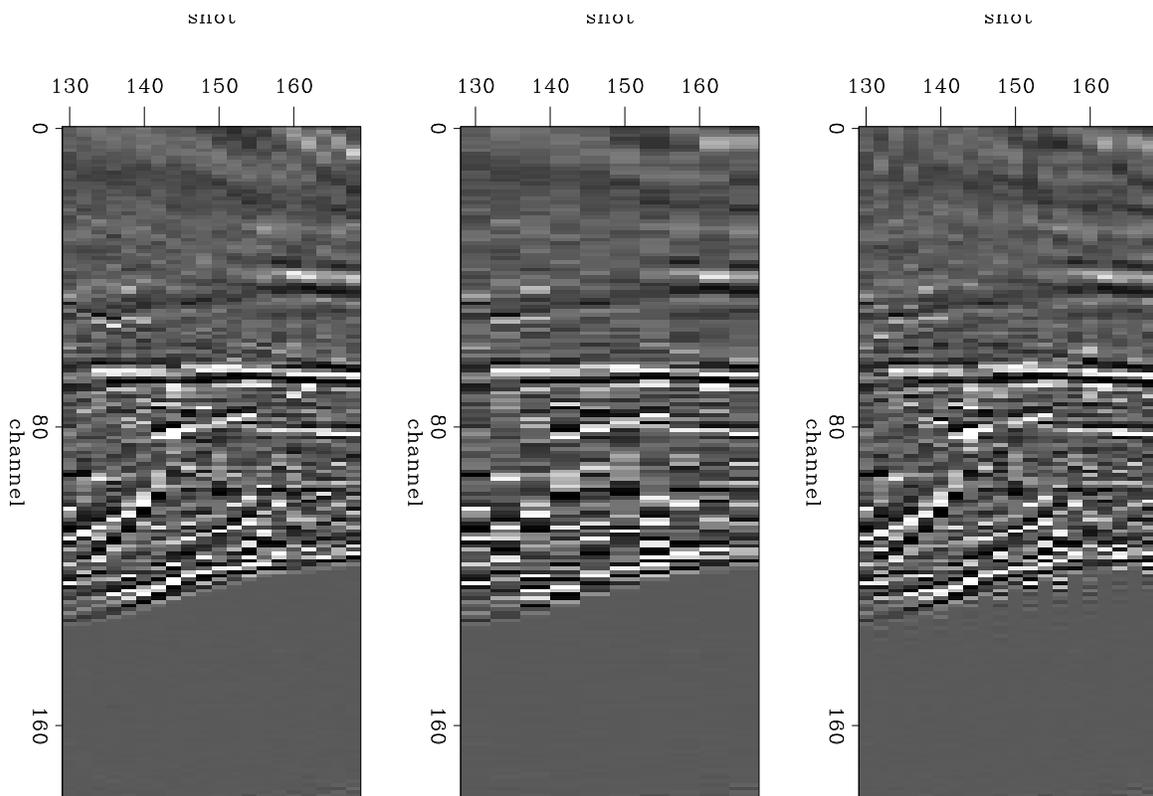


Figure 5: Interpolation test. Left panel shows the correct (recorded) time slice, center panel shows the subsampled input to the interpolation, right panel shows the output. sean1-355Fig1 [CR]

(1997), data \mathbf{d} can be decomposed into known plus missing parts, $\mathbf{d} = \mathbf{k} + \mathbf{m}$. With known and unknown data selectors \mathbf{K} and \mathbf{M} , write the data as

$$\mathbf{d} = \mathbf{Kd} + \mathbf{Mm} \quad (6)$$

where \mathbf{Kd} is zero-padded known data and all the components of \mathbf{m} are freely adjustable. The goal is to fill in the appropriate parts of \mathbf{m} .

The data can also be decomposed into signal plus noise, $\mathbf{d} = \mathbf{s} + \mathbf{n}$. Thus

$$\mathbf{s} + \mathbf{n} = \mathbf{Kd} + \mathbf{Mm} \quad (7)$$

If we write fitting goals for signal and noise and then get rid of the noise with equation (7) we have

$$\mathbf{0} \approx \mathbf{Nn} = \mathbf{N}(\mathbf{Kd} + \mathbf{Mm} - \mathbf{s}) \quad (8)$$

$$\mathbf{0} \approx \mathbf{Ss} = \mathbf{Ss} \quad (9)$$

Solving (8) and (9) gives estimates for the signal component \mathbf{s} of both the known and missing data, as well as an estimate of the noisy missing traces, \mathbf{Mm} . This requires a signal predictor and a noise predictor. For the signal predictor I use the PEFs estimated from the data. In the example to follow I will throw out alternating midpoint gathers and attempt to interpolate them back, so I define noise as whatever is incoherent across midpoints and choose the noise predictor to be an average in a small window along the midpoint axis.

Example

Figure 6 shows a CMP gather from a land seismic survey, and a time slice through several CMP gathers. There is a visible “noise cone” defined by some low velocity. Inside the cone there is very little coherency along the midpoint axis. As an experiment with noisy data interpolation, I reduced the data to every second CMP, then performed NMO and stack to produce the top left panel of Figure 7. Then I further reduced the data to every fourth CMP and interpolated with and without noise estimation, and again did NMO and stack to try to reproduce the stack in the top left of Figure 7. The top right panel, the result without any noise estimation, has significant problems with reflector continuity. In particular, strong reflectors in the center of the section, at 1.2 seconds and at .9 seconds, have undesirable gaps in the midpoint direction. The bottom left stack, which is really a stack of the interpolated and extracted signal \mathbf{s} , is a much more coherent stack, but is slightly lower in temporal frequency content. The bottom right stack is the same, after matched filtering to restore the temporal spectrum. The spectra of the original (top left) stack and the interpolated stack without matched filtering (bottom left) are shown in the left side of Figure 9. After matched filtering, the interpolated, noise suppressed stack has the same temporal frequencies as the stack of the original data; without the filtering step, it is boosted at low frequencies and suppressed at high frequencies. The right side shows the spectra of the original (top left) stack and the interpolated, matched filtered stack (bottom right). The effects of signal and noise separation are shown in Figure 8. The left panel shows the signal component estimated from one of the input CMP gathers, the right shows signal component for an interpolated CMP gather.

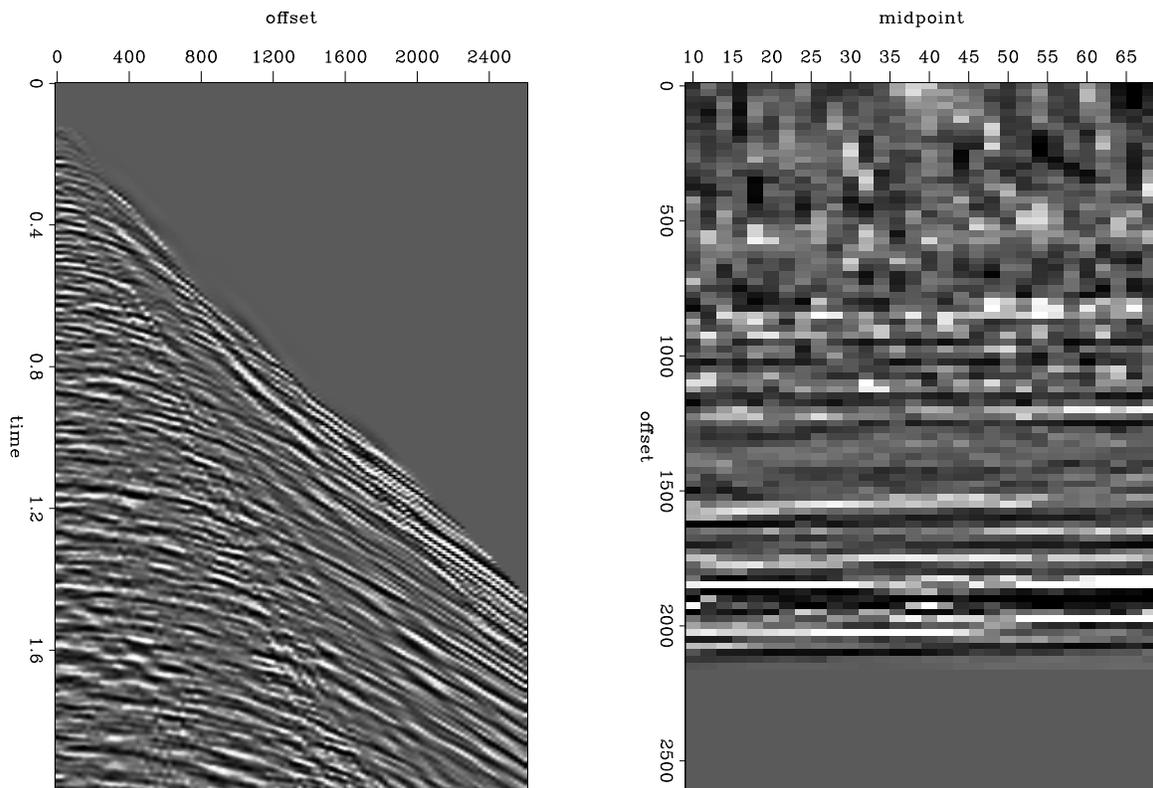


Figure 6: Land seismic data. Left panel shows a CMP gather, right shows a time slice through several CMP gathers. The data are incoherent along the midpoint axis, close to zero offset.

sean1-landCmp [CR]

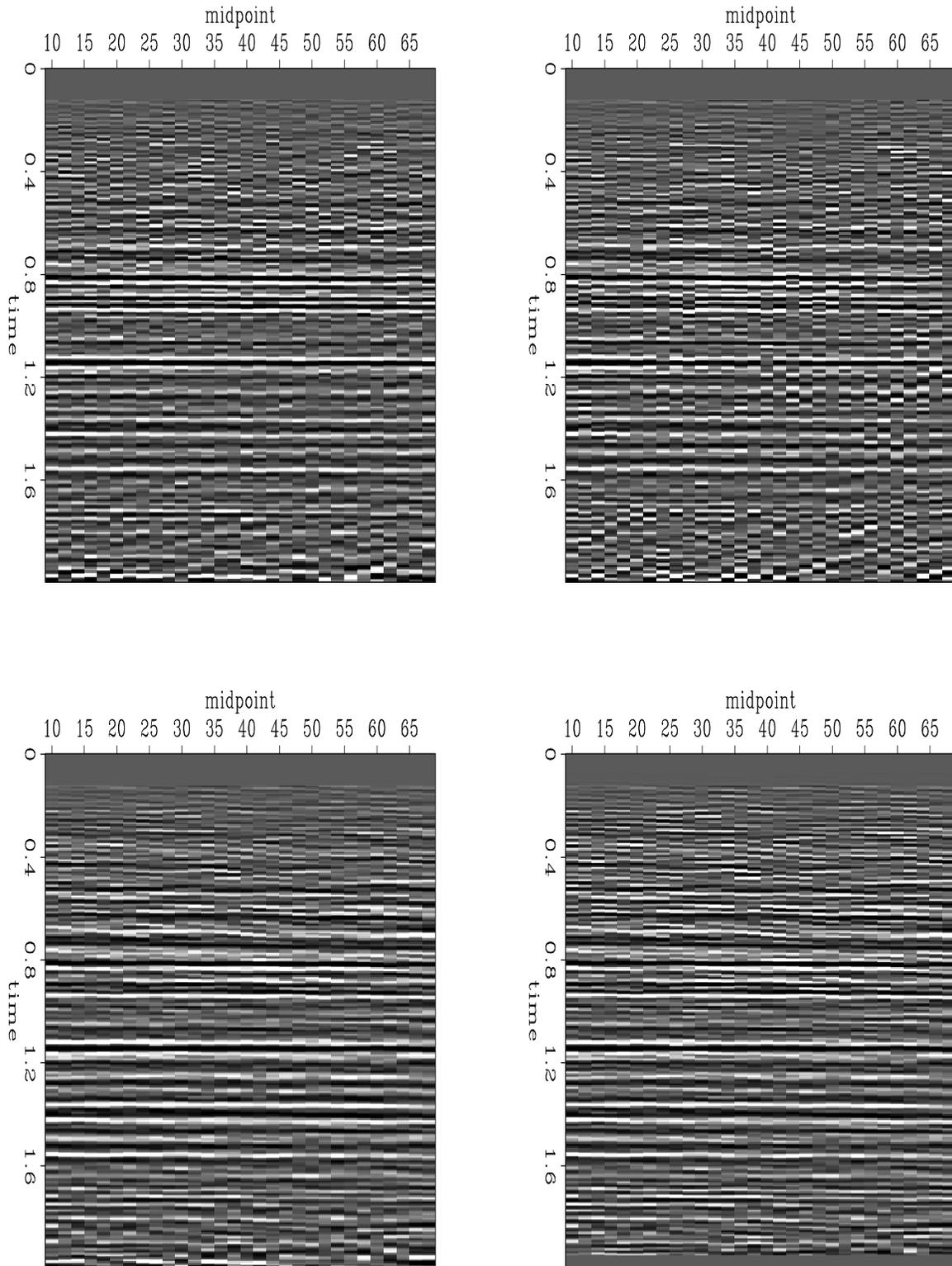


Figure 7: Land data interpolation. Top left panel is a stack of the known data. Top right panel is the stack after throwing away half the CMP gathers and reinterpolating them without attempting to compensate for noise. Bottom left is the stack of the signal component extracted from the known data and interpolated into the missing data. Bottom right is the same stack after matched filtering to get back the temporal frequency content of the known data. sean1-stacks [CR]

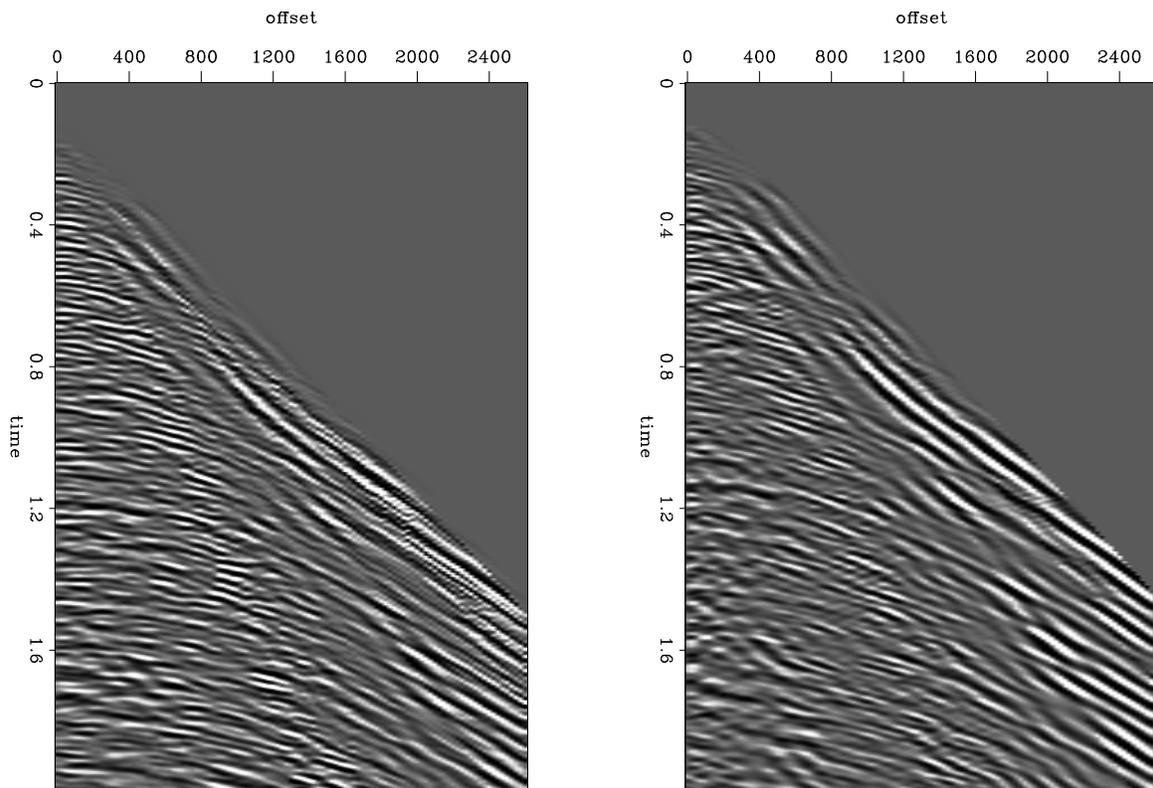


Figure 8: Interpolation and noise removal. Left side shows the estimated signal component of an input CMP gather, right side shows the signal component of an interpolated gather. `sean1-outCmp` [CR]

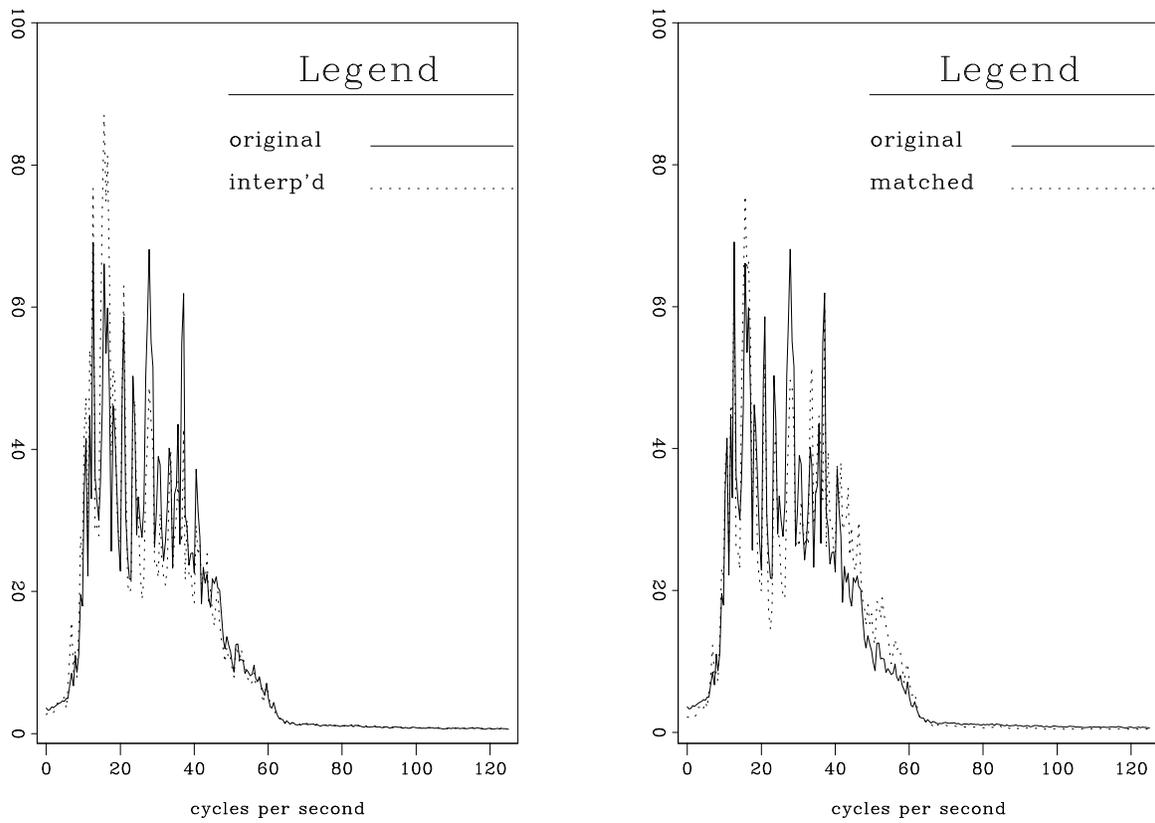


Figure 9: Spectra of stacked data. Left panel shows spectra of the original stack and the stack of the interpolated data, without matched filtering. Right panel shows spectra of the original stack and the matched filtered stack. Solid lines represent original data. `sean1-specs` [CR]

CONCLUSIONS

I describe a method for interpolating missing data with smoothly varying PEFs. Forcing the PEFs to vary smoothly enables us to use very small patches, even a single data sample, and produces good interpolation results on complicated data. I also describe a method for separating signal from noise and interpolating just the signal.

ACKNOWLEDGEMENTS

Thanks go to Bob Clapp, Sergey Fomel, and Jon Claerbout for numerous useful discussions.

REFERENCES

- Claerbout, J. F., 1992, *Earth Soundings Analysis: Processing Versus Inversion*: Blackwell Scientific Publications.
- Claerbout, J. F., 1997, *Geophysical exploration mapping: Environmental soundings image enhancement*: Stanford Exploration Project.
- Claerbout, J., 1998a, Factorization of cross spectra: *SEP-97*, 337–342.
- Claerbout, J. F., 1998b, Multi-dimensional recursive filtering via the helix: *Geophysics*, **63**, no. 5, 1532–1541.
- Clapp, R. G., Fomel, S., Crawley, S., and Claerbout, J. F., 1999, Directional smoothing of non-stationary filters: *SEP-100*, 197–209.
- Clapp, R. G., 1998, Regularizing velocity estimation using geologic dip information: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1851–1854.
- Crawley, S., 1998, Shot interpolation for radon multiple suppression: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1238–1241.
- Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: *SEP-95*, 15–25.
- Spitz, S., 1991, Seismic trace interpolation in the f-x domain: *Geophysics*, **56**, no. 6, 785–794.
- van Borselen, R. G., Thorbecke, J., Fokkema, J. T., and van den Berg, P. M., 1991, Surface related multiple elimination based on reciprocity: Surface related multiple elimination based on reciprocity:, 61st Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1339–1342.

Directional smoothing of non-stationary filters

Robert G. Clapp, Sergey Fomel, Sean Crawley, and Jon F. Claerbout¹

ABSTRACT

Space-varying prediction error filters are an effective tool in solving a number of common geophysical problems. To estimate these filters some type of regularization is necessary. An effective method is to smooth the filters along radial lines in CMP gathers where dip information is relatively unchanging.

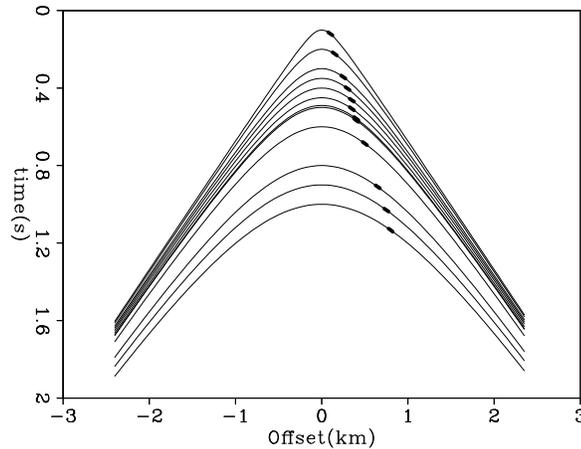
INTRODUCTION

Estimating filters is routine in seismic processing. The simplest example might be deconvolution, but filter estimation is also valuable in many other aspects of seismic processing: interpolation (Spitz, 1991; Crawley, 1998), noise attenuation (Canales, 1984; Soubaras, 1994; Abma, 1995), missing data (Claerbout, 1998a; Fomel et al., 1997), and coherency estimation (Schwab, 1998; Bednar, 1997) to name just a few. All of these processes are based on the concept of finding a filter that minimize the energy when it is applied to a given set of data. The fundamental assumption is that that statistics of the data does not change spatially. This is often not the case. One solution to this problem is to separate the data into a number of overlapping patches (Claerbout, 1992d) where the stationary statistic assumption is more valid. Unfortunately, there is a limit to how small we can make our patches and still gather sufficient statistics.

A way around this limitation is to estimate a space varying prediction error filter (PEF) (Crawley et al., 1998). In the extreme case you can think of estimating a filter at every data location, or more realistically, at a coarser grid spacing. With so many filters and, as result, so many filter coefficients, our estimation can quickly turn into an undetermined or at least poorly determined problem. Therefore we must impose some type of regularization to our estimation problem. Choosing an appropriate regularization then becomes an issue. In this paper we argue that when estimating filters on seismic CMP data, you should smooth along radial lines. In a constant velocity medium the dip along a radial trace does not change, but in a more complex media it will vary slowly (Ottolini, 1982). By limiting filter variation in the radial direction we gather more data in our filter estimation thus enhancing stability. Here we show how to estimate the appropriate smoothing direction, and how to build and apply the appropriate regularization.

¹email: bob@sep.Stanford.EDU, sean@sep.Stanford.EDU, sergey@sep.Stanford.EDU,
jon@sep.stanford.edu

Figure 1: Constant velocity curves. The thick lines are the same dip on all the reflectors. Note how they form a line. `bob3-dips.constant` [ER]



WHY SMOOTH RADIALLY

Dips change quickly along every axis in seismic data. As a result a single PEF has trouble characterizing it, even in small patches (Crawley, 1999). By estimating a space-varying PEF, we can overcome this deficiency. Unfortunately, this changes our estimation problem from something overdetermined to something, at times, grossly underdetermined. To stabilize our filter estimation we must apply some type of regularization to the standard PEF estimation optimization goals:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{Y}\mathbf{a} \\ \mathbf{0} &\approx \epsilon\mathbf{F}\mathbf{a} \end{aligned} \quad (1)$$

where \mathbf{a} is our space-varying filter, \mathbf{Y} is convolution with our data, and \mathbf{F} is a roughener. To speed up convergence, we can take advantage of helix theory (Claerbout, 1998c) and reformulate our regularized problem into a preconditioned one

$$\begin{aligned} \mathbf{0} &\approx \mathbf{Y}\mathbf{F}^{-1}\mathbf{A}^{-1}\mathbf{p} \\ \mathbf{0} &\approx \epsilon\mathbf{p} \end{aligned} \quad (2)$$

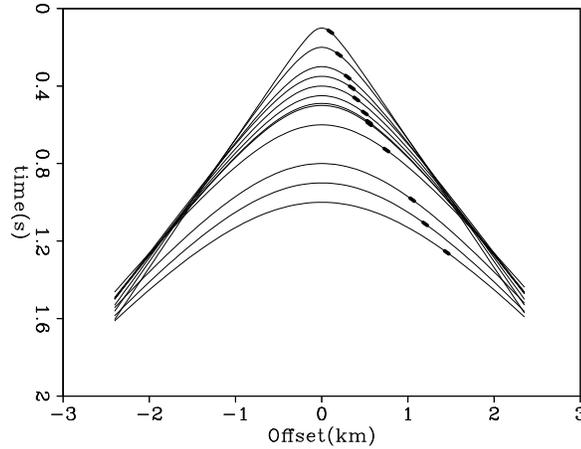
where

$$\mathbf{p} = \mathbf{F}\mathbf{a}. \quad (3)$$

Our choice for \mathbf{F} can have significant influence on both the speed and quality of our filter estimation. The character of seismic data itself gives us a clue on what type of regularization we should use. A PEF filter is most successful when the statistics of the data it is being estimated from are stationary. Logically, our rougher \mathbf{F} , or \mathbf{F}^{-1} , should tend to smooth along a region with consistent dips, or along Snell traces (Claerbout, 1978). Figure 1 shows several constant velocity hyperbolas, with the same dips highlighted. These dips all fall along a radial line through zero time and zero offset. If we look at hyperbolas in $v(z)$, Figure 2, we see that there is deviation from a simple line, but generally this trend is preserved.

Figure 2: $V(z)$ medium curves. The thick lines represent the same dip. Note how they are not perfectly linear but generally lay along a line.

`bob3-dips.vz` [ER]



CHOOSING SMOOTHING DIRECTIONS

Prediction-error filters work best on predicting local plane waves (Claerbout, 1992c; Canales, 1984). With non-stationary filters, it is possible to predict data with variable slopes. For preconditioning the filter estimation problem, such filters can be smoothed along the direction where the slope stays locally constant. To put this principle into a mathematical form, let us denote the monodop data as $P(x, y)$, where x and y are the coordinate values. On a seismic data section, the y coordinate would have the meaning of time, but here we would like to develop a general method that would work on different kinds of data. The local dip field of the data can be defined by the formula

$$D(x, y) = -\frac{P_x}{P_y}, \quad (4)$$

where P_x and P_y denote the first partial derivatives: $P_x = \frac{\partial P}{\partial x}$, $P_y = \frac{\partial P}{\partial y}$. To validate formula (4), consider a plane-wave model with the slope s :

$$P(x, y) = P_0(y - sx). \quad (5)$$

Substituting (5) into formula (4), we can see that the $D(x, y)$ indeed produces an estimate of s (Claerbout, 1992a). In the general case, $D(x, y)$ corresponds to the tangent of the local plane wave angle, measured from the x axis in the direction of the y axis. Bednar (1997) describes an application of formula (4) for computing coherency attributes. Instead of using formula (4) explicitly, we intend to estimate prediction-error filters that would destroy local plane waves in the data (Claerbout, 1992c; Schwab, 1998). To precondition the filter estimation problem we can smooth the filters in the direction of the least change in the slope. By analogy with (4), the smoothing direction can be defined as follows:

$$S(x, y) = -\frac{D_x}{D_y}, \quad (6)$$

or, substituting formula (4),

$$S(x, y) = -\frac{P_x P_{xy} - P_y P_{xx}}{P_y P_{xy} - P_x P_{yy}}, \quad (7)$$

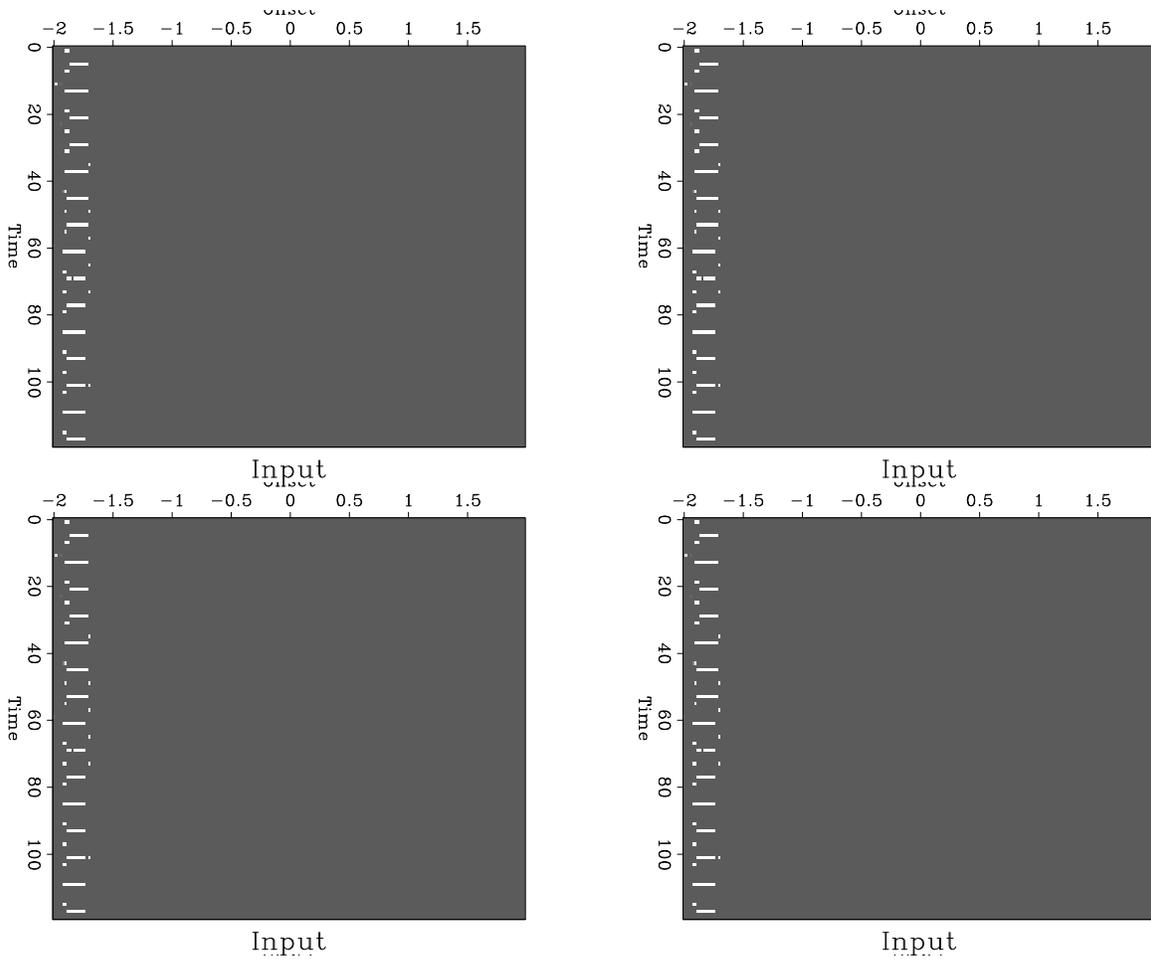


Figure 3: The effect of dip smoothing. The top-left panel is the input, the top-right is the result of applying the forward operator, bottom-left is the adjoint response; and bottom-right is the cascade of forward and the adjoint. [bob3-random] [ER]

where P_{xx} , P_{yy} , and P_{xy} are the corresponding second-order partial derivatives. An important analytical test case is a constant-velocity CMP gather, composed of reflection hyperbolas:

$$P_{\text{hyper}}(x, y) = P_0 \left(\sqrt{y^2 - s^2 x^2} \right). \quad (8)$$

Substituting (8) into formula (7) leads to the expression

$$S_{\text{hyper}}(x, y) = \frac{y}{x}, \quad (9)$$

which suggests smoothing the estimated prediction-error filters along radial lines on the $\{x, y\}$ plane (Crawley et al., 1998). Figure 4 and 5 illustrate a practical application of formulas

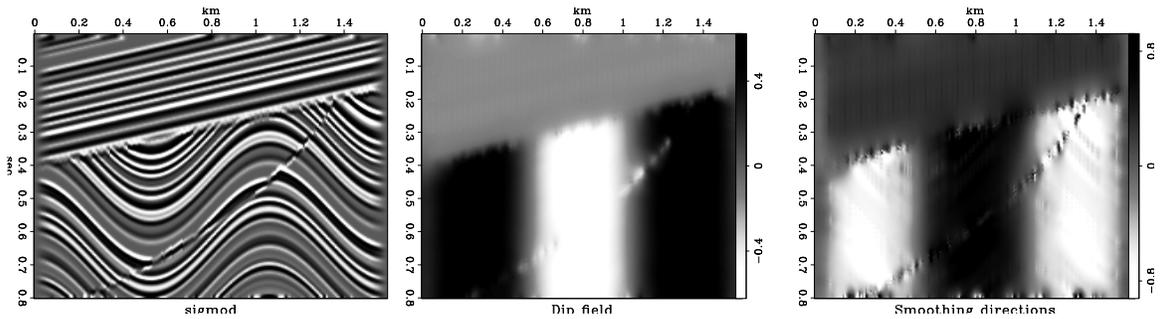


Figure 4: Synthetic model from *Basic Earth Imaging* (left), its estimated dip field (center), and estimated smoothing directions (right). `bob3-sigmod` [ER,M]

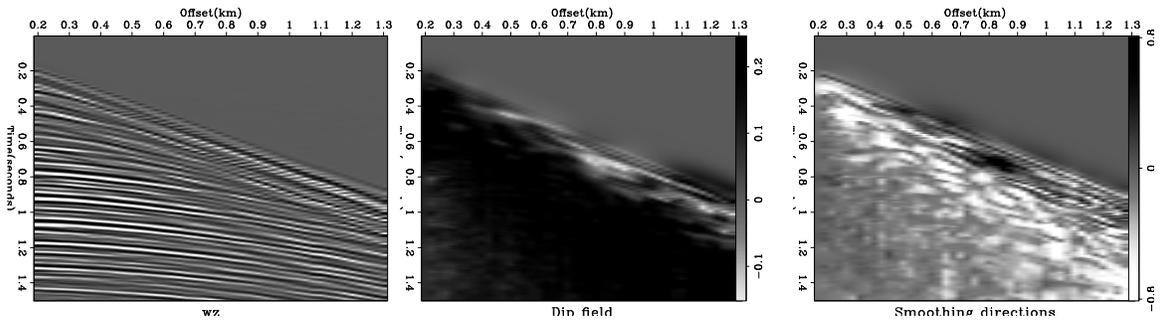


Figure 5: Seismic shot gather (left), its estimated dip field (center), and estimated smoothing directions (right). `bob3-wz` [ER,M]

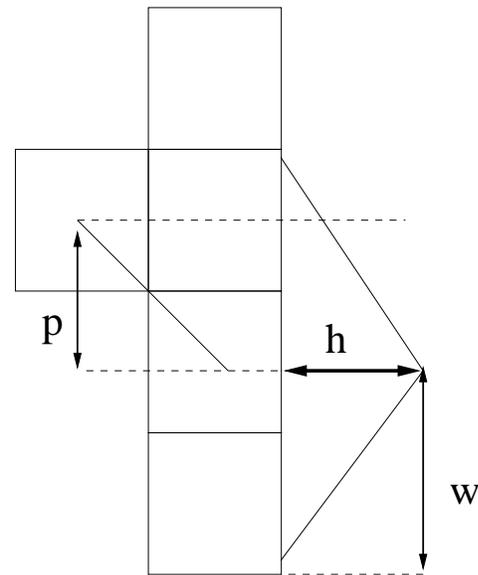
(4) and (6) on a synthetic reflectivity model from *Basic Earth Imaging* (Claerbout, 1995) and on a shot gather from the Yilmaz collection (Yilmaz, 1987). In both cases the first- and second-derivative operators were computed with simple finite-difference schemes. To avoid a non-stable division in formulas (4) and (6), we solve the regularized least-square system

$$\begin{cases} \mathbf{D}\mathbf{x} \approx \mathbf{N} \\ \epsilon \nabla \mathbf{x} \approx \mathbf{0} \end{cases}, \quad (10)$$

where \mathbf{D} and \mathbf{N} denote the denominator and the numerator respectively, ϵ is the scalar regularization parameter, and \mathbf{x} is the estimated regularized ratio. Our simple two-point finite-difference scheme does not handle correctly the aliased dips on the seismic gather in Figure

Figure 6: A finite-difference star for a monoplane rejection filter. The left column contains a '1'. The right column contains samples off a triangle. The desired slope is represented by p , the smaller w the more precise the dip smoothed, and the larger h the bigger the area the smoother acts on.

bob3-steering [NR]



5. Nevertheless it produces a reasonable output, which we can use as a rough estimate of the smoothing directions.

HOW TO SMOOTH RADIALY

Once we know what directions we wish to smooth in, we must build an operator that can smooth in the desired directions. We want to minimize the cost of smoothing, so we would like the filters to be small. As discussed in Clapp (1997) an effective method is to build a series of small plane-wave annihilation filters (Claerbout, 1992b) and then combine them into a single operator.

Constructing a filter

The basic idea in building a steering filter is to create a filter that destroys a given slope p . Further, we would like to keep differences of the bandwidth response for filters oriented at different slopes to a minimum. We can achieve both these goals by constructing a triangle centered at the appropriate slope (Figure 6.) Every grid cell center which the triangle passes through is assigned a negative value proportional to the height of the triangle at that location. The wider the triangle base, the less precise, and more Gaussian-like our smoother becomes, Figure 7. By decreasing the sum of the coefficients (with a hard limit of -1 to ensure filter stability when applying polynomial division (Claerbout, 1976)), we can spread information larger distances.

Control

The number of adjustable parameters in the filter construction is both a curse and a blessing. Whenever you add parameters to your problem, the model space that you have to search increases exponentially. With two adjustable parameters, taken to the extreme, at every model point, the task can seem daunting. Generally, the smartest course is to keep these two parameters constant throughout the whole model space. But, this freedom also opens up interesting possibilities. In certain regions of the data you might feel that the radial assumption is not quite valid, or that dips aren't changing quite as fast. In this region you could consider making your triangle bigger, smoothing your filter coefficients over a wider angle range, while keeping it small in areas where dip changes quickly. The sum of the non-zero lag coefficients opens up another intriguing freedom. As Figure 8 shows, when the sum of the non-zero lag coefficients gets close to -1 , the area over which the smoother operates increases greatly. This is similar to increasing the ϵ value over only a portion of your model space. This gives you the freedom to easily smooth regions where filter stability is questionable, while allowing high frequency changes in areas of good data.

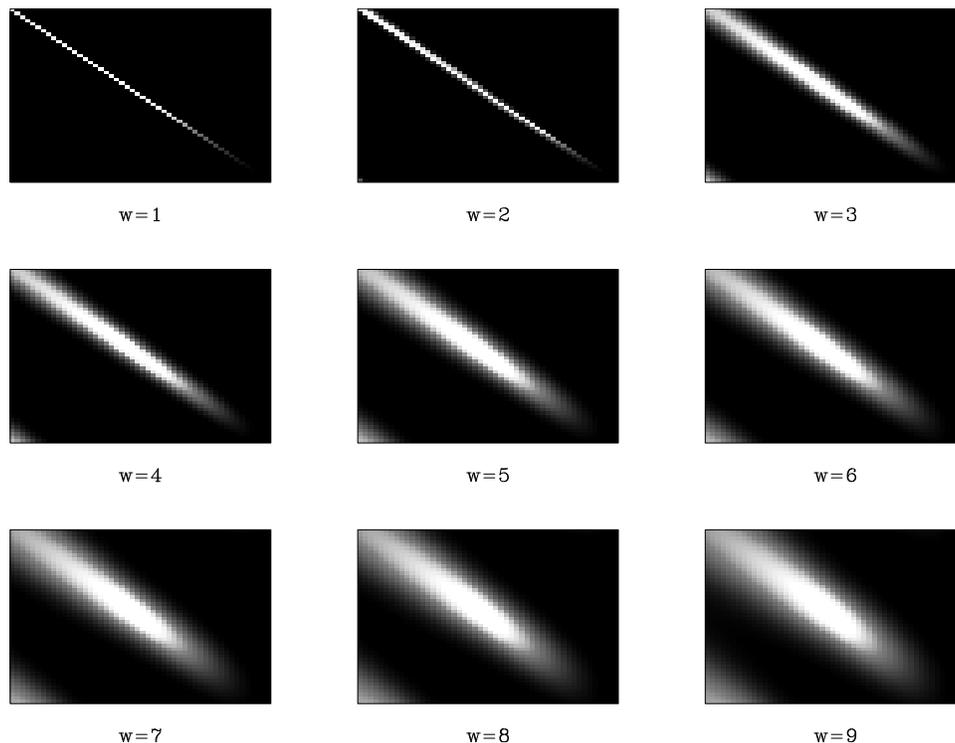


Figure 7: The impulse response of the smoothing filter as function of the triangle base. Note the wider the base, the less precise the dip smoothing. `bob3-width` [ER]

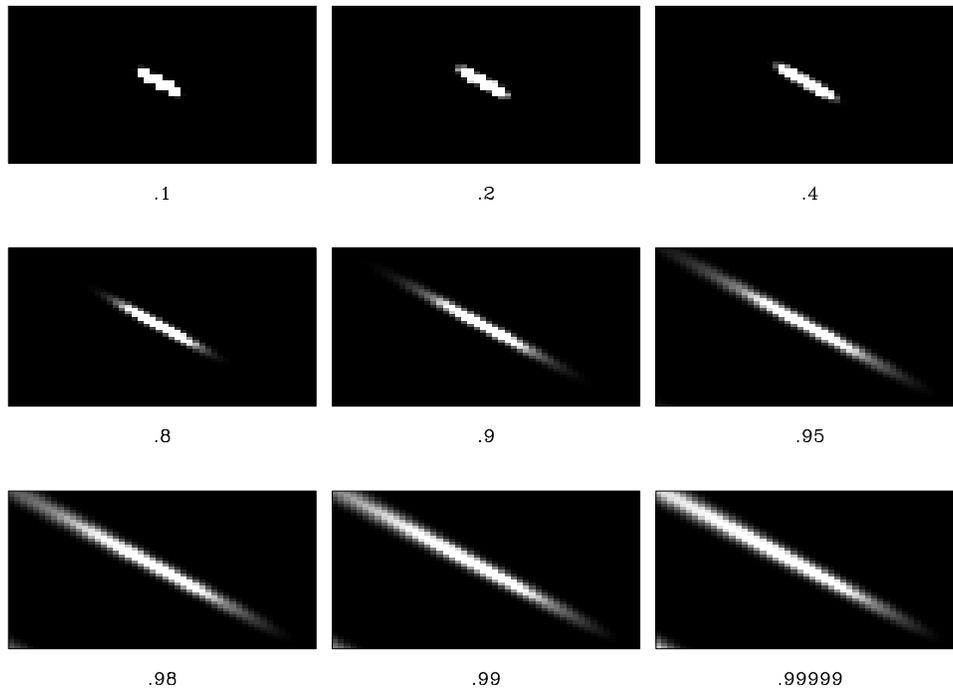


Figure 8: The impulse response of the smoothing filter as the sum of the non-zero lag coefficients get closer to 1. `bob3-distance` [ER]

Applying filter

As discussed by Claerbout (1998a), by defining our filters in helix space we can use polynomial division to apply their inverse. This same principal holds true for space varying filters. The basic algorithm is:

```
integer function npolydiv(adj,add,model,data){
logical          :: adj,add
real            :: xx(:),yy(:)
integer         :: ia, ix, iy, ip
integer, dimension(:), pointer :: lag
real,    dimension(:), pointer :: flt,tt
allocate(tt(size(yy)))
tt = 0.
if( adj) {
    tt = yy
    do iy= nd, 1, -1 { ip = aa%pch( iy)
        lag => aa%hlx( ip)%lag; flt => aa%hlx( ip)%flt
        do ia = 1, size( lag) {
            ix = iy - lag( ia);    if( ix < 1) cycle
            tt( ix) -= flt( ia) * tt( iy)
        }
    }
    xx += tt
} else {
    tt = xx
}
```

```

do iy= 1, nd { ip = aa%pch( iy)
  lag => aa%hlx( ip)%lag; flt => aa%hlx( ip)%flt
  do ia = 1, size( lag) {
    ix = iy - lag( ia);      if( ix < 1) cycle
    tt( iy) -= flt( ia) * tt( ix)
  }
}
yy += tt
}
allocate(tt(size(yy)))
}

```

PREDICTING A CMP GATHER

To show how radial smoothing can be valuable, we constructed a synthetic CMP gather using a Kirchhoff modeling code. To these CMP gathers we added two lines, one in a radial direction and one at constant time (left panel of Figure 9.) The constant time line can be thought of as noise, while the radial line represents conflicting information that fits our model of valid data. We then attempted to estimate the shot gather using fitting goals (3) with filters every 20 points in time and every 5 points in offset using two different types of preconditioners. The center panel shows the residual after using an inverse Laplacian (Claerbout, 1998b) and the right panel, radial smoothers. Generally, the two approaches did approximately the same job in predicting the data. The difference comes where the lines intersect the hyperbolas. If we examine the intersection points, more closely, Figure 10, we see that in the case of the Laplacian we did an equal job of predicting the hyperbolas and the constant time line. When using steering filters, the constant time line is much stronger (we avoid predicting noise).

INTERPOLATING A CMP GATHER

Once filters are estimated, one of their potential uses is missing data interpolation. Systematic gaps in data acquisition may cause data aliasing sufficient to make some processing steps difficult (Spitz, 1991; Crawley, 1998). Adding more traces can dealias the data. To add more traces, we require that the original data and the new data have the same dips (Claerbout, 1997). The dip information is carried in the PEFs. The missing data estimation is formulated just like the filter estimation, except that the PEFs are known and the data unknown. Also, we constrain the data by specifying that the originally recorded traces do not change. To separate the known and unknown data we have a known data selector \mathbf{K} and an unknown data selector \mathbf{U} , with $\mathbf{U} + \mathbf{K} = \mathbf{I}$. These multiply by 1 or 0 depending on whether the data was originally recorded or not. With \mathbf{A} signaling convolution with the PEF and \mathbf{y} the vector of data, the regression is $0 \approx \mathbf{A}(\mathbf{U} + \mathbf{K})\mathbf{y}$, or $\mathbf{A}\mathbf{U}\mathbf{y} \approx -\mathbf{A}\mathbf{K}\mathbf{y}$. Filters at every data point are cumbersome to estimate, so we estimate filters over small areas. This is just like patching (Claerbout, 1992d) except that now the patches are not independent. If the patches are independent, there is a

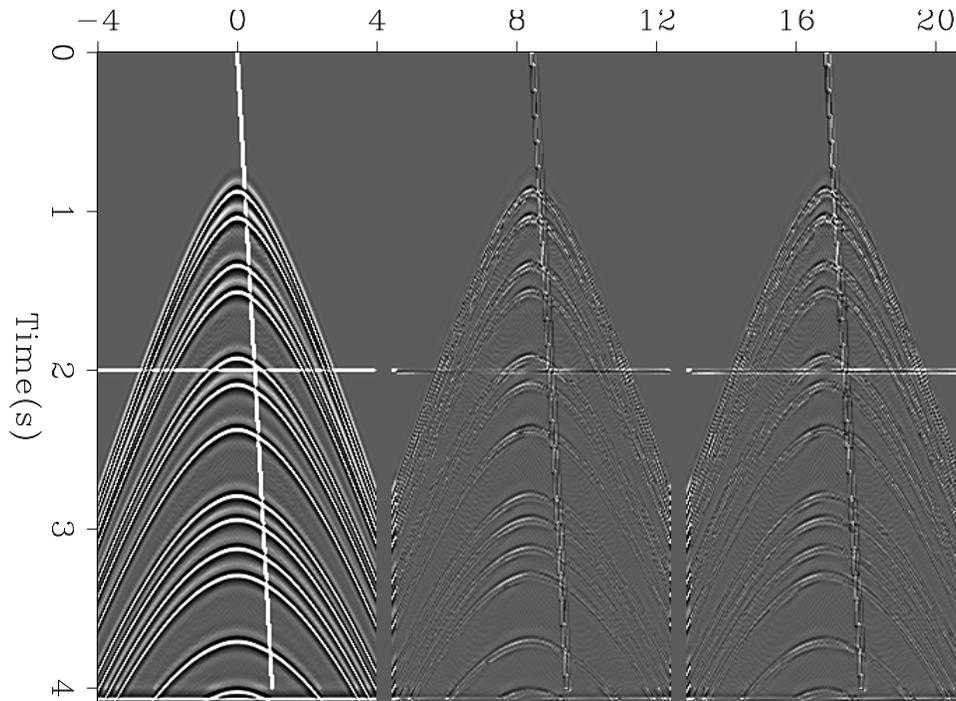


Figure 9: The result after 15 conjugate gradient steps of fitting goals (3). The left panel is the input, the center is using an inverse Laplacian preconditioner, the right panel is using radial smoothing. `bob3-comparison1` [ER]

lower limit on the patch size, because a patch must contain plenty of data to provide enough fitting equations to determine all the filter coefficients. Experience shows that where the data have curvature, the minimum patch size tends to be too large for the assumption of stationarity to be reasonable. Smoothing the filters allows us to make the patches much smaller, so that stationarity assumptions are workable. We arrange the new patches in polar coordinates, to take advantage of the notion of radial smoothing. An illustration is given in Figure 11. The cmp gather is overlaid by lines which delineate patch boundaries. Degree of smoothing in r and θ is adjustable. The patches shown are fairly large. Crawley and Claerbout(1999) explains further this method and shows the result of interpolating using radial patches and smoothers.

CONCLUSIONS

As the progress report deadline arrived, the authors were uncertain among themselves whether the results were correct. The prediction-error filters have clearly reduced the output variance, but the results do not clearly show the dip dependences that we expected. Generally we expected to see strong energy locally where events cross, and we expected to see weak energy where the data was locally monodip. It is not clear that this happened.

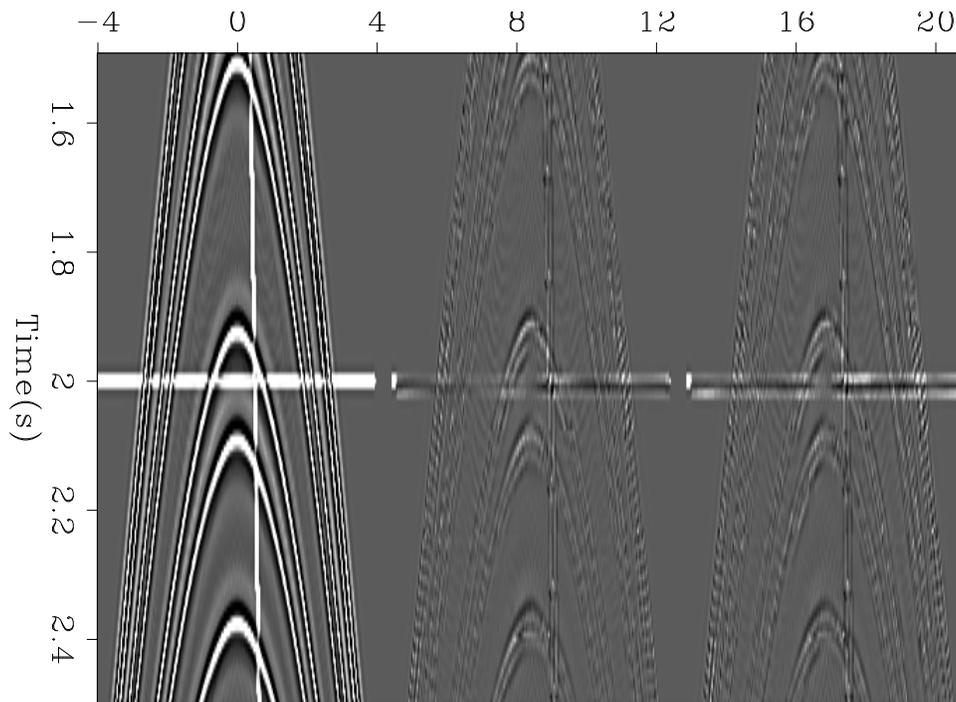


Figure 10: An enlargement of Figure 9. Note that the constant time line, what we consider noise, is much better predicted by the inverse Laplacian (center panel) than by radial smoothing. [bob3-comparison2](#) [ER]

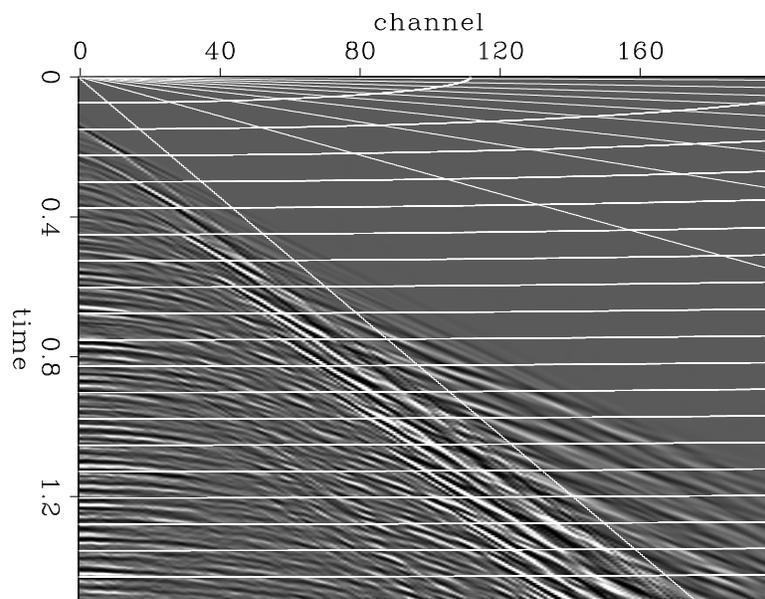


Figure 11: Example CMP gather overlaid by patch boundaries. Smoothing of filter coefficients is adjustable in r and θ . [bob3-web](#) [ER]

ACKNOWLEDGMENTS

We thank Dave Nichols for encouraging discussions.

REFERENCES

- Abma, R., 1995, Least-squares separation of signal and noise with multidimensional filters: Ph.D. thesis, Stanford University.
- Bednar, J. B., 1997, Least squares dip and coherency attributes: *SEP-95*, 219–225.
- Canales, L. L., 1984, Random noise reduction: 54th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Session:S10.1.
- Claerbout, J. Fundamentals of Geophysical Data Processing: <http://sepwww.stanford.edu/sep/prof/>, 1976.
- Claerbout, J. F., 1978, Snell waves: *SEP-15*, 57–72.
- Claerbout, J. F., 1992a, Earth Soundings Analysis: Processing Versus Inversion: Blackwell Scientific Publications.
- Claerbout, J. F., 1992b, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.
- Claerbout, J. F., 1992c, Information from smiles: Mono-plane-annihilator weighted regression: *SEP-73*, 409–420.
- Claerbout, J. F., 1992d, Nonstationarity and conjugacy: Utilities for data patch work: *SEP-73*, 391–400.
- Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.
- Claerbout, J. F., 1997, Geophysical exploration mapping: Environmental soundings image enhancement: Stanford Exploration Project.
- Claerbout, J. Geophysical Estimation by Example: Environmental soundings image enhancement: <http://sepwww.stanford.edu/sep/prof/>, 1998.
- Claerbout, J., 1998b, Multidimensional recursive filters via a helix: *SEP-97*, 319–335.
- Claerbout, J. F., 1998c, Multi-dimensional recursive filtering via the helix: *Geophysics*, **63**, no. 5, 1532–1541.
- Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: *SEP-95*, 27–42.
- Crawley, S., Clapp, R., and Claerbout, J., 1998, Deconvolution and interpolation with nonstationary filters: *SEP-97*, 183–192.

- Crawley, S., 1998, Shot interpolation for radon multiple suppression: SEP-**97**, 173–182.
- Crawley, S., 1999, Interpolation with smoothly nonstationary prediction-error filters: SEP-**100**, 181–196.
- Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: SEP-**95**, 15–25.
- Ottolini, R., 1982, Migration of reflection seismic data in angle-midpoint coordinates: Ph.D. thesis, Stanford University.
- Schwab, M., 1998, Enhancement of discontinuities in seismic 3-D images using a Java estimation library: Ph.D. thesis, Stanford University.
- Soubaras, R., 1994, Signal-preserving random noise attenuation by the f-x projection: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1576–1579.
- Spitz, S., 1991, Seismic trace interpolation in the f-x domain: *Geophysics*, **56**, no. 6, 785–794.
- Yilmaz, O., 1987, *Seismic data processing*: Soc. Expl. Geophys., Tulsa, OK.

Texture synthesis and prediction error filtering

Morgan Brown¹

ABSTRACT

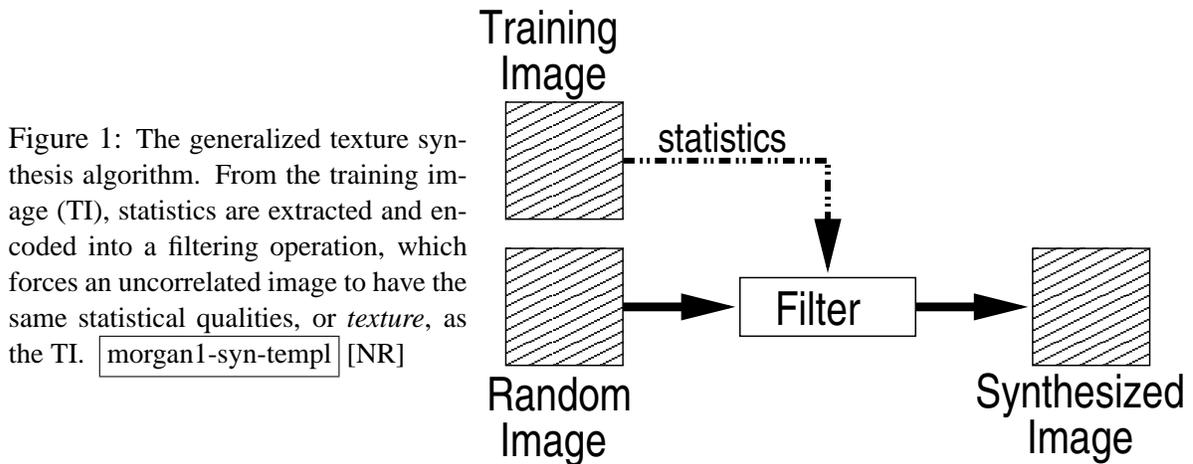
The spectrum of a prediction-error filter (PEF) tends toward the inverse spectrum of the data from which it is estimated. I compute 2-D PEF's from known "training images" and use them to synthesize similar-looking textures from random numbers via helix deconvolution. Compared to a similar technique employing Fourier transforms, the PEF-based method is generally more flexible, due to its ability to handle missing data, a fact which I illustrate with an example. Applying PEF-based texture synthesis to a stacked 2-D seismic section, I note that the residual error in the PEF estimation forms the basis for "coherency" analysis by highlighting discontinuities in the data, and may also serve as a measure of the quality of a given migration velocity model. Last, I relate the notion of texture synthesis to missing data interpolation and show an example.

INTRODUCTION

In terms of digital images, the word *texture* might be defined as, "an attribute representing the spatial arrangement of gray levels of the pixels in a region," (IEEE, 1990). In the same context, I define *texture synthesis* as the process of first estimating the spatial statistical properties of a known image and then imparting these statistics onto a second (random) image. Figure 1 illustrates the general approach taken here: an uncorrelated image is transformed into one with the same statistical qualities as a known "training image" (TI), through an as-yet undefined filtering operation.

Texture synthesis is an active area of research in the computer graphics community, owing to the need for realistic, quickly generated surface textures (Simoncelli and Portilla, 1998; Heeger and Bergen, 1995; Mao), but the same notion of texture applies to the earth sciences as well. Physically measurable quantities, be they geology, gravity, or topography, behave in certain repeatable ways as a function of space, i.e., these quantities have a given texture. Inversion problems are often underdetermined, hampered by a lack of "hard" measurements, causing a nullspace of high dimension. A priori "soft" constraints on functional form of the unknown model help in suppressing the nullspace of modeling operators. These a priori constraints can be conceptualized as textures. For instance, in velocity analysis and tomography, the earth's velocity field is sometimes assumed to have a "blocky" texture (Clapp et al., 1998). Underdetermined inverse interpolation problems are often regularized by assuming "smooth" model texture (Claerbout, 1999).

¹email: morgan@sep.stanford.edu



The prediction-error filter (PEF) is an autoregressive filter which has the distinction of capturing the inverse spectrum of the data it is regressed upon. Because it captures this essential statistical property of the data, the PEF is a candidate for the generic "filter" operation shown in Figure 1.

This paper is intended as a follow-up to the earlier work by Claerbout and Brown (1999), which presented a texture synthesis technique utilizing 2-D PEF's and 2-D deconvolution via the helix transform (Claerbout, 1998). First I motivate the texture synthesis problem by applying a Fourier transform-based technique to create synthetic textures of everyday objects, then introduce and apply a PEF-based technique to synthesize the same images. I compare the results of the two methods and conclude that the PEF-based method is the better choice because it more naturally handles missing data. Next I apply the PEF-based method to a 2-D stacked seismic section. The nature of the residual error in the PEF estimation of this example suggests application to seismic discontinuity detection and migration velocity analysis. Last, I solve a simple missing data problem to illustrate how regularization with a PEF imparts a reasonable "texture" onto the nullspace.

FOURIER TRANSFORM METHOD

The texture synthesis methodology of this paper really boils down to one of spectral estimation. An image's amplitude spectrum contains the relative weights between frequency components, while the phase spectrum localizes these frequency components in space (Castleman, 1996). Therefore, it stands to reason that the texture of stationary, loosely correlated images is adequately modeled using the amplitude spectrum alone. This idea is the basis for the Fourier Transform method of texture synthesis: all "realizations" of texture synthesis are forced to have the same amplitude spectrum, differing only in phase. The following is an outline of the method.

1. Given a training image, $t(x, y)$, compute its amplitude spectrum:

$$R(k_x, k_y) = T^*(k_x, k_y)T(k_x, k_y) \quad (1)$$

2. Create random phase function: $\phi_r(k_x, k_y) = \text{random numbers}$.
3. Reconstruct by substituting random phase:

$$t_{\text{recon}}(x, y) = \mathcal{F}^{-1} \left\{ \sqrt{|R(k_x, k_y)|} e^{i\phi_r(k_x, k_y)} \right\} \quad (2)$$

Figures 2 and 3 illustrate the Fourier transform method of texture synthesis. Clockwise from top-left: the training image, the synthesized image, the TI's amplitude spectrum, and the TI's phase spectrum.

Figure 2: Smoothed random image and Fourier transform synthesis. The TI is stationary, so the synthesis result is convincing. Notice that the true phase, in the regions where the modulating amplitude spectrum is nonzero, is quite random in appearance.

`morgan1-rand2d-ftsyn` [ER]

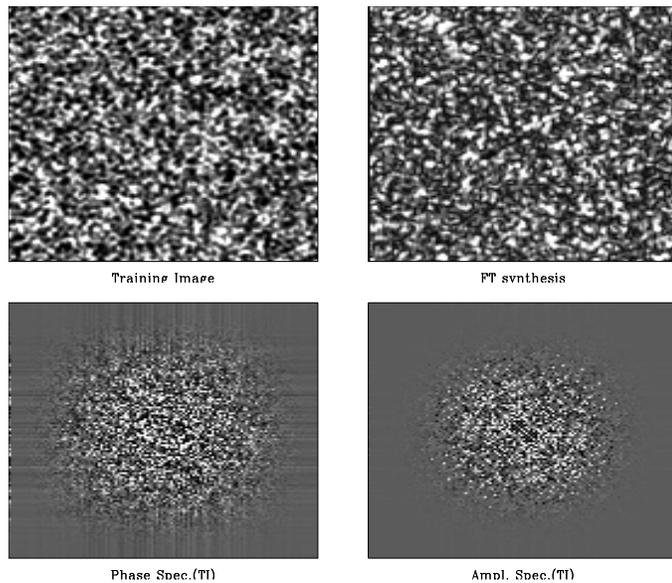
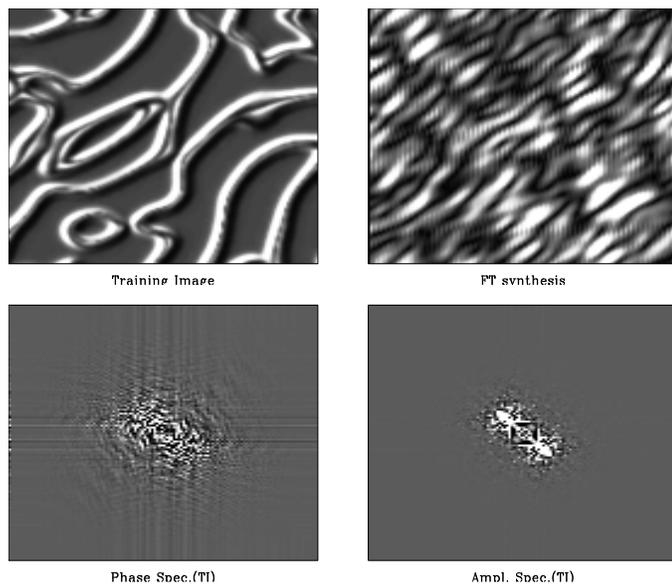


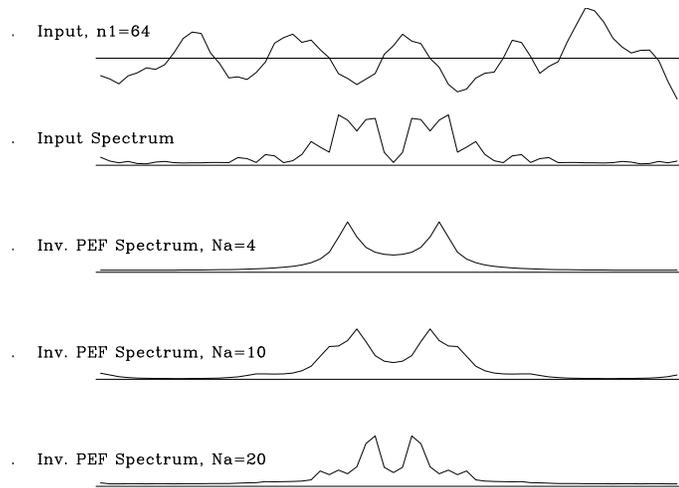
Figure 3: "Ridges" image and Fourier transform synthesis. The correlation is both long-range and extremely complicated - quite like a meandering network of fluvial channels. Though the synthesized image has the same *general* character as the TI, not all of the structures are modeled, proving the inadequacy of the amplitude spectrum for modeling nonstationary, highly correlated images. The TI phase spectrum shows some ordering, so the random phase substitution was ill-advised. `morgan1-ridges-ftsyn` [ER]



PEF-BASED METHOD

Theoretically, the convolution of data (N_d points) and a PEF (N_a coefficients) estimated from the data is approximately uncorrelated in the limit $N_a \rightarrow N_d \rightarrow \infty$: a spike at zero lag plus Gaussian, independent identically distributed (iid) noise elsewhere. Thus the spectrum of this residual error is approximately white. The frequency response of the “inverse PEF”, as computed by deconvolution, is an N_a -point parameterization of the N_d -point inverse amplitude spectrum, as illustrated in Figure 4. As the size of the filter increases, the parameterization becomes more accurate, as expected from theory (Claerbout, 1976). The notion of PEF as “decorrelator” is quite akin to decomposition by principal components (Castleman, 1996), where the number of principal components used in computation determines the degree of decorrelation.

Figure 4: Frequency response of “inverse PEF” (deconvolution) as a function of filter size. As expected, as the filter length increases, the approximation improves. `morgan1-rand1d-spec` [ER]



The following is an outline of the PEF-based texture synthesis method.

1. Given training image $t(x, y)$, estimate unknown PEF $a(x, y)$ via least squares minimization:

$$\min \| t * a \|^2 \quad (3)$$

2. The residual $r = t * a$ is approximately uncorrelated, with the same dimension as the TI, since we use an "internal" convolution algorithm (Claerbout, 1999). It can be proved that a is a minimum phase filter, (Claerbout, 1976) so deconvolution (polynomial division) robustly and stably reconstructs t given r . Generate a random residual r' with the same dimension as r . To create the synthetic texture, simply deconvolve r' by a :

$$t_{\text{syn}} = r' / a \quad (4)$$

where the “/” refers to polynomial division, our preferred method of deconvolution.

Though the residual is uncorrelated, it does contain “phase” information. Deconvolution of a random image blindly spreads scaled copies of the impulse response of the inverse PEF

across the output space. If the residual r is not sufficiently whitened, then the replacement of r with r' will lead to an ineffective representation of t by t_{syn} .

Figures 5 through 7 illustrate the PEF-based texture synthesis process. The left-hand panel shows the training image, the center panel shows the residual $r = t * a$, and the right-hand panel shows the synthesized image, $t_{\text{syn}} = r' / a$. A 10x10 PEF is used in each case. The blank areas in the residual panel correspond to regions where the PEF falls outside the bounds of the known data.

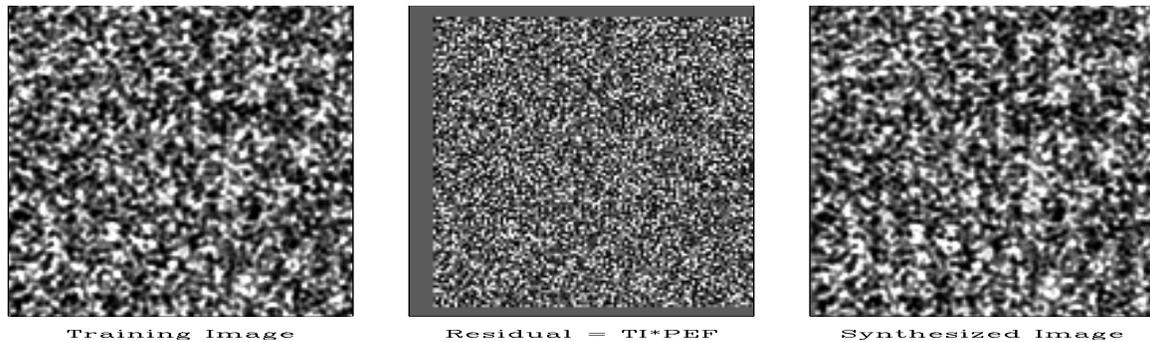


Figure 5: Smoothed random 2-D image and PEF-based texture synthesis result. The TI is quite simple (stationary, low correlation), so as expected, the synthesized image and the TI are almost indistinguishable. To the naked eye, the residual appears effectively white. `morgan1-rand2d-pefsyn` [ER]

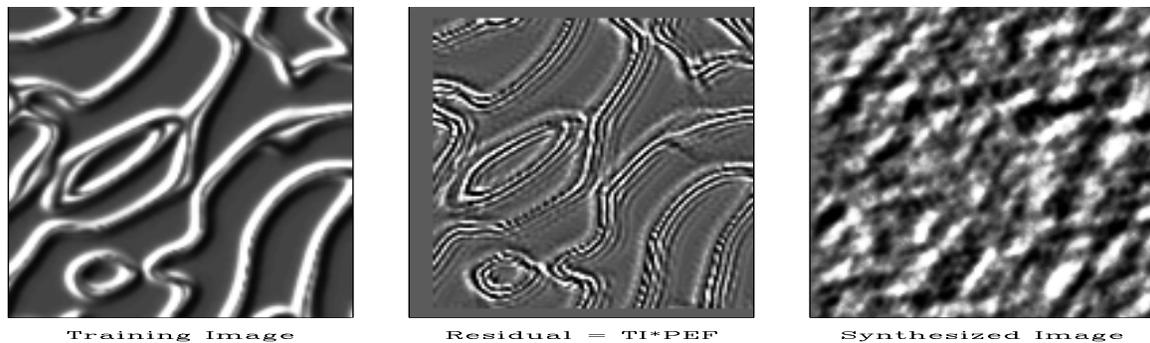


Figure 6: “Ridges” image and PEF-based texture synthesis result. Recall that the complicated connected features of this image were not completely synthesized by the Fourier transform method (Figure 3), of which the PEF method is an approximation. This synthesized image bears even less resemblance to the TI, exhibiting only a general southwest-to-northeast trend. The wavy, ridge-like features have many different dips, making them difficult to predict with a PEF, and with two point statistics in general. The same can be said for the ubiquitous hyperbolic features of reflection seismology. `morgan1-ridges-pefsyn` [ER]

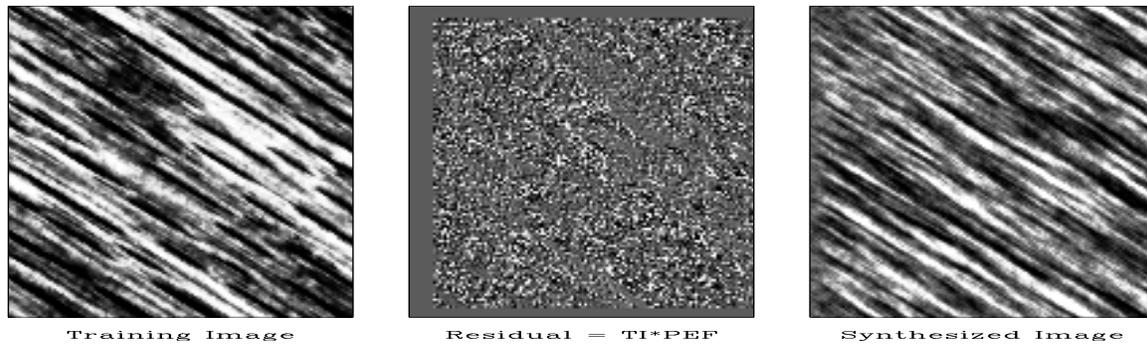


Figure 7: “Wood” image and PEF-based texture synthesis result. The synthesis result is pleasing. The PEF-based method preserves the general trend and relative scale length of the lineations in the TI. The correlation of the TI is relatively long-range, in that the lineations cross a large portion of the image, but the features are merely straight lines at one dip. [morgan1-wood-pefsyn](#) [ER]

WHY USE THE PEF?

PEF-based texture synthesis can only achieve the results of the Fourier transform method (Figures 2 and 3) in the limit $N_d \rightarrow N_d$, which is unrealistic in practical situations, where N_d is very large. Least squares estimation of the filter in this case is certainly costlier than three Fast Fourier transforms. On the other hand, if the filter size can be limited without compromising quality, which is the case for stationary, simply correlated images, then the PEF-based method is more flexible. Unlike the Fourier transform a PEF can be estimated easily when data are missing. Figure 8 shows that the PEF estimated from the incomplete data captures enough features of the data’s spectrum to make a fairly convincing texture synthesis result. The output of the PEF-based method can be of any size, while the output of a Fourier transform is generally constrained to be the same size as the input.

APPLICATIONS

PEF Estimation with incomplete data

Modern reservoir characterization efforts take a pragmatic view of collected data. Rather than wait for collection of the elusive “perfect” dataset, the desire is to incorporate a wide variety of possibly incomplete data types into a single inversion scheme (Caers and Journel, 1998). Often the only data available is spatially incomplete. Figure 8 shows the result of texture synthesis on training images with large void regions. As noted earlier, the blank areas in the center panels of the figure correspond to regions where the filter can’t fit without falling on one or more missing points. Each of the “in-bounds” data points contributes one equation to the LS estimation of the 100 or so filter coefficients. Even when well over half of the data points are removed from the training image this result shows that we can still safely estimate a filter and synthesis a believable texture.

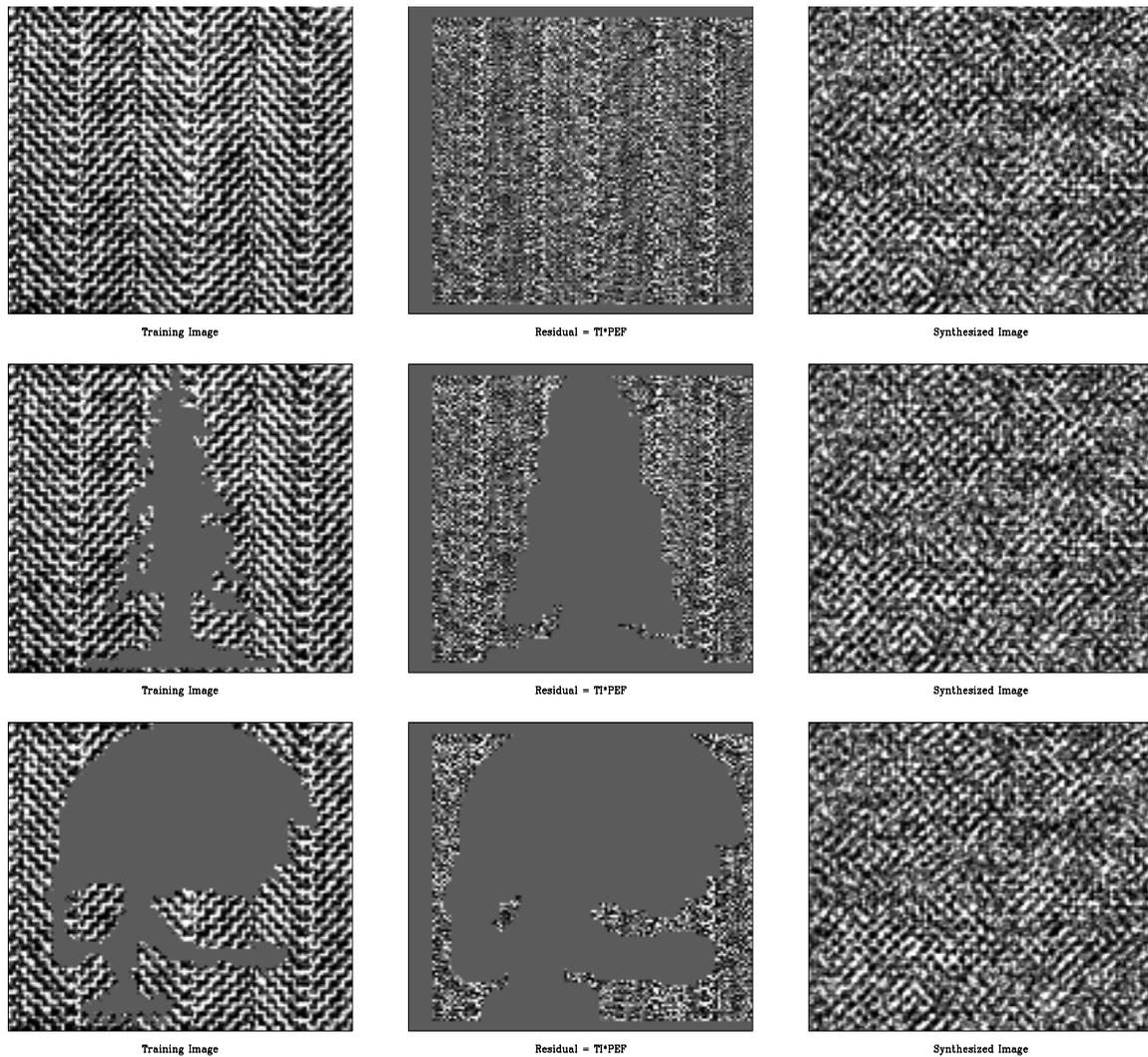


Figure 8: Comparison showing the effects of missing data on the PEF texture synthesis result, for two different “holes”. Although half or more of the equations are removed from the PEF estimation problem, the synthesized textures still capture the character of the training image. Fourier transforms are ill-defined on irregular coordinate systems, but the PEF makes an estimate of the known data’s spectrum regardless. morgan1-holes [ER]

2-D Stacked Seismic Section

Figure 9 shows the result of applying PEF texture synthesis to a 2-D stacked seismic section. The residual panel is interesting; notice uncollapsed diffraction hyperbolas, two highlighted fault planes, and also statics-like artifacts in the earlier times. PEF's easily predict straight lines (plane waves) and sinusoids, but hyperbolas and discontinuities are quite another matter.

Matthias Schwab used the “plane wave prediction” property of the PEF in his Ph.D. thesis (Schwab, 1998) to create so-called “coherency cubes” from 3-D seismic data by nonstationary convolution with small PEF's. Development of viable seismic coherency attributes merits considerable industrial interest, as evidenced by the concentration of related articles in the March, 1999 edition of *The Leading Edge*.

If a good velocity model is used, poststack migration should collapse these hyperbolas, so one measure of the fitness of a given velocity model could be the relative amount of residual energy in the data*PEF panel. Additionally, to the same end, this technique could be used to measure the relative amount of residual curvature in common reflection point (CRP) gathers, which are flattened when the correct migration velocity is used (Biondi, 1997). This preprocessing could be done quickly, for the necessary PEF's are small.

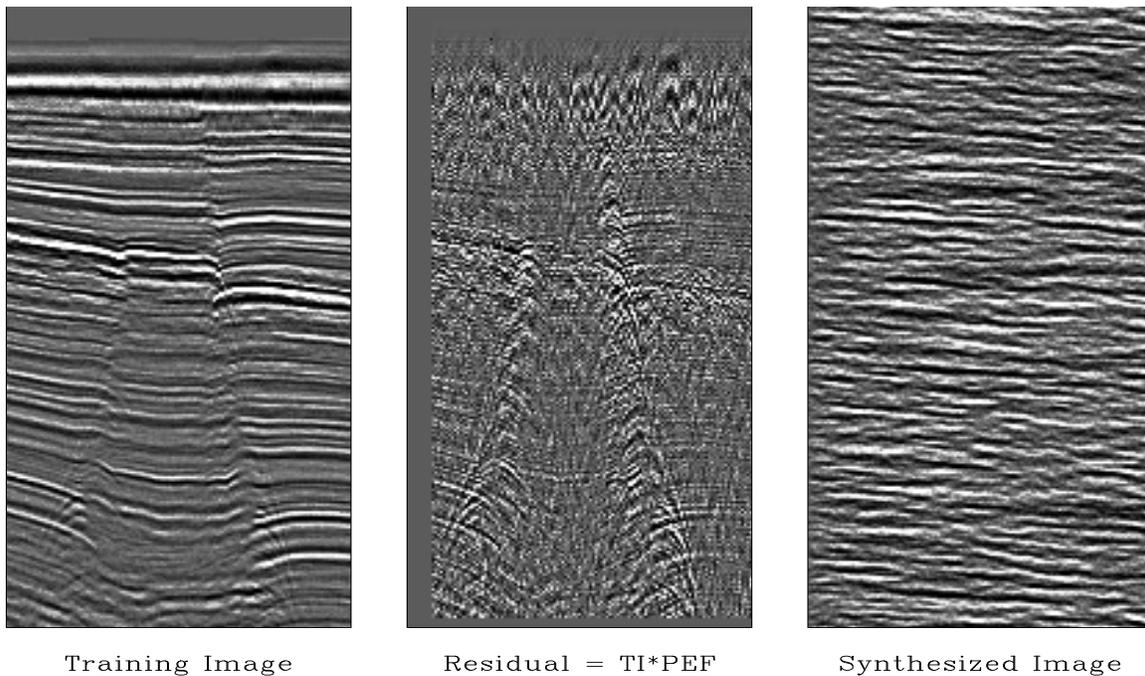


Figure 9: Stacked 2-D seismic section. morgan1-WGstack-pefsyn [ER]

Preconditioned Missing Data Infill

To fill “holes” in collected data, we have the familiar SEP formulation (Claerbout, 1999):

$$\mathbf{K}m - \mathbf{d} \approx 0 \quad (5)$$

$$\epsilon \mathbf{A} \mathbf{m} \approx 0 \quad (6)$$

[5] is the “data matching” goal, which states that the model \mathbf{m} must match the known data \mathbf{d} , while [6] is the “model smoothness” goal, where \mathbf{A} is an arbitrary roughening operator. To combat slow convergence, Claerbout (1999) preconditions with the inverse of the convolutional operator \mathbf{A} (multidimensional *deconvolution*). Provided that \mathbf{A} is minimum phase or factorizable into the product of minimum phase filters (Sava et al., 1998), the helix transform now permits stable multidimensional deconvolution. Making the change of variables $\mathbf{m} = \mathbf{A}^{-1} \mathbf{x}$, we have the equivalent preconditioned problem:

$$\mathbf{K} \mathbf{A}^{-1} \mathbf{x} - \mathbf{d} \approx 0 \quad (7)$$

$$\epsilon \mathbf{x} \approx 0 \quad (8)$$

The operator \mathbf{K} effectively maps vectors in model space into a smaller-dimension “known data space”, so it has a nonempty nullspace. Missing points in model space are completely unconstrained by \mathbf{K} , so our choice of \mathbf{A} wholly determines the behavior of the missing model points, i.e., their *texture* (Fomel et al., 1997). The PEF is a perfect choice for \mathbf{A} , as shown in Figure 10. The preconditioned, PEF-regularized result fills the hole quite believably after only 20 iterations, as opposed to the case where $\mathbf{A} = \nabla^2$, which imposes an unrealistically smooth texture on the missing model points.

DISCUSSION

The goal of this paper is not to make slick surface textures for computer games. Nonetheless, as a tutorial device, texture synthesis using the PEF is valuable, since it concretely and intuitively illustrates in two dimensions some of the fundamental concepts of autoregression which are proved only in the one dimensional case (Claerbout, 1976). In fact, some of the results shown here and in Claerbout and Brown (1999) have recently been incorporated into Jon Claerbout’s textbook, *Geophysical Estimation by Example* (1999).

Both the Fourier transform and PEF-based texture synthesis operate under the assumption that the training image is sufficiently well characterized by amplitude spectrum alone. For some images (Figures 2, 5, and 7) the assumption holds, but for others (Figures 3, 6) it is obviously violated. Real digital images and earth phenomena alike often exhibit complex spatial correlation which are modelable only with multiple point templates (Caers and Journel, 1998; Malzbender and Spach, 1993). Additionally, I have ignored the interesting subjects of nonstationarity and spatial scale variance. By scale-variant, I mean that the characteristic scale of an image’s features is not constant with respect to spatial frequency. Many methods for characterizing scale-variant images appeal to the world of wavelets for a methodology known as *multiresolution analysis* (Simoncelli and Portilla, 1998; Heeger and Bergen, 1995; Strang and Nguyen, 1997). The notion of texture synthesis for nonstationary images is ill-defined, since it amounts to a random reordering of filters estimated on locally-stationary patches, followed by deconvolution on the corresponding patches.

When the training image has missing values, as in Figure 8, the PEF-based texture synthesis method performs favorably. As shown in the missing data interpolation example (Figure

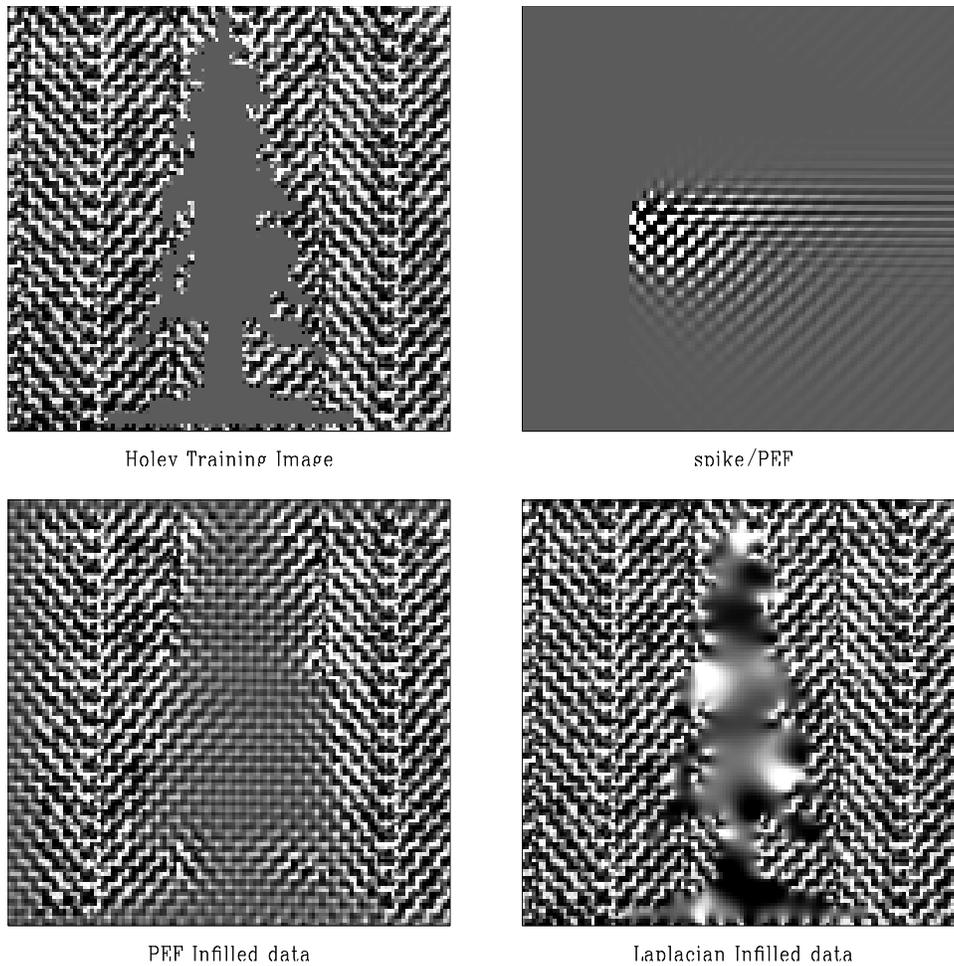


Figure 10: Clockwise from top left: Data with hole, impulse response of “inverse PEF” (deconvolution of the PEF estimated from the data and a spike), data in-filled using ∇^2 regularization, data in-filled using preconditioned PEF regularization. `morgan1-tree-hole-filled` [ER]

10), the ability of the PEF to reliably estimate the data spectrum, even with missing data, makes it an ideal regularization operator. Figure 9 illustrates the fact that the PEF primarily predicts plane waves. I proposed using a PEF residual measure to determine the viability of a given migration velocity. In general, PEF estimation/convolution might have value as a pre-processing step for a variety of applications. For instance, a very small PEF (2 columns) has a relatively large residual in the presence of conflicting dips, and thus may help in determining local filter size or patch size.

ACKNOWLEDGEMENTS

All the results in this paper were generated quite easily using Sergey Fomel's helix inversion library. All important programs have been included in the most recent release of SEPLib. The programs are conducive to curious experimentation, so I encourage the reader to use the makefile which accompanies the source code for this paper as a template.

REFERENCES

- Biondi, B. 3-d seismic imaging: <http://sepwww.stanford.edu/sep/biondo/Lectures>, 1997.
- Caers, J., and Journel, A. G., 1998, Stochastic reservoir modeling using neural networks trained on outcrop data.: SPE Annual Technical Conference and Exhibition, 49026.
- Castleman, K. R., 1996, Digital image processing: Prentice-Hall, Inc., Upper Saddle River, NJ.
- Claerbout, J., and Brown, M., 1999, Two-dimensional textures and prediction-error filters: 61st Ann. Mtg., Eur. Assoc. Expl. Geophys.
- Claerbout, J. F., 1976, Fundamentals of Geophysical Data Processing: Blackwell.
- Claerbout, J. F., 1998, Multi-dimensional recursive filtering via the helix: Geophysics, **63**, no. 5, 1532–1541.
- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, <http://sepwww.stanford.edu/sep/prof/>.
- Clapp, R. G., Sava, P., and Claerbout, J. F., 1998, Interval velocity estimation with a null-space: SEP-97, 147–156.
- Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: SEP-95, 15–25.
- Heeger, D. J., and Bergen, J. R., 1995, Pyramid-based texture analysis/synthesis: Computer Graphics Proceedings, 229–238.
- IEEE, 1990, IEEE standard glossary of image processing and pattern recognition terminology, standard 610.4-1990: IEEE Press, New York.

Malzbender, T., and Spach, S., 1993, A context sensitive texture nib: Proceedings of Computer Graphics International, 151–163.

Mao, S. Psych 267 final project - texture synthesis:.

Sava, P., Rickett, J., Fomel, S., and Claerbout, J., 1998, Wilson-Burg spectral factorization with application to helix filtering: SEP-97, 343–351.

Schwab, M., 1998, Enhancement of discontinuities in seismic 3-D images using a Java estimation library: Ph.D. thesis, Stanford University.

Simoncelli, E., and Portilla, J., 1998, Texture characterization via joint statistics of wavelet coefficient magnitudes: 5th IEEE Int'l Conf on Image Processing.

Strang, G., and Nguyen, T., 1997, Wavelets and Filter Banks: Wellesley-Cambridge Press.

Short Note

Polarity and PEF regularization

Jon Claerbout¹

INTRODUCTION

We address the puzzle of seismic polarity. Why do we rarely observe it clearly and how could we be more systematic about trying to observe polarity? This puzzle will lead us to long prediction-error filters. Being long, they require many data samples. If such a filter is nonstationary, we might have an inadequate number of fitting equations. Then we need regularization. Here we consider some examples and consider an efficient way to regularize the filter estimation.

EXAMPLES

We sometimes observe seismic signal polarity. We rarely observe it in soft clastic areas like much of the Gulf of Mexico, but often see it in more hardened areas like the North Sea. Generally, polarity is much more likely to be recognized in the ideal conditions we encounter with marine data than with land data. Even where it is possible to recognize, it is not especially noticeable because even an ideal seismic impulse is spread out by the shot and receiver ghosts to the (low cut) three-phase event $(1, -2, 1)$. An attractive example that I recently encountered is that of Christine Ecker in Figure 1.

In principle, a causal double integration converts this triplet into an impulse which theoretically should make for ready recognition of polarity. On the other hand, we need to think about whether the recording equipment actually records the low frequencies that double integration would amplify and we need to think about low-frequency noise.

I have noticed that marine recording systems often record seismic energy right down to zero frequency. This is evidenced by the appearance of water surface gravity waves with speeds of about 30 km/hour, speeds so slow that we can hardly distinguish their speed from zero. I have not made quantitative measurements on Figure 2, but I believe it to be consistent with long ocean swells moving about 20 miles/hour or 30 km/hour. Distance traveled in meters over four seconds will be $4 \times 30 \times 1000/3600 = 33\text{m}$. These become visible late on the data where the t^2 gain brings them up. Their time signature is roughly a growing ramp function.

¹email: jon@sep.stanford.edu, <http://sepwww.stanford.edu/sep/jon/>

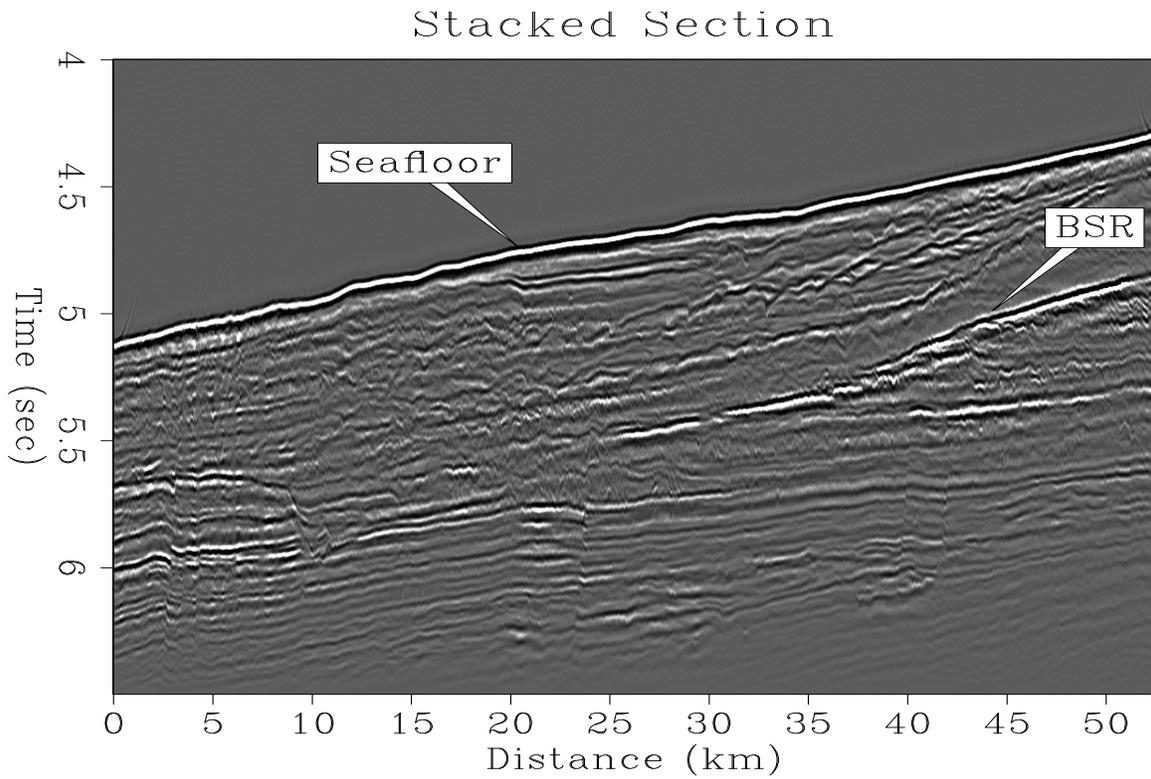


Figure 1: Christine Ecker's BSR data from her PhD dissertation. Most reflectors are recognizable as "black-white-black". The exception is the BSR reflector which is "white-black-white".
 [jon3-christine] [NR]

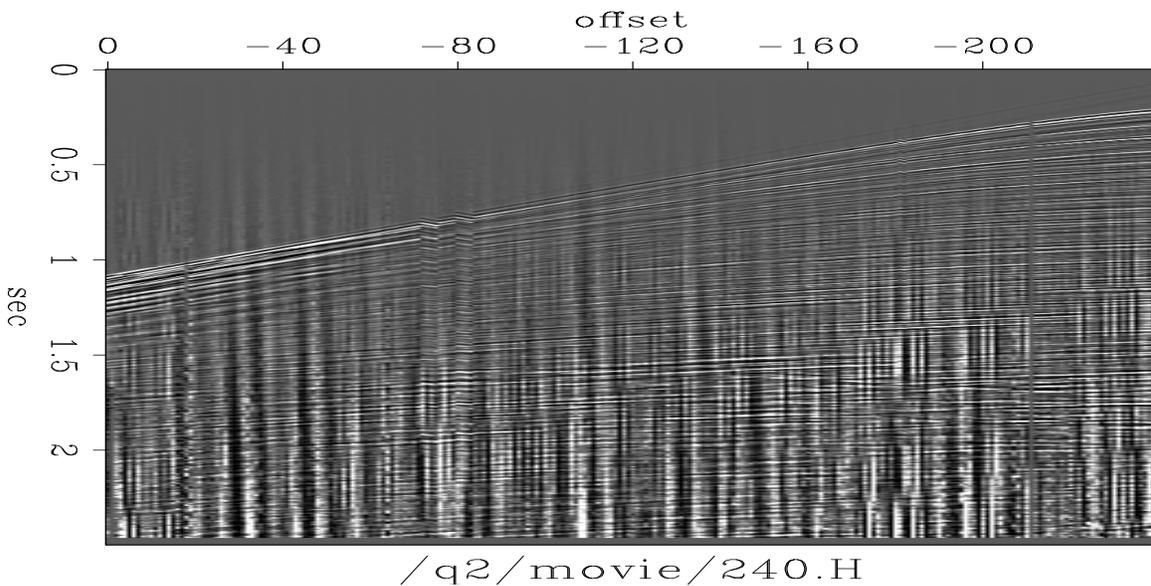


Figure 2: Marine data with t^2 gain. [jon3-gravity] [ER]

From these two data sets we recognize two different goals: One data set says we should use double integration to convert the $(1, -2, 1)$ ghost to a pulse. The other data set says we should use double differentiation to suppress the growing ramp function of the surface gravity waves.

On more careful study we can see that the goals are not *exactly* opposite: The ramp function rises over the full two seconds of the data. Let us say its frequency is about .2 Hz. On the other hand, the triplet $(1, -2, 1)$ is not merely spread over three samples but perhaps 10 samples or 40ms so we associate it with a frequency of about 25 Hz. Now imagine a data set with both the phenomena we see on Figure 1 and on Figure 2. It would show both phenomena. We can analyze this in the frequency domain or the time domain. The composite spectrum would be the sum of their spectra. The spectrum would grow near zero frequency as $1/\omega^2$ and grow near .2 Hz as ω^2 . We need to raise the spectrum between those two limits (or suppress it beyond those limits). In the time domain we notice that while the ramp function is approximately as long as the seismogram itself, the duration of the triplet $(1, -2, 1)$ is about 10 samples or 40ms. Perhaps the filter with the correct spectral characteristic in the time domain would begin as a ramp for 40ms and then gently bend toward negative values a second or so later. Notice that this is a fairly long filter

We could examine the specifics more carefully but fortunately, the autoregression method of deconvolution addresses the general problem without requiring specific information. Notice a key ingredient here: The required autoregression filter, the PEF, is *long* because any approximation to double integration must be long. The filter length raises interesting issues that I have previously not given adequate attention to.

Basically we are talking about a decon filter whose length is a significant fraction of the trace length. There is a real danger that we might “over fit”, i.e., use insufficient data to estimate the filter and find our filter adapting to the geology instead of adapting to the data-recording environment. The way to overcome this problem is to use a lot of data.

MINIMUM-PHASE EQUIVALENT TRAINING DATA SET

This leads to a technique that is new to me. I’ll describe it first in the one-dimensional world. (In real life, multidimensional cases might be more interesting, for example where dip spectra change rapidly.) The basic problem is to define the appropriate regularization for a prediction-error filter (PEF). Regularization is ordinarily regarded as supplying a prior statement about the model, in this case, about the autoregression filter. We don’t think about PEFs as being “physical” and the correct prior model and its covariance are not immediately obvious. The answer is that the prior PEF is nothing more and nothing less than the solution to the autoregression equations for a prior “universal” data set. In practice, it amounts to having a “training” data set. I have noticed an efficient way to merge the information of the training data set with the “too-small” local data set. Given a data set packed in an operator \mathbf{D} and likewise a training data set \mathbf{T} , we formulate the fitting goals for finding the PEF \mathbf{a} by using a

constraint matrix \mathbf{K} (an identity matrix except for the (1,1) element which is zero).

$$\begin{aligned} \mathbf{0} &\approx \mathbf{DKa} \\ \mathbf{0} &\approx \mathbf{TKa} \end{aligned} \quad (1)$$

In principle, the training data (and hence the matrix \mathbf{T}) is very large. Consider however a spectral factorization of the training data set. Say $\mathbf{T}'\mathbf{T} = \mathbf{B}'\mathbf{B}$ where \mathbf{b} is a minimum-phase spectral factorization of the training data set (and \mathbf{B} is the packing of \mathbf{b} into a convolution operator). For me, this is a new idea, that we express the prior information as a “training wavelet” \mathbf{b} that we find by spectral factorization of a “universal” data set. The idea is that we then find our “local” PEF by fitting the goals

$$\begin{aligned} \mathbf{0} &\approx \mathbf{DKa} \\ \mathbf{0} &\approx \mathbf{BKa} \end{aligned} \quad (2)$$

The result of fitting (2) is theoretically equal to that of (1) but computationally (2) is potentially much easier training wavelet is much more compact (because it is minimum-phase) than the full training data set.

DISCUSSION

Sometimes data is stationary. Then Fourier analysis gives a more efficient approach. The ideas above are more relevant to cases where stationarity is less valid. Likewise, when the required filter is very long, comparable to a trace length, Fourier analysis would be more appropriate.

On the other hand, with spatial filtering applications a local PEF is more appropriate. In my book GEE, I explain how to build a time-variable PEF. It seems an alternative could be based on a training data set that varies locally. I see this as perhaps theoretically superior. In the GEE example, the filter itself is stated to vary smoothly. Now I would be proposing that the training data set be varying smoothly. In general, PEFs tend to “look bad” because their frequency content is inverse to that of the signal. This could mean that smoothing a PEF is not nearly such a good idea as using a training data set. Imagine we seek a new PEF upon the arrival of each new trace, or perhaps even upon the arrival of each new data point. Naturally, the Wilson-Burg spectral factorization method might be helpful. Generally however, I am not sure how to proceed.

Spectral factorization revisited

Paul Sava and Sergey Fomel¹

ABSTRACT

In this paper, we review some of the iterative methods for the square root, showing that all these methods belong to the same family, for which we find a general formula. We then explain how those iterative methods for real numbers can be extended to spectral factorization of auto-correlations. The iteration based on the Newton-Raphson method is optimal from the convergence stand point, though it is not optimal as far as stability is concerned. Finally, we show that other members of the iteration family are more stable, though slightly more expensive and slower to converge.

INTRODUCTION

Spectral factorization has been recently revived by the advent of the helical coordinate system. Several methods are reported in the literature, ranging from Fourier domain methods, such as Kolmogoroff's (Claerbout, 1992; Kolmogoroff, 1939), to iterative methods, such as the Wilson-Burg method (Claerbout, 1999; Wilson, 1969; Sava et al., 1998).

In this paper, after reviewing the general theory of root estimation by iterative methods, we derive a general square root relationship applicable to both real numbers and to auto-correlation functions. We introduce a new spectral factorization relation and show its relation to the Wilson-Burg method.

THE SQUARE ROOT OF REAL NUMBERS

This section briefly reviews some well known square root iterative algorithms, and derives the Newton-Raphson and Secant methods. It also shows that Muir's iteration for the square root (Claerbout, 1995) belongs to the same family of iterative methods, if we make an appropriate choice of the generating function.

¹email: paul@sep.stanford.edu, sergey@sep.stanford.edu

Root-finding recursions

Given a function $f(x)$ and an approximation for one of its roots x_n , we can find a better approximation for the root by linearizing the function around x_n

$$f(x) \approx f(x_n) + (x_{n+1} - x_n)f'(x_n)$$

and by setting $f(x)$ to be zero for $x = x_{n+1}$. We find that

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (1)$$

1. Newton-Raphson's method for the square root

A common choice of the function f is $f(x) = x^2 - s$. This function has the advantage that it is easily differentiable, with $f'(x) = 2x$. The recursion relation thus becomes

$$x_{n+1} = x_n - \frac{x_n^2 - s}{2x_n} = \frac{x_n}{2} + \frac{s}{2x_n}$$

or

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{s}{x_n} \right)$$

or, after rearrangement,

$$x_{n+1} = \frac{s + x_n^2}{2x_n} \quad (2)$$

The recursion (2) converges to $\pm\sqrt{s}$ depending on the sign of the starting guess $x_0 \neq 0$.

2. Secant method for the square root

A variation of the Newton-Raphson method is to use a finite approximation of the derivative instead of the differential form. In this case, the approximate value of the derivative at step n is

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

For the same choice of the function f , $f(x) = x^2 - s$, we obtain

$$x_{n+1} = x_n - \frac{x_n^2 - s}{x_n + x_{n-1}}$$

and

$$x_{n+1} = \frac{s + x_n x_{n-1}}{x_n + x_{n-1}} \quad (3)$$

In this case, recursion (3) also converges to $\pm\sqrt{s}$ depending on the sign of the starting guesses x_0 and x_1 .

3. Muir’s method for the square root

Another possible iterative relation for the square root is Francis Muir’s, described by Jon Claerbout (1995):

$$x_{n+1} = \frac{s + x_n}{x_n + 1} \tag{4}$$

This relation belongs to the same family of iterative schemes as Newton and Secant, if we make the following special choice of the function $f(x)$ in (1):

$$f(x) = |x + \sqrt{s}|^{\frac{\sqrt{s}-1}{2\sqrt{s}}} |x - \sqrt{s}|^{\frac{\sqrt{s}+1}{2\sqrt{s}}} \tag{5}$$

Figure 1 is a graphical representation of the function $f(x)$.

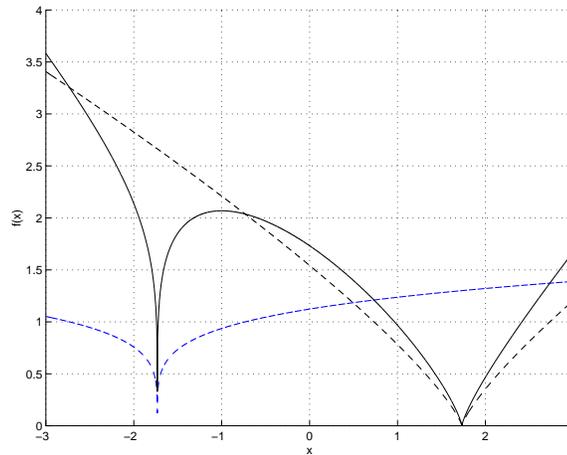


Figure 1: The graph of the function defined in Equation (1) used to generate Muir’s iteration for the square root (solid line). The dashed lines are the plot of the two factors in the equation . paul3-muf [CR]

4. A general formula for the square root

From the analysis of equations (2), (3), and (4), we can derive the following general form for the square root iteration:

$$x_{n+1} = \frac{s + x_n \gamma}{x_n + \gamma} \tag{6}$$

where γ can be either a fixed parameter, or the value of the iteration at the preceding step, as shown in Table 1. The parameter γ is the estimate of the square root at the given step (Newton), the estimate of the square root at the preceding step (Secant), or a constant value (Muir). Ideally, this value should be as close as possible to \sqrt{s} .

The convergence rate

We can now analyze which of the particular choices of γ is more appropriate as far as the convergence rate is concerned.

Table 1: Recursions for the square root

	γ	Recursion
Muir	1	$x_{n+1} = \frac{s+x_n}{x_n+1}$
Secant	x_{n-1}	$x_{n+1} = \frac{s+x_n x_{n-1}}{x_n+x_{n-1}}$
Newton	x_n	$x_{n+1} = \frac{s+x_n^2}{2x_n}$
Ideal	\sqrt{s}	$x_{n+1} = \frac{s+x_n \sqrt{s}}{x_n+\sqrt{s}}$

If we consider the general form of the square root iteration

$$x_{n+1} = \frac{s + x_n \gamma}{x_n + \gamma}$$

we can estimate the convergence rate by the difference between the actual estimation at step $(n + 1)$ and the analytical value \sqrt{s} . For the general case, we obtain

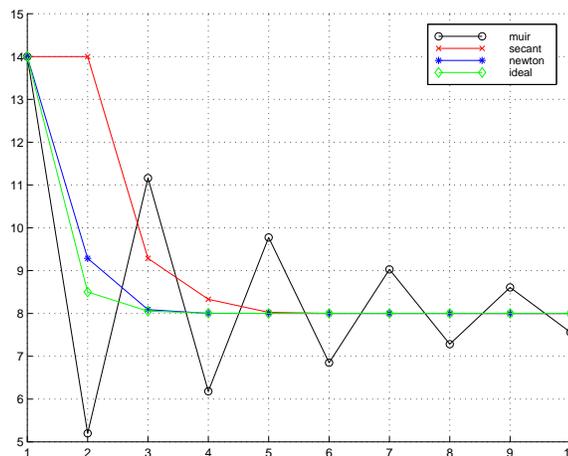
$$x_{n+1} - \sqrt{s} = \frac{s + \gamma x_n - x_n \sqrt{s} - \gamma \sqrt{s}}{x_n + \gamma}$$

or

$$x_{n+1} - \sqrt{s} = \frac{(x_n - \sqrt{s})(\gamma - \sqrt{s})}{x_n + \gamma} \quad (7)$$

The possible selections for γ from Table 1 clearly show that the recursions described in the

Figure 2: Convergence plots for different recursive algorithms, shown in Table 1. `paul3-sqroot` [CR]



preceding subsection generally have a linear convergence rate (that is, the error at step $n + 1$ is proportional to the error at step n), but can converge quadratically for an appropriate selection of the parameter γ , as shown in Table 2. Furthermore, the convergence is faster when γ is closer to \sqrt{s} .

Table 2: Convergence rate

	γ	Convergence
Muir	1	linear
Secant	x_{n-1}	quasi-quadratic
Newton	x_n	quadratic

We therefore conclude that Newton's iteration has the potential to achieve the fastest convergence rate. Ideally, however, we could use a fixed γ which is a good approximation to the square root. The convergence would then be slightly faster than for the Newton-Raphson method, as shown in Figure 2.

SPECTRAL FACTORIZATION

We can now extend the equations derived for real numbers to polynomials of Z , with $Z = e^{i\omega t}$, and obtain spectral factorization algorithms similar to the Wilson-Burg method (Sava et al., 1998), as follows:

$$X_{n+1} = \frac{S + X_n \bar{G}}{\bar{X}_n + \bar{G}} \quad (8)$$

If L represents the limit of the series in (8),

$$L\bar{L} + L\bar{G} = S + L\bar{G}$$

and so

$$L\bar{L} = S$$

Therefore, L represents the causal or anticausal part of the given spectrum $S = X\bar{X}$.

Table 3 summarizes the spectral factorization relationships equivalent to those established for real numbers in Table 1.

The convergence properties are similar to those derived for real numbers. As shown above, the Newton-Raphson method should have the fastest convergence.

A COMPARISON WITH THE WILSON-BURG METHOD

For reasons of symmetry, we can take Newton's relation from Table 3

$$X_{n+1} = \frac{S + X_n \bar{X}_n}{2\bar{X}_n}$$

Table 3: Spectral factorization

General	$X_{n+1} = \frac{S+X_n\bar{G}}{\bar{X}_n+\bar{G}}$
Muir	$X_{n+1} = \frac{S+X_n}{\bar{X}_n+1}$
Secant	$X_{n+1} = \frac{S+X_n\bar{X}_{n-1}}{\bar{X}_n+\bar{X}_{n-1}}$
Newton	$X_{n+1} = \frac{S+X_n\bar{X}_n}{2\bar{X}_n}$
Ideal	$X_{n+1} = \frac{S+X_n\sqrt{S}}{\bar{X}_n+\sqrt{S}}$

and convert it to

$$\frac{X_{n+1}}{2X_n} = \frac{S + X_n\bar{X}_n}{(2X_n)(2\bar{X}_n)}.$$

We can then consider a symmetrical relation where on the left side we insert the anticausal part of the spectrum, and obtain

$$\frac{\bar{X}_{n+1}}{2\bar{X}_n} = \frac{S + X_n\bar{X}_n}{(2X_n)(2\bar{X}_n)}.$$

Finally, we can sum the preceding two equations and get

$$\boxed{\frac{X_{n+1}}{2X_n} + \frac{\bar{X}_{n+1}}{2\bar{X}_n} = \frac{2S + X_n\bar{X}_n + \bar{X}_nX_n}{(2X_n)(2\bar{X}_n)}} \quad (9)$$

which can easily be shown to be equivalent to the Wilson-Burg relation

$$\frac{X_{n+1}}{X_n} + \frac{\bar{X}_{n+1}}{\bar{X}_n} = 1 + \frac{S}{X_n\bar{X}_n} \quad (10)$$

In an analogous way, we can take the general relation from Table 3

$$X_{n+1} = \frac{S + X_n\bar{G}}{\bar{X}_n + \bar{G}}$$

and convert it to

$$\frac{X_{n+1}}{X_n + G} = \frac{S + X_n\bar{G}}{(X_n + G)(\bar{X}_n + \bar{G})}$$

We can then consider a symmetrical relation where on the left side we insert the anticausal part of the spectrum, and obtain

$$\frac{\bar{X}_{n+1}}{\bar{X}_n + \bar{G}} = \frac{S + \bar{X}_nG}{(X_n + G)(\bar{X}_n + \bar{G})}$$

Finally, we can sum the preceding two equations and get

$$\frac{X_{n+1}}{X_n + G} + \frac{\bar{X}_{n+1}}{\bar{X}_n + \bar{G}} = \frac{2S + X_n \bar{G} + \bar{X}_n G}{(X_n + G)(\bar{X}_n + \bar{G})} \quad (11)$$

Equation (11) represents our general formula for spectral factorization. If we consider the particular case when G is X_n , we obtain equation (10), which we have shown to be equivalent to the Wilson-Burg formula.

From the computational standpoint, our equation is more expensive than the Wilson-Burg because it requires two more convolutions on the numerator of the right-hand side. However, our equation offers more flexibility in the convergence rate. If we try to achieve a quick convergence, we can take G to be X_n and get the Wilson-Burg equation. On the other hand, if we worry about the stability, especially when some of the roots of the auto-correlation function are close to the unit circle, and we fear losing the minimum-phase property of the factors, we can take G to be some damping function, more tolerant of numerical errors.

Moreover, by using the Equation (11), we can achieve fast convergence in cases when the auto-correlations we are factorizing have a very similar form, for example, in nonstationary filtering. In such cases, the solution at the preceding step can be used as the G function in the new factorization. Since G is already very close to the solution, the convergence is likely to occur quite fast.

CONCLUSIONS

The general iterative formula for the square root that we derived can be extended to the factorization of the auto-correlation functions. The Wilson-Burg algorithm is a special case of our more general formula. Using such a general formula provides flexibility in choosing between fast convergence and stability. We can achieve fast convergence when factorizing auto-spectra that have a very similar form. This improvement in convergence rate can have a useful application, for instance, in nonstationary preconditioning.

ACKNOWLEDGMENTS

We thank Jon Claerbout, who brought Muir's iterative scheme to our attention, and suggested its application to spectral factorization.

REFERENCES

Claerbout, J. F., 1992, *Earth Soundings Analysis: Processing Versus Inversion*: Blackwell Scientific Publications.

Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, <http://sepwww.stanford.edu/sep/prof/>.

Kolmogoroff, A. N., 1939, Sur l'interpolation et extrapolation des suites stationnaires: C.R. Acad.Sci., **208**, 2043–2045.

Sava, P., Rickett, J., Fomel, S., and Claerbout, J., 1998, Wilson-Burg spectral factorization with application to helix filtering: SEP-97, 343–351.

Wilson, G., 1969, Factorization of the covariance generating function of a pure moving average process: SIAM J. Numer. Anal., **6**, no. 1, 1–7.

Helix derivative and low-cut filters' spectral feature and application

Yi Zhao¹

ABSTRACT

A helix derivative filter can be used to roughen an image and thus enhance its details. Unlike the conventional derivative operator, the helix derivative filter has no direction orientation. I present the enhanced helix/low-cut derivative filters, in which the zero frequency response is adjustable. I analyze the quantitative effects of the adjustable parameters on the filter spectrum and propose guidelines for choosing parameters. I also show some roughened images created by the enhanced filters.

INTRODUCTION

The derivative operator is a useful tool in removing the low frequency components to enhance the details of an image (Claerbout, 1998a). However, the conventional derivative operator has a particular derivative direction. This is not desirable in some cases where the contour map has circular features, so what needed is a derivative operator without a specific direction. One isotropic alternative, the Laplacian operator, often cuts low frequencies too strongly.

The introduction of helical coordinate system (Claerbout, 1997) gives us the ability to solve the problem. The Kolmogoroff algorithm of spectral factorization (Claerbout, 1976) enables us to derive a helix derivative filter from its autocorrelation, the negative Laplacian operator (Claerbout, 1998b). The Wilson-Burg method of spectral factorization (Sava et al., 1998) makes it more convenient to compute the coefficients of a finite-length helix derivative filter.

The helix low-cut filter is another operator useful in enhancing the digital images and also has the circularly symmetric spectrum. Unlike the helix derivative filter, the helix low-cut filter removes the frequency components below an adjustable cut-off frequency from the image, and has a flat response at high frequencies.

In this paper, I present the enhanced helix derivative and low-cut filters with more adjustable parameters. Then I analyze the influence on the filter spectrum for these parameters, and present guidelines for choosing the parameters for image processing, using the examples from GEE.

¹email: yi@sep.stanford.edu

ENHANCED HELIX FILTERS

The autocorrelation of the helix derivative filter \mathbf{H} is the negative of the finite-difference representation of the Laplacian operator ∇^2 ,

$$\mathbf{H}'\mathbf{H} = \mathbf{R} = -\nabla^2 \quad (1)$$

The coefficients of the causal helix filter \mathbf{H} can be found by spectral factorization.

The helix low-cut filter H/D is designed by doing two spectral factorizations, one for the numerator of H , and another for the denominator of D . It is expressed by

$$\frac{k^2}{k^2 + k_0^2} \approx \frac{-\nabla^2}{-\nabla^2 + k_0^2} \approx \frac{\overline{H}H}{\overline{D}D} \quad (2)$$

where k is the frequency², k_0 affects the cut-off frequency, \overline{H} and \overline{D} are the conjugate anti-causal filters of H and D .

Both the filters do not remove the zero frequency completely, degrading the contrast and details of the roughened image. A way to solve this problem as suggested by Claerbout, is to rescale all the coefficients of H with nonzero lags by a . If s is the sum of all the coefficients with nonzero lag (which are all negative), a is expressed by

$$a = 1 + \left(\frac{1}{|s|} - 1 \right) \rho \quad (3)$$

$\rho = 0$ denotes the original unscaled filter, while $\rho = 1$ guarantees the filter really removes the zero frequency component.

Now I have the enhanced helix derivative with adjustable parameters n_a and ρ , the enhanced helix low-cut filter with n_a , k_0 and ρ . Here n_a is the half length of the helix filter.

Compared with the conventional helix filters, the enhanced filters have a new adjustable parameter, ρ .

HELIX DERIVATIVE FILTER

First I apply the enhanced helix derivative filter to some familiar images and check the effects of adjustable parameters n_a and ρ .

Figure 1 shows the views of the Sea of Galilee created by helix derivative filters with different n_a and ρ . The top two plots (with $\rho = 0$) indicate that as the half filter length n_a increases, the contrast of the image increases. This means the longer filter keeps less low frequency components, indicating a lower zero-frequency response. The vertical plots (with same n_a) indicate that as ρ increases, the contrast increases also, as expected from the definition of the enhanced helix filter equation.

²The frequencies in this paper are all scaled frequencies, π is the Nyquist frequency.

To compare the properties of the filters with different n_a other than the zero-frequency response, I set $\rho = 1$ in the bottom two plots, so that they have the same zero-frequency response. The plots are very similar, and the difference between them is very weak.

As the examples above show, the zero-frequency response of the enhanced helix derivative filter decreases as n_a or ρ increases. Thus filters with different n_a may create the same results by adjusting ρ , which provides the possibility of using a short filter instead of a long filter in image processing and reduce the computational cost greatly. By adjusting ρ , the enhanced short filter can reach a low level of zero-frequency response, which is done conventionally by using a long filter.

Since the short filter is equivalent to the long filter when adjusting ρ , now ρ plays the key role in image processing for the enhanced helix derivative filter, and n_a is not an important parameter. When the high frequency component is too weak compared with the low frequency, a large ρ (≈ 1) should be chosen in order resolve details in the image. Otherwise, a small ρ would be suitable. Based on the balance of the computation cost and response symmetry, I recommend the use of $n_a \approx 8$.

From the quantitative analysis of spectra of the enhanced helix derivative filter (see the appendix A), I can find ρ according to empirical formula

$$\rho = 1 - \frac{I_0 n_a}{0.44} \quad (4)$$

where I_0 is the factor of the zero-frequency component to be preserved.

For the view of the Sea of Galilee, the high frequency component is very weak, so I chose $I_0 = 0.002$, $n_a = 8$, and $\rho = 0.96$. Figure 2 shows the preferred and gradient roughened results.

HELIX LOW-CUT FILTER

I check the effects of adjustable parameters k_0 , n_a and ρ on roughened images. The quantitative analysis of the effects on the filter spectrum is provided in appendix A.

When comparing the effects of k_0 and n_a , I set $\rho = 1$ to make sure the filters have the same zero-frequency response. Figure 3 shows the Bay Area map created with different k_0 . As k_0 increases, more low-frequency components were removed and the detailed structure turns out to be the main focus of the map. This indicates the cut-off frequency increases as k_0 increases. In other words, k_0 governs the cut-off frequency. When k_0 remains the same, filters with different n_a can create very similar results if the zero-frequency response is the same by adjusting ρ , as shown in the middle and bottom plots in Figure 4. As expected, ρ controls the zero-frequency response. The larger ρ leads to higher contrast, as shown in the top and middle plots in Figure 4.

Among the three adjustable parameters, n_a is the least important one because the long filter can be replaced with a short one by adjusting ρ . k_0 controls the cut-off frequency and

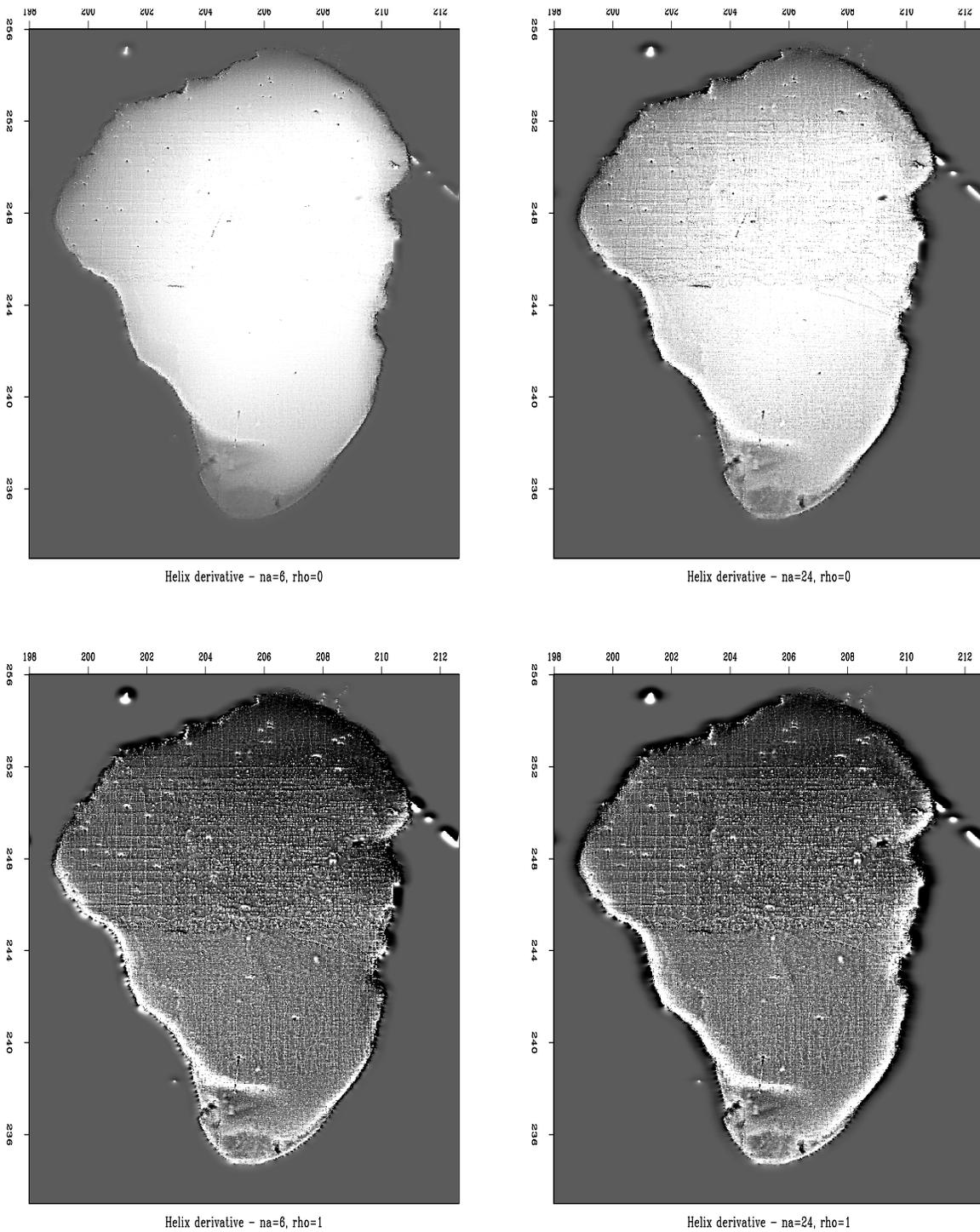


Figure 1: The Views of the Sea of Galilee roughened with helix derivative filters. Each two horizontal plots have the same ρ ; each two vertical plots have the same n_a . As n_a or ρ increases, the zero-frequency response decreases. yi1-gal-drv-ar [ER]

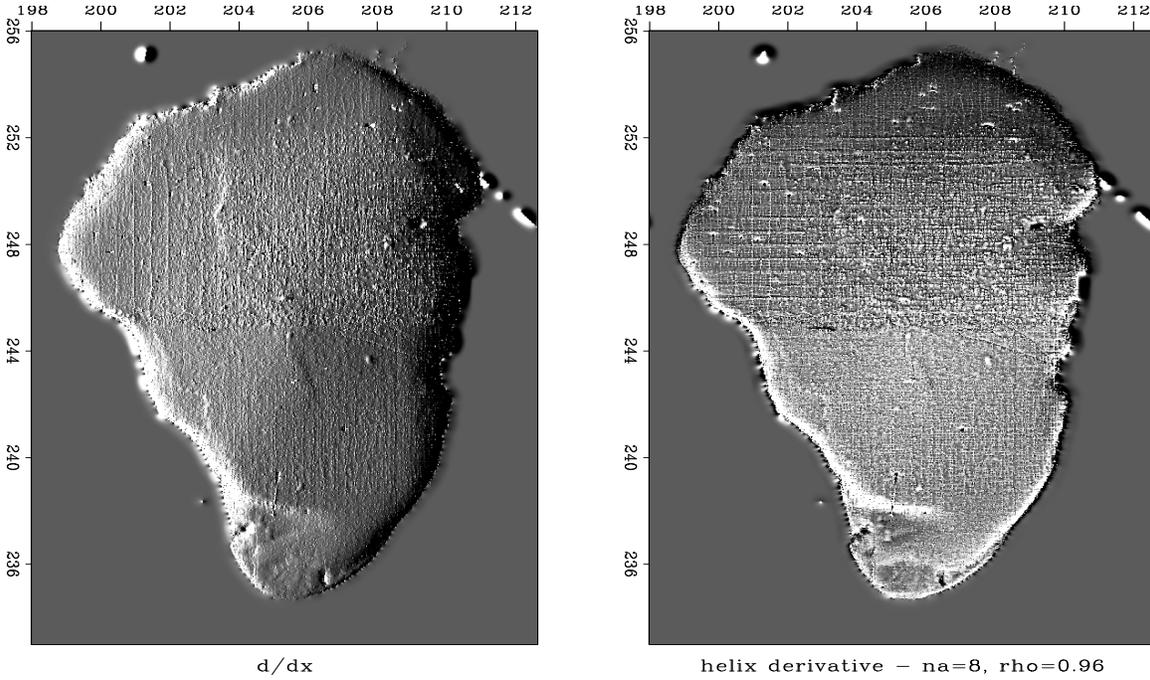


Figure 2: The views of the Sea of Galilee. The left is with the gradient operator d/dx ; the right is with the helix derivative filter. The helix filter enhances the details both along the vertical and horizontal direction. `yil-gal-drv-res` [ER]

ρ controls the zero-frequency response. It is hard to tell which one is more important if I use only this information. From the quantitative analysis, I know k_0 affects zero-frequency response significantly, but ρ does not have such an influence on cut-off frequency. So k_0 is the most important parameter.

For the enhanced helix low-cut filter, it is very reasonable to choose parameter k_0 first, then ρ and n_a .

When roughening the image with the helix low-cut filter, the key point is to choose the cut-off frequency f_0 or the adjustable parameter k_0 . My suggestion is that if the lowest frequency component to be preserved is f_L , k_0 should be

$$k_0 = \frac{2}{3} f_L \quad (5)$$

Therefore, the frequency far below f_L would be cut off completely, and the component near f_L would not be affected too much.

Since the short filter is made equivalent to the long filter by adjusting ρ , I suggest the use of $n_a \approx 16$ based on the balance of computation costs and symmetric features.

With k_0 and n_a determined, I can find ρ according to Equation (A-7)

$$\rho = 1 - \frac{I_0 n_a k_0}{0.82} \quad (6)$$

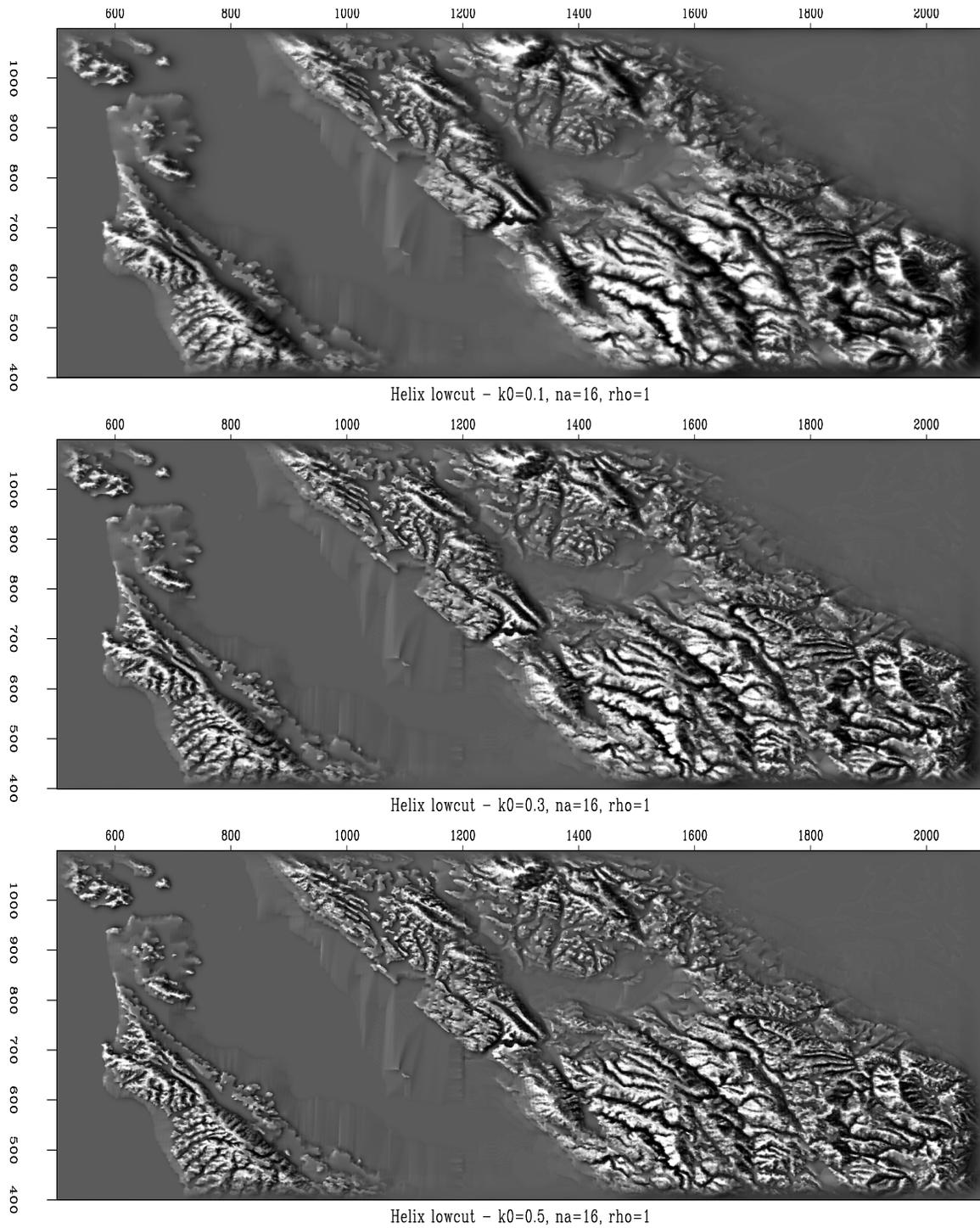


Figure 3: Bay Area maps roughened by helix low-cut filters with different k_0 . The top is $k_0 = 0.1$, the middle is $k_0 = 0.3$, the bottom is $k_0 = 0.5$. $n_a = 16$, $\rho = 1$. As k_0 increases, the cut-off frequency increases. `yi1-bay-lct-k-a16r1` [ER]

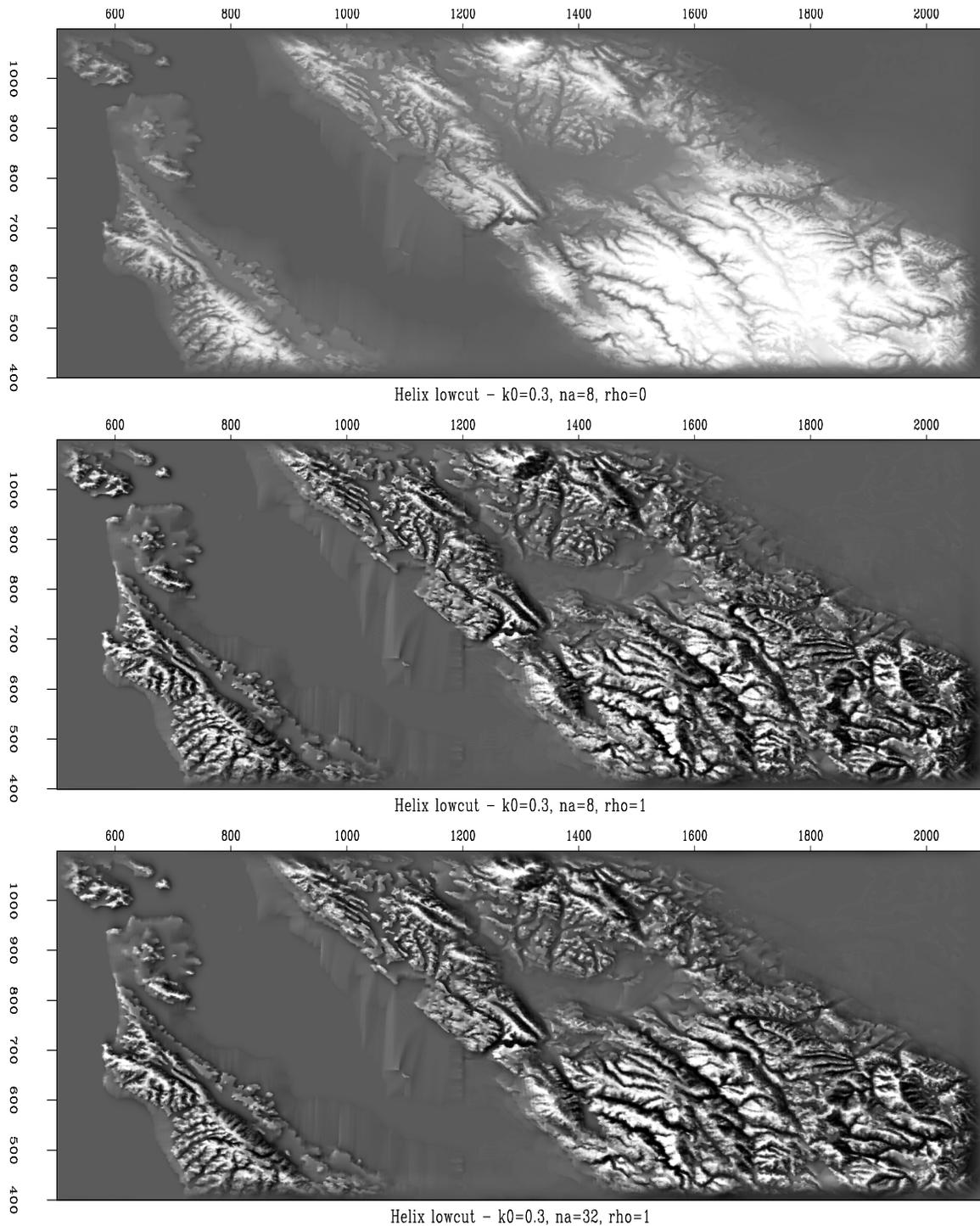


Figure 4: Bay Area maps roughened by helix low-cut filters with different filter length. The top is $n_a = 8$ and $\rho = 0$, the middle is $n_a = 16$ and $\rho = 1$, the bottom is $n_a = 32$. $k_0 = 0.3$, $\rho = 1$. $k_0 = 0.3$. As ρ increases, the zero-frequency response decreases. When the zero-frequency response remains the same, the difference of n_a does not affect results. yi1-bay-lct-ar-k3 [ER]

If $\rho = 1$, the zero-frequency is removed completely and leads to the highest contrast in the roughened image.

Figure 5 consists of three maps of the Bay Area. The top portion is a topographic map of the Bay Area. The bottom plot is the preferred result. From one slice of the Bay Area topographic map, I know one main low frequency component is about 0.4. So I chose $k_0 = 0.3$ to remove the lower frequency. I chose $n_a = 16$. In order to obtain the highest contrast, I chose $\rho = 1$. The middle one is the reference plot with $k_0 = 0.1$, $n_a = 16$ and $\rho = 1$. I notice that the bottom plot removes more low frequency components than the middle one and has clearer details, as predicted by the theory.

Figure 6 consists of a normal mammogram and the roughened images. The main low frequency component of the mammogram slice is about 0.3, so I choose $k_0 = 0.2$ in the right plot as the preferred result, and use $k_0 = 0.1$ in the middle as reference. I use $\rho = 1$ to achieve the highest contrast.

DERIVATIVE VERSUS LOW-CUT

Both helix derivative and low-cut filters can be used to enhance the details of images. But which one is better? It depends on the situation.

- Focus of the structure

Both helix derivative and low-cut filters cut off the zero-frequency component, but the functions of the frequency are different. The derivative filter's response is approximately a linear function of frequency, while the low-cut filter's response has a hole below the cut-off frequency and is flat above it. This leads to the main difference between derivative and low-cut filters: the derivative filter enhances the small-scale structures more, while the low-cut filter makes medium-scale structure much clearer.

In Figure 2, a long line structure in the middle of the sea is very clear in the plot created by $\frac{d}{dx}$ operator; in the plot created by helix derivative filter, it is too weak to be seen. The helix low-cut filter preserves this structure quite well, as shown in Figure 7.

- Computational cost

Another difference is the cost of image processing:

1. The helix low-cut filter needs to do the deconvolution with \mathbf{D} besides the convolution with \mathbf{H} ;
2. The worse spectral symmetry leads helix low-cut filter to the use of a long filter.

These two points add the burden of computation for the helix low-cut filter.

So if the main interests are the small-scale structures or the high costs of computation are not affordable, I should choose the helix derivative filter. Otherwise, the helix low-cut filter is a better choice.

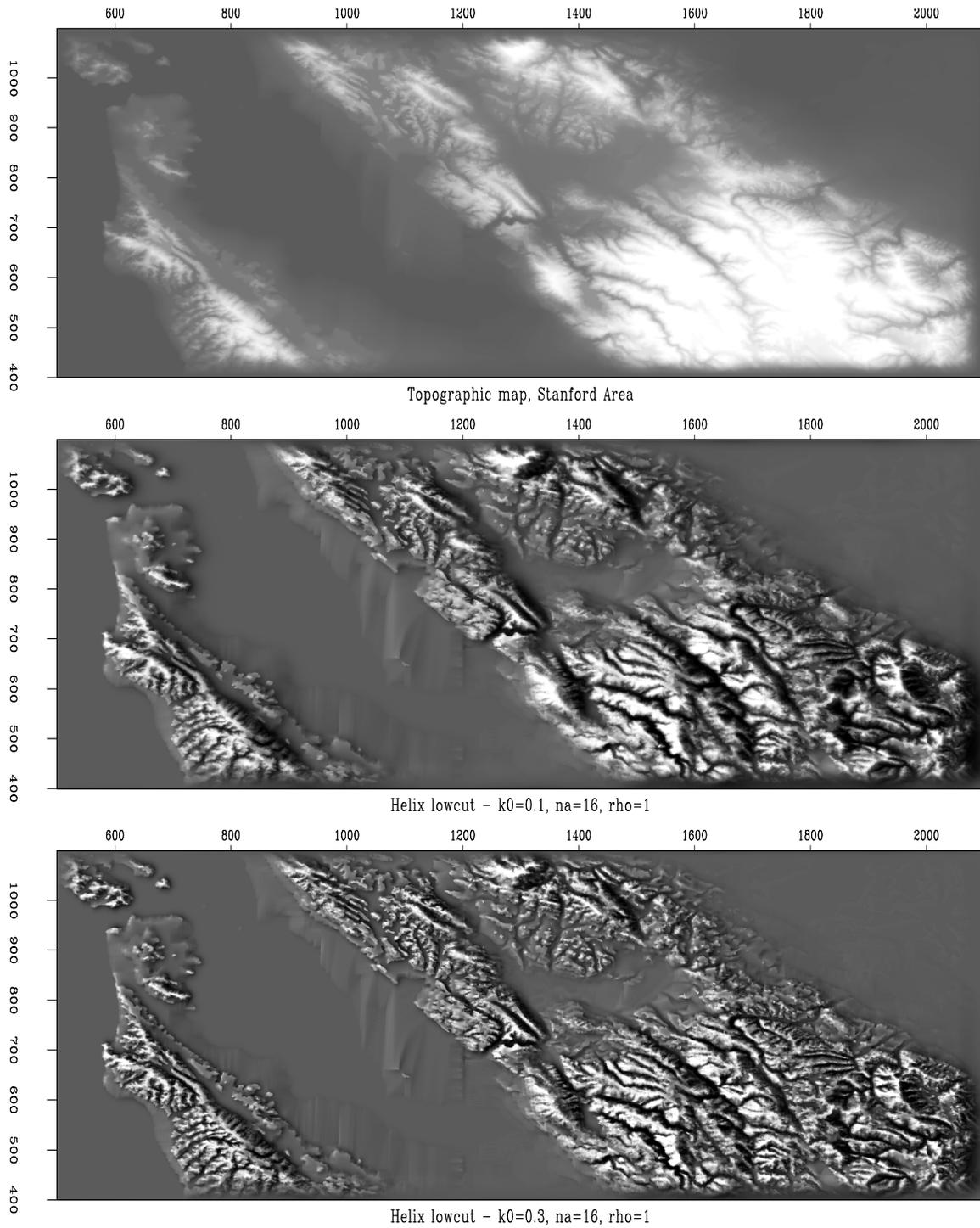


Figure 5: Bay Area maps. The top is the topographic map; the other two are roughened with helix low-cut filter. The middle is $k_0 = 0.1$; and the bottom is $k_0 = 0.3$. $n_a = 16$, $\rho = 1$.

`yi1-bay-lct-res` [ER]

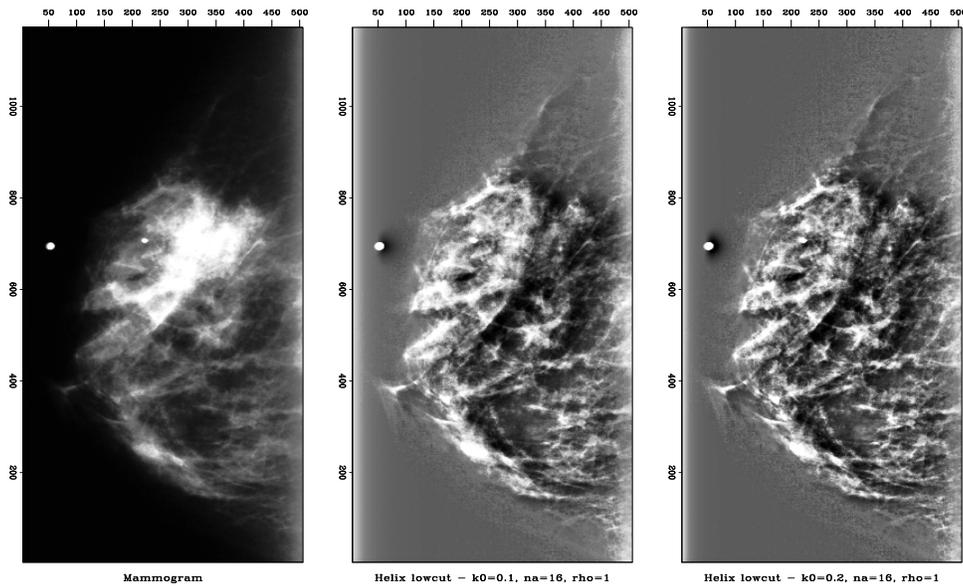


Figure 6: Mammogram (medical X-ray). The left figure is the origin map; the right two are filtered with helix low-cut filter. $n_a = 16$, $\rho = 1$, the middle is $k_0 = 0.1$, the right is $k_0 = 0.2$. `y11-mam-lct-res` [ER]

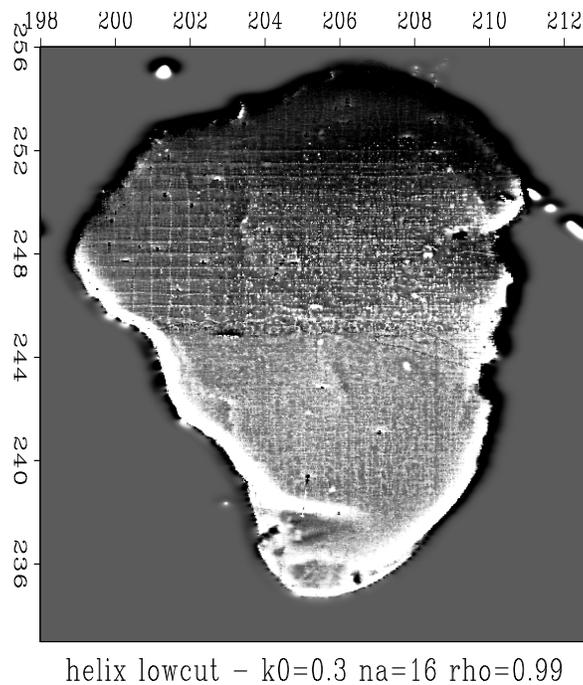


Figure 7: View of the Sea of Galilee roughened with a helix low-cut filter. $k_0 = 0.3$, $n_a = 16$ and $\rho = 0.99$. The edge of the sea fades away, but in the middle of the sea, the long line structure is persevered quite well. `y11-gal-lct-res` [ER]

CONCLUSIONS

The enhanced derivative and low-cut filters with adjustable parameters improve our ability to roughen an image for details. The introduction of adjustable parameter ρ makes it possible to gain computational savings by using a short filter instead of a long one.

For the enhanced helix derivative filter, the filter length determines its spectral symmetry and controls the zero-frequency response with ρ . For the low-cut filter, the cut-off frequency is mainly determined by k_0 , and the zero-frequency response is affected by all three parameters. n_a is an unimportant parameter for helix filters.

The enhanced helix derivative filter emphasizes the small-scale structures in the image, while the low-cut filter emphasizes the medium-scale structures and costs more than derivative filter.

Based on our analysis of the effects of the parameters on the roughened image and filter spectrum, I present guidelines for choosing the adjustable parameters for image processing application.

ACKNOWLEDGMENTS

I thank Jon Claerbout for giving me the motivation and many useful suggestions for this paper.

REFERENCES

- Claerbout, J., 1976, Fundamentals of Geophysical Data Processing: <http://sepwww.stanford.edu/sep/prof/>.
- Claerbout, J., 1997, Multidimensional recursive filters via a helix: SEP-95, 1-13.
- Claerbout, J., 1998a, Geophysical Estimation by Example: <http://sepwww.stanford.edu/sep/prof/>.
- Claerbout, J., 1998b, Multidimensional recursive filters via a helix: SEP-97, 319-335.
- Sava, P., Rickett, J., Fomel, S., and Claerbout, J., 1998, Wilson-Burg spectral factorization with application to helix filtering: SEP-97, 343-351.

APPENDIX A

QUANTITATIVE EFFECT OF THE ADJUSTABLE PARAMETERS

Helix derivative filter

For the helix derivative filter, the response is nearly a linear function of $|\mathbf{k}| = \sqrt{k_x^2 + k_y^2}$ and does not have the “cut-off frequency”. The main feature here is the zero-frequency response. Figure A-1 shows the zero-frequency response R_0 as the function of filter length when $\rho = 0$. n_a is the half size of the filter. I find that R_0 decreases as n_a increases and $R_0 \propto n_a^{-1}$. The empirical relationship between R_0 and n_a is

$$R_0 \approx \frac{0.44}{n_a} \quad (\text{A-1})$$

R_0 is the sum of the helix filter’s coefficients, so when $\rho \neq 0$, according to Equation (3), the zero-frequency response is

$$R_0 \approx \frac{0.44}{n_a}(1 - \rho) \quad (\text{A-2})$$

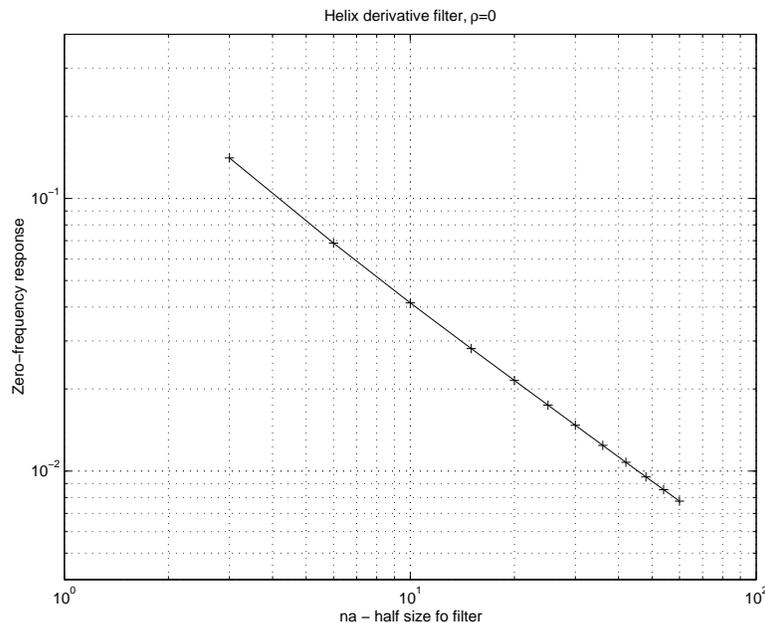


Figure A-1: The zero-frequency response of helix derivative filter when $\rho = 0$. The approximate zero-frequency response is $\frac{0.44}{n_a}$. [yi1-drv-r0f-ar0] [CR]

The zero-frequency response of the enhanced helix derivative filter is controlled by both n_a and ρ , I can compute the approximate value of R_0 using Equation (A-2).

Helix low-cut filter

For the helix low-cut filter, the main features are the cut-off frequency and zero-frequency response. All the frequencies in this paper are scaled values, taking π as the Nyquist frequency.

- Effect of k_0

Figure A-2 is the zero-frequency response and cut-off frequency of a 100-point low-cut filter for different k_0 when $\rho = 0$. I find that cut-off frequency f_0 is almost the same as k_0 .

$$f_0 \approx k_0 \quad (\text{A-3})$$

This satisfies the Equation (2) very well. In Equation (2), if I set $k = k_0$, the expression of $\frac{HH}{DD}$ is 0.5. According to the definition of cut-off frequency, the energy spectrum of the filter should be 0.5 at the cut-off frequency.

The difference of the cut-off frequencies at various azimuths is very small when $k_0 > 0.1$ and can be ignored. For $k_0 < 0.1$, the difference is obvious.

In Figure A-2, the zero-frequency response decreases as k_0 increases. For $\rho = 0$ and $k_0 > 0.03$, $R_0 \propto k_0^{-1}$, the empirical relationship between the zero-frequency response of 100-point and k_0 is

$$R_0 \approx \frac{0.017}{k_0} \quad (\text{A-4})$$

For $k_0 < 0.02$, the zero frequency response curve becomes flat, and reaches the limit of 1. This is easily derived from Equation (2). When k_0 turns to zero, the difference between numerator H and denominator and D becomes smaller.

- Effect of n_a

Figure A-3 shows the effect of n_a on helix low-cut filter when $k_0 = 0.3$ and $\rho = 0$. For small n_a , the numerical anisotropy is very strong. Although the mean value of the cut-off frequency remains the same, the azimuthal difference becomes larger when n_a becomes smaller. The zero-frequency response increases as n_a decreases, and when $k_0 = 0.3$ and $\rho = 0$, the empirical expression is

$$R_0 \approx \frac{2.7}{n_a} \quad (\text{A-5})$$

- Effect of ρ

ρ directly controls the zero-frequency response and affects the cut-off frequency as well. Figure A-4 shows the cut-off frequency of the 100-point helix low-cut filter as the function of k_0 when $\rho = 0.5$ and $\rho = 1$. For larger ρ , the numerical anisotropy is stronger, especially when k_0 is small. Compared with Figure A-2, the cut-off frequency at small k_0 increases slightly with ρ .

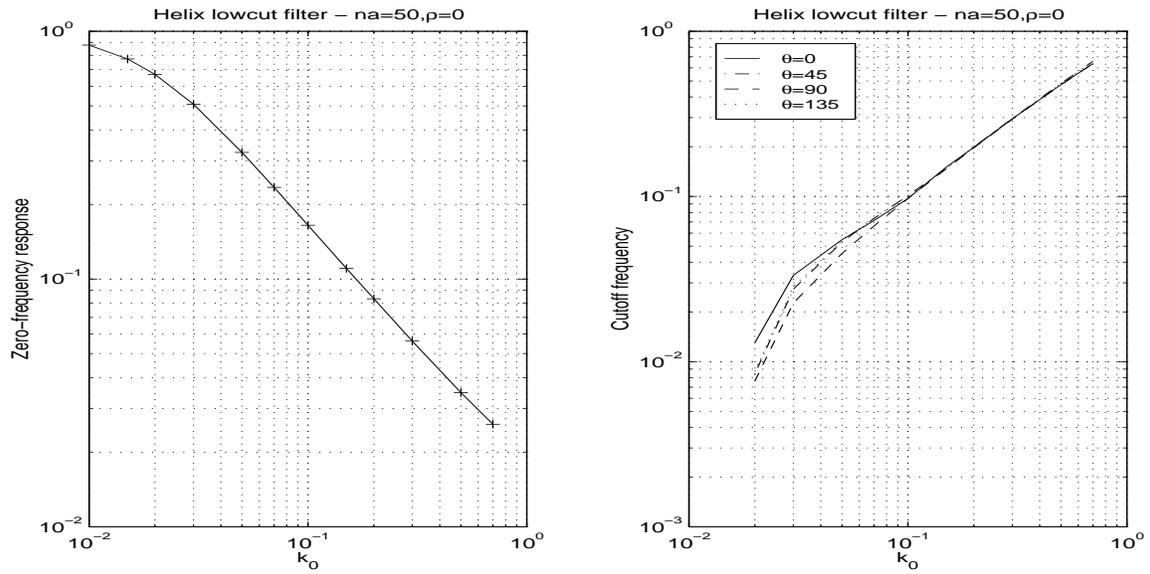


Figure A-2: The zero-frequency response and cut-off frequency of the helix low-cut filter with $n_a = 50$, $\rho = 0$. The zero-frequency response here is about $\frac{0.017}{k_0}$; the cut-off frequency is about k_0 . yi1-lcut-r00-ka50r0 [CR]

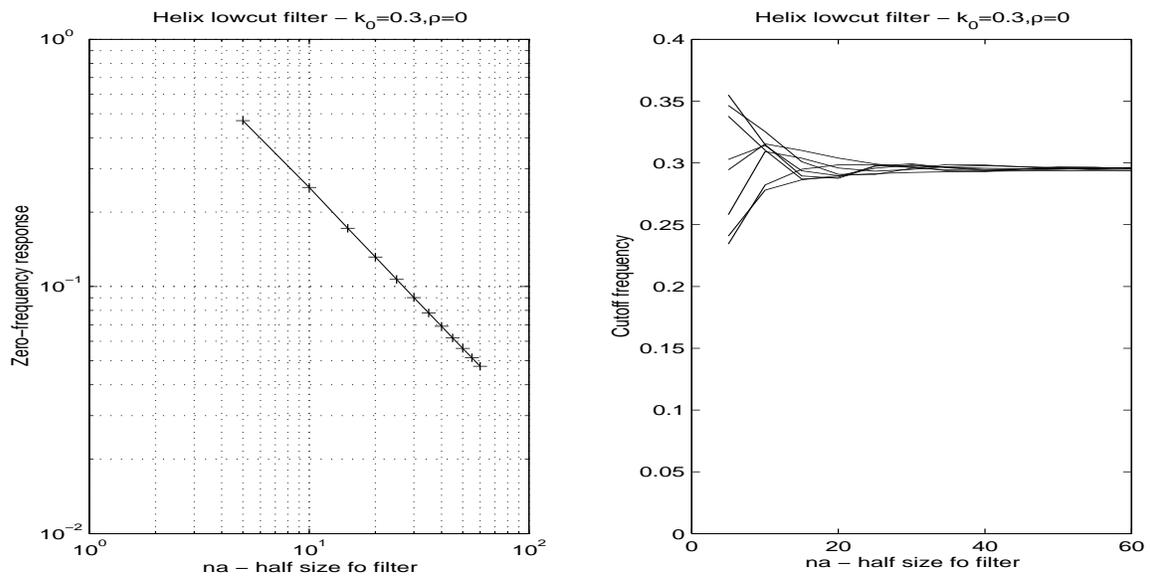


Figure A-3: The zero-frequency response and cut-off frequency of the helix low-cut filter with $k_0 = 0.3$, $\rho = 0$. The zero-frequency response here is about $\frac{2.7}{n_a}$. yi1-lcut-r00-ak3r0 [CR]

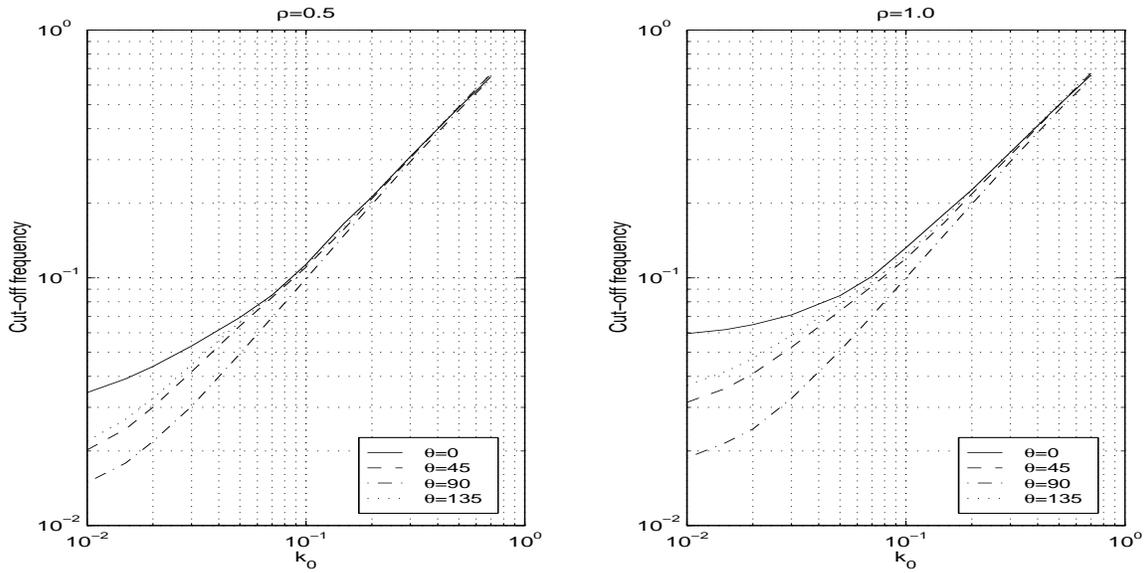


Figure A-4: The zero-frequency response and cut-off frequency of helix low-cut filter. `yi1-lcut-rf0-kra50` [CR]

- Composed effects

Based on the proceeding analysis, I can derive the composed effects of the adjustable parameters on the helix low-cut filter.

The cut-off frequency is mainly governed by k_0 .

$$f_0 \approx k_0 \quad (\text{A-6})$$

For large n_a , f_0 is almost the same as k_0 . Both nonzero ρ and small n_a leads to the anisotropy of cut-off frequency. However, there is a difference between them: ρ causes the average cut-off frequency to increase slightly; small n_a intends to keep it.

The zero-frequency response R_0 is under the direct control of ρ and influenced by n_a and k_0 . If I assume that the influences of n_a and k_0 are independent, the empirical expression of R_0 would be

$$R_0 \approx \frac{0.82}{n_a k_0} (1 - \rho) \quad (\text{A-7})$$

Equations (A-2), (A-6) and (A-7) describe the quantitative effects of the helix derivative / low-cut filter's adjustable parameters. These empirical formulas make it quantitative for us to choose the adjustable parameters of the helix filter in practice.

Helical meshes on spheres and cones

Jon Claerbout¹

ABSTRACT

We embed a helix in the two-dimensional surface of a sphere; likewise, in the two-dimensional surface of a cone. This provides a one-dimensional coordinate system on a two-dimensional surface. Although mesh points are exactly evenly spaced along the helix and approximately evenly spaced in the crossline dimension, unfortunately, the angles between neighboring points are continuously changing. We seem to lose the concepts of two-dimensional autoregression that we have in cartesian space.

INTRODUCTION

Geophysical practice is filled with spheres and cones. Wave fronts are spheres. The Kirchhoff impulse response has a conical asymptote. Surface wave noises fill a conical surface. Likewise linear moveout in 3-D can be taken to be a conical surface. Because of the many applications that I found for a helix mapping of a cartesian space, I chose to examine spheres and cones. (To a small boy with a new hammer, everything looks like a nail.)

HELICAL COORDINATE ON A SPHERE

I set out to find the equations describing a rope that begins from the north pole and spirals its way around a sphere neatly covering it and ending at the south pole. Taking uniform samples along this rope gives a fairly uniform covering of the sphere. The surface of a sphere is two dimensional, but the rope gives a one dimensional covering that is uniformly sampled in that dimension. I wondered about the sampling in the other dimension so I set out to plot it. I found the “generalized spiral set” of Saff and Kuijlaars(1997). In spherical coordinates (θ, ϕ) , for $0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$, they set

$$\theta_k = \arccos(h_k), \quad h_k = -1 + \frac{2(k-1)}{(N-1)}, \quad 1 \leq k \leq N \quad (1)$$

$$\phi_k = \phi_{k-1} + \frac{3.6}{\sqrt{N}} \frac{1}{\sqrt{1-h_k^2}}, \quad 2 \leq k \leq N-1, \quad \phi_1 = \phi_N = 0 \quad (2)$$

My plot of these equations is shown in Figure 1. There is no interesting pattern in the crossline direction. Although my plot looks reasonable, Saff and Kuijlaars(1997) show a curious pattern

¹email: claerbout@stanford.edu

in the crossline direction that my plots do not show. A few tests with various values of N and various rotations failed to show any curious pattern.

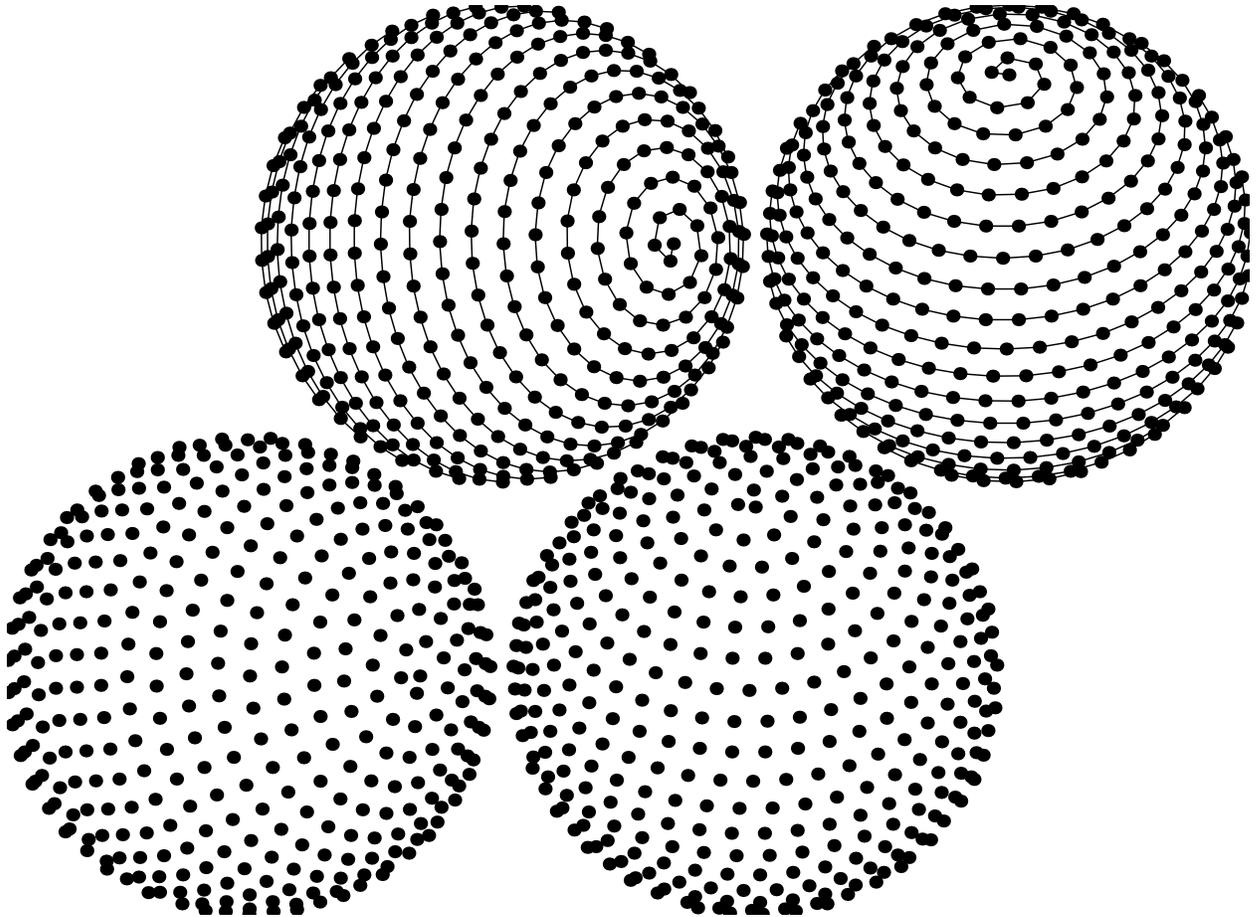


Figure 1: Helix on a sphere. Top shows the embedded helix. Bottom hides it. An interesting pattern of points that appears in the article in the Mathematical Intelligencer is inexplicably absent here (even though I tested several rotations and several values of N). jon2-sphere
[ER]

HELICAL COORDINATE ON A CONE

I could not find the equations for a helix on a cone, so I derive them below. An example of results is Figure 2.

Define ϕ to be the angle from the axis of the cone to its surface. I call this the apex angle. I discovered by accident that certain apex angles give interesting patterns in the crossline directions while most do not. I found these patterns to be insensitive to the choice of N and show them here for about $N = 1400$. Another value of apex angle with an interesting pattern is $\phi = 1/3$. It gives the charming pattern in Figure 3. (I've seen this pattern before on a party

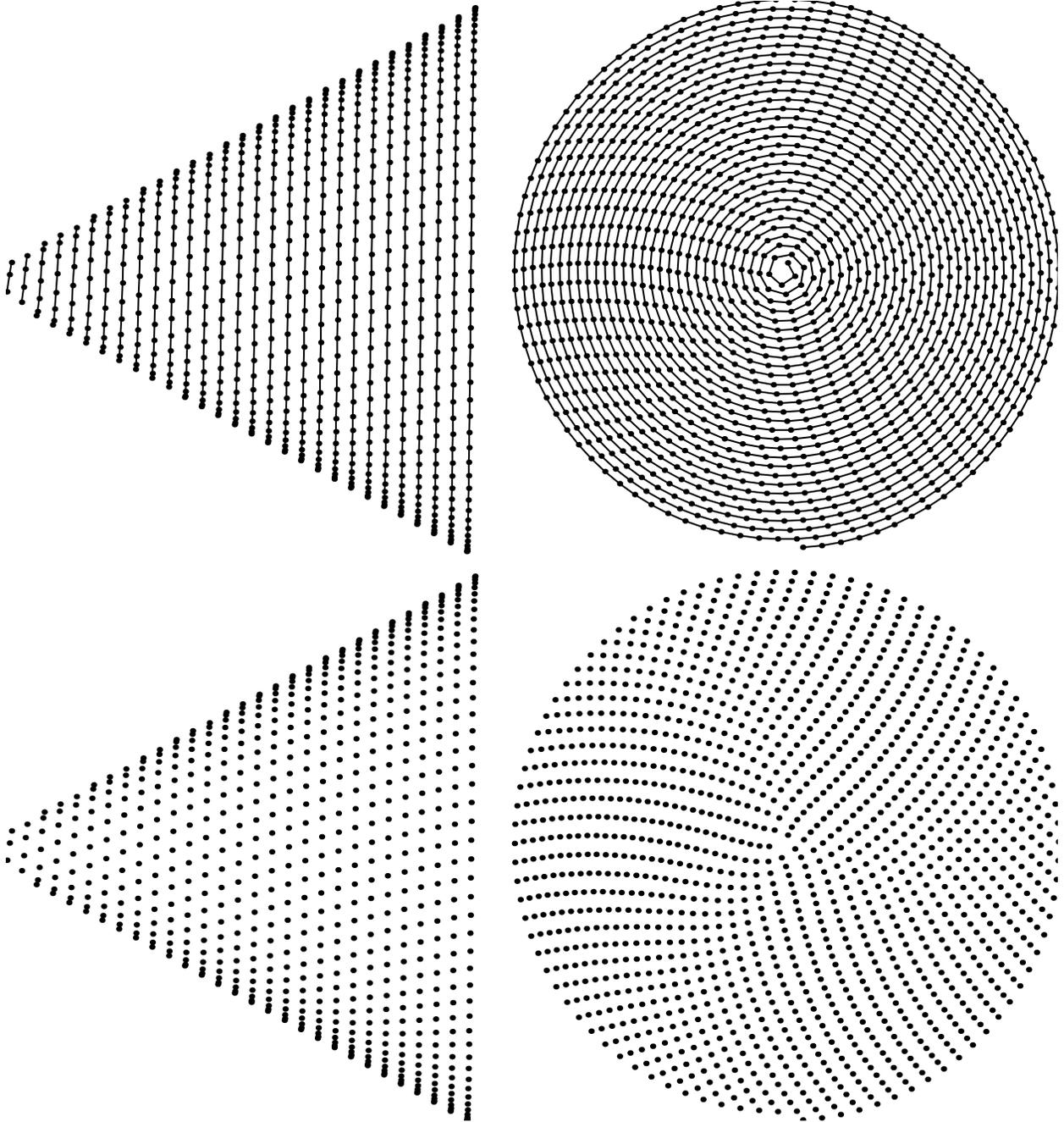


Figure 2: Helix on a cone. Top shows the embedded helix. Bottom hides it. Apex angle $\phi = 1/2$. [jon2-trycart] [ER]

hat. I attributed it to an ingenious artist. Now I realize that like all mathematics, this art existed before the big bang.)

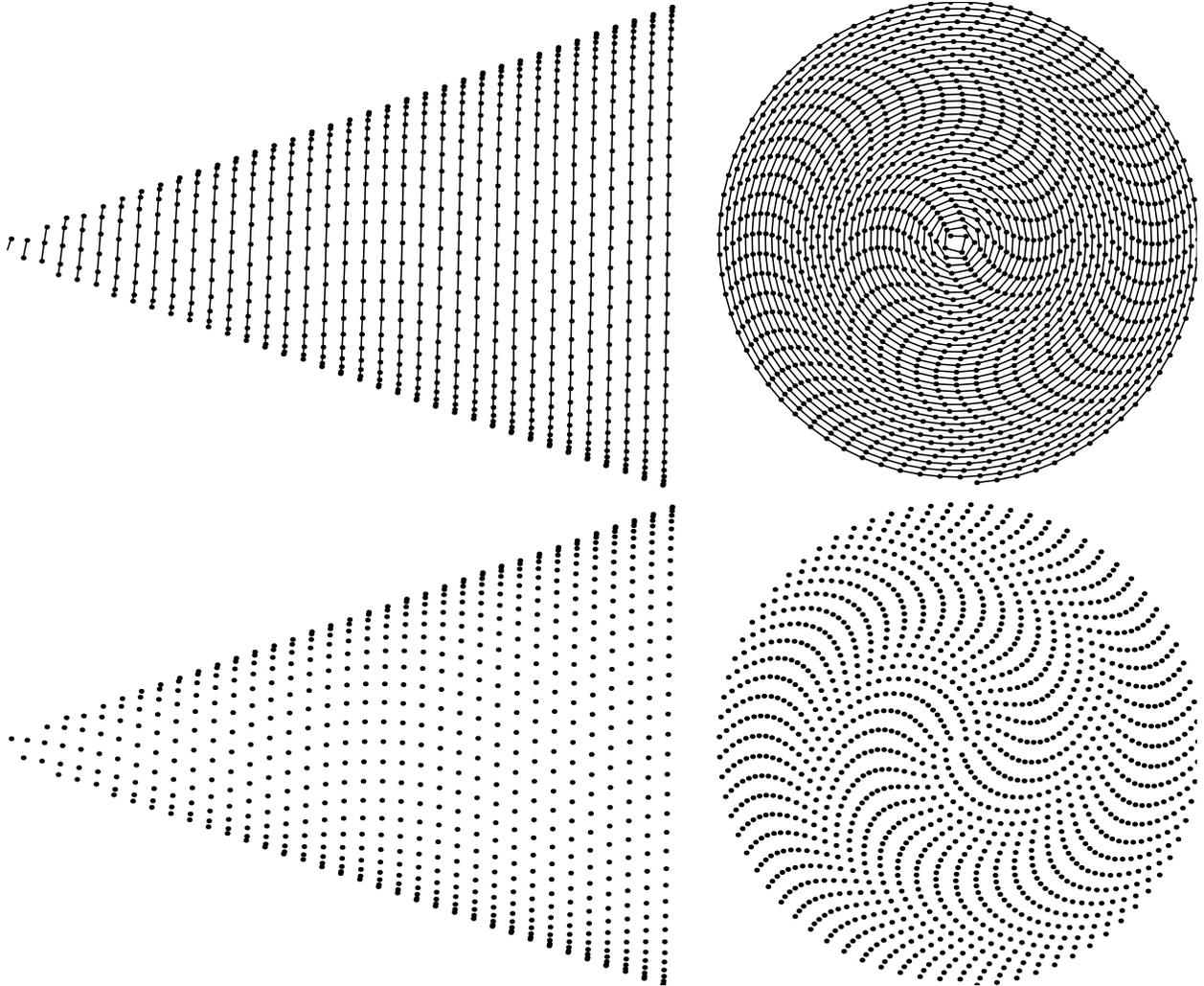


Figure 3: Helix on a cone. Top shows the embedded helix. Bottom hides it. Apex angle equals one third radian. [jon2-YinYang](#) [ER]

The radius r of the cone divided by its altitude z has a ratio $r/z = \tan \phi$ where ϕ is the angle from the axis of the cone to its surface. The area of a circle is πR^2 . The surface area of a cone is $\pi R^2 / \sin \phi$. Dividing the area into N cells, the surface area per cell is $\pi R^2 / (N \sin \phi)$. Taking any cell area to be square, the length of a side is

$$\Delta s = R \sqrt{\frac{\pi}{N \sin \phi}} \quad (3)$$

The number of points running one cycle around a rim is

$$N_{\text{rim}} = \frac{2\pi r}{\Delta s} = \frac{2r}{R} \sqrt{N\pi \sin \phi} \quad (4)$$

On one cycle around the rim, the radius must change by $\Delta r = \Delta h \sin \phi$ where the hypotenuse Δh of the triangle lies on the surface of the cone. Thus, for each mesh point going around the rim

$$\Delta r = \frac{\Delta h \sin \phi}{N_{\text{rim}}} \quad (5)$$

Since we want “square” mesh points, Δh must equal Δs , hence

$$\Delta r = \frac{R^2}{2rN} \quad (6)$$

The algorithm starts on the rim at $\theta = 0$ and $r = R$. At each step, update r and θ with $-\Delta r$ and $\Delta \theta = r \Delta s$. Stop before $r - \Delta r$ becomes negative.

FILLING VOLUMES

Nested spherical shells can fill a volume. A question is whether the “string” that winds around each shell should wind from north to south on every layer suffering a discontinuity at the poles, or whether layers of string should alternate between winding south and winding north.

Likewise cones can nest with cones, as a stack of ice-cream cones.

CONCLUSIONS

The useful property of all these fillings of space is that the spacing between the grid points is basically constant. Two-dimensional “pixels” have about equal area, likewise three-dimensional voxels have about equal volume.

Although mesh points are exactly evenly spaced along the helix and approximately evenly spaced in the crossline dimension, unfortunately, the angles between neighboring points are continuously changing.

I hope I am wrong, but this appears to defeat the many useful tricks we play in cartesian space such as using autoregression for multidimensional prediction and using prediction-error filters to characterize multidimensional spectra. Thus we cannot interpolate by expressing physical stationarity as stationarity of filter coefficients.

REFERENCES

Saff, E., and Kuijlaars, A., 1997, Distributing many points on a sphere: *Mathematical Intelligencer*, **19**, no. 1, 5–11.

3-D traveltimes computation by Huygens wavefront tracing

Paul Sava¹

ABSTRACT

In this paper, I present a 3-D implementation of Huygens wavefront tracing. The three-dimensional version of the method retains the characteristics of the two-dimensional one: stability, accuracy, and efficiency. The major difficulty of the 3-D extension is related to the handling of triplications. An easy to implement solution is to approximate the wavefronts at the triplications as planes orthogonal to the incident ray.

INTRODUCTION

The goal of obtaining efficient and robust multiple-arrival traveltimes has yet to be accomplished in practice. The typical traveltimes methods—eikonal solvers, ray tracing and wavefront construction—either only compute the first-arrival traveltimes (eikonal solvers) or are expensive and not always robust in regions of high velocity variation (ray tracing and wavefront construction). However, Huygens wavefront tracing (HWT), introduced in a previous paper (Sava and Fomel, 1998), does offer a robust and very efficient method of computing multiple-arrival traveltimes.

This paper presents my extension of Huygens wavefront tracing to three dimensions. I start by briefly reviewing the theory, then I discuss how to handle the triplications of the wavefronts and present one 2-D and two 3-D examples. I continue with a brief comparison to other traveltimes methods, and end with conclusions and possible directions for future work.

REVIEW OF HWT THEORY

Given an isotropic heterogeneous medium, wavefronts are represented by surfaces of equal traveltimes, constrained by the eikonal equation

$$\left(\frac{\partial\tau}{\partial x}\right)^2 + \left(\frac{\partial\tau}{\partial y}\right)^2 + \left(\frac{\partial\tau}{\partial z}\right)^2 = \frac{1}{v^2(x,y,z)} \quad (1)$$

and appropriate boundary conditions.

¹email: paul@sep.stanford.edu

Each point on a wavefront can be parametrized by either its Cartesian coordinates x , y , and z , or its ray coordinates, which consist of the travelttime τ , and the two shooting angles at the source, γ and ϕ .

For complex velocity fields, the ray coordinates as a function of the Cartesian coordinates become multi-valued, in other words, there is more than one ray going through a given point in the subsurface. In contrast, the Cartesian coordinates as a function of the ray coordinates remain single-valued, that is, there is one unique position in the subsurface where a ray, shot with two particular shooting angles, arrives at a given time. Figure 1 illustrates the difference between the two representations of the wavefronts.

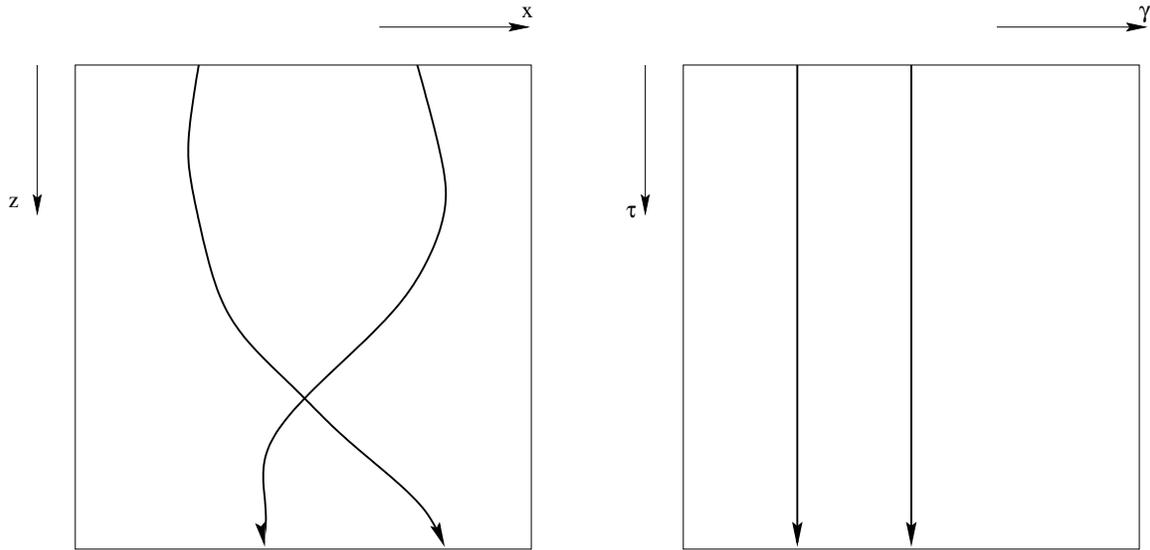


Figure 1: The ray coordinates as a function of the Cartesian coordinates are multi-valued (left). The Cartesian coordinates as a function of the ray coordinates are single-valued (right). paul2-coord [NR]

Since $x(\tau, \gamma, \phi)$, $y(\tau, \gamma, \phi)$, and $z(\tau, \gamma, \phi)$ are uniquely defined for arbitrarily complex velocity fields, the eikonal equation (Equation 1) can be transformed to another form that is better suited for analysis in ray coordinates (Sava and Fomel, 1998):

$$\left(\frac{\partial x}{\partial \tau}\right)^2 + \left(\frac{\partial y}{\partial \tau}\right)^2 + \left(\frac{\partial z}{\partial \tau}\right)^2 = v^2(x, y, z). \quad (2)$$

Converting equation (2) to a finite-difference equation using a first-order discretization scheme, we obtain

$$\left(x_{j+1}^{i,k} - x_j^{i,k}\right)^2 + \left(y_{j+1}^{i,k} - y_j^{i,k}\right)^2 + \left(z_{j+1}^{i,k} - z_j^{i,k}\right)^2 = \left(r_j^{i,k}\right)^2, \quad (3)$$

where j is the index of the current wavefront, $j + 1$ is the index of the new wavefront to be computed, and i and k are the indices of the shooting angles. This equation represents a

sphere, the wavefront of a secondary Huygens source placed at $(x_j^{i,k}, y_j^{i,k}, z_j^{i,k})$ on the current wavefront.

According to the Huygens principle, the new wavefront is the envelope of all the secondary wavefronts. Mathematically, the position of the new wavefront is described by a system of three equations composed of Equation (3) and the following two equations (Sava and Fomel, 1998):

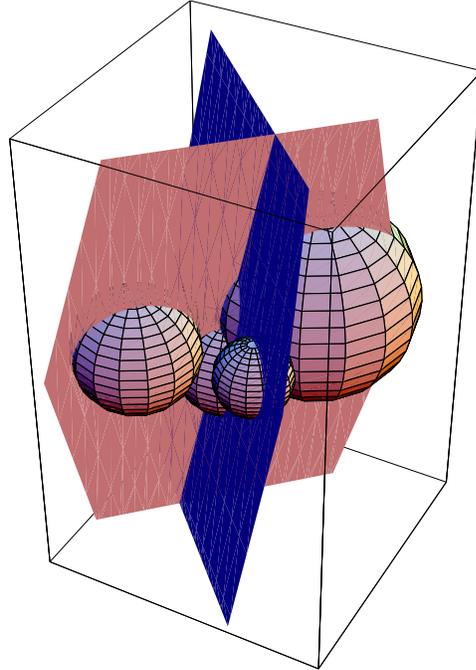
$$\begin{aligned} (x_j^{i,k} - x_{j+1}^{i,k}) (x_j^{i+1,k} - x_j^{i-1,k}) + (y_j^{i,k} - y_{j+1}^{i,k}) (y_j^{i+1,k} - y_j^{i-1,k}) + \\ (z_j^{i,k} - z_{j+1}^{i,k}) (z_j^{i+1,k} - z_j^{i-1,k}) = r_j^{i,k} (r_j^{i+1,k} - r_j^{i-1,k}) \end{aligned} \quad (4)$$

and

$$\begin{aligned} (x_j^{i,k} - x_{j+1}^{i,k}) (x_j^{i,k+1} - x_j^{i,k-1}) + (y_j^{i,k} - y_{j+1}^{i,k}) (y_j^{i,k+1} - y_j^{i,k-1}) + \\ (z_j^{i,k} - z_{j+1}^{i,k}) (z_j^{i,k+1} - z_j^{i,k-1}) = r_j^{i,k} (r_j^{i,k+1} - r_j^{i,k-1}) . \end{aligned} \quad (5)$$

Figure 2 contains a simple geometrical interpretation of the system described by Equations (3), (4), and (5). Five points on the current wavefront, represented by the five spheres, not all

Figure 2: A geometrical updating scheme for 3-D HWT in the physical domain. Five points on the current wavefront, represented by the five spheres, not all visible, with radii defined by the velocities at the corresponding points of the wavefront, are used to compute a point on the next wavefront. The sphere in the middle represents equation (3), and the planes represent equations (4) and (5). `paul2-huygens3d` [CR]



visible, with radii defined by the velocities at the corresponding points of the wavefront, are used to compute a point on the next wavefront. The sphere in the middle represents equation (3), while the planes represent equations (4) and (5).

Huygens wavefront tracing, based on the system of equations (3), (4), and (5), is nothing but an explicit finite-difference method in the ray coordinate system. The coordinates of the new wavefronts are computed according to those of the current wavefront. A three-point stencil is needed in two dimensions to compute the centered finite-difference representation

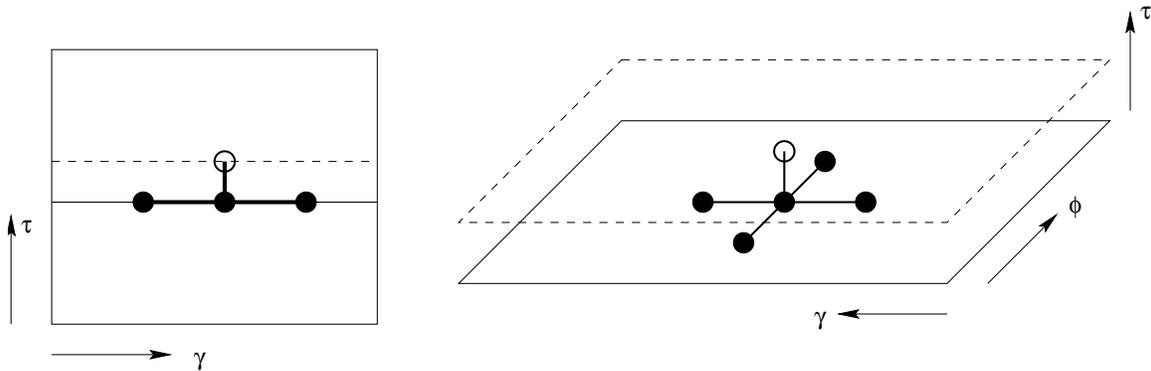


Figure 3: Finite-difference travelttime computation scheme. A 3-point stencil is needed in 2-D to compute the centered finite-difference representation of the derivative with respect to the shooting angle (left). A 5-point stencil is needed in 3-D to compute the centered finite-difference representation of the derivatives with respect to the shooting angles (right)

`paul2-scheme` [NR]

of the derivative with respect to the shooting angle, while a five-point stencil is needed in three dimensions to compute the centered finite-difference representation of the derivatives with respect to the shooting angles. Figure 3 is a graphical illustration of the finite-difference stencils.

BOUNDARIES AND TRIPLICATIONS

This section presents a short discussion of the special treatment required by the boundaries of the computation domain. These boundaries are of two kinds: exterior boundaries, represented by the edges of the computational domain, and interior boundaries, represented by the triplication lines. Because of the centered finite-difference scheme, HWT cannot be used at the boundaries of the computational domain. This means that the boundaries need to be treated differently from the rest of the domain. Also, the centered finite-difference scheme cannot be used when the wavefronts create triplications. Triplications represent points of discontinuity of the derivative along the wavefront, and, therefore, the centered finite-difference representation of the derivative is inappropriate. Figure 4 describes a point of triplication represented in both the physical (Cartesian) domain (left) and the ray coordinate domain (right).

One possible solution for the boundaries is to make a local approximation of the wavefront. Instead of considering the actual points on the wavefront, we can create an approximate wavefront that is locally orthogonal to the ray arriving at the cusp point, as depicted in Figure 5. We can then pick an appropriate number of points (two in 2-D or four in 3-D) on this approximate wavefront, and use the HWT scheme without any change. A new search for the cusp points is then needed on the new wavefront before we can proceed any further.

Figure 4: The centered finite-difference representation of the derivative along the wavefront cannot be used at the cusps. These points represent discontinuities in the derivative, and need to be treated separately. `paul2-cusp` [NR]

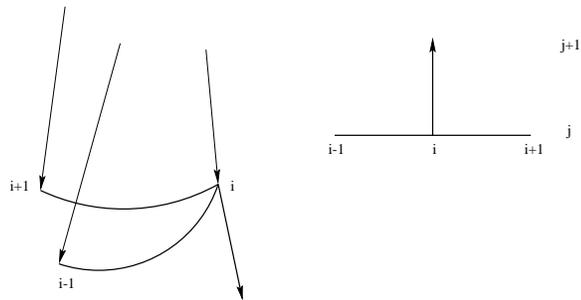
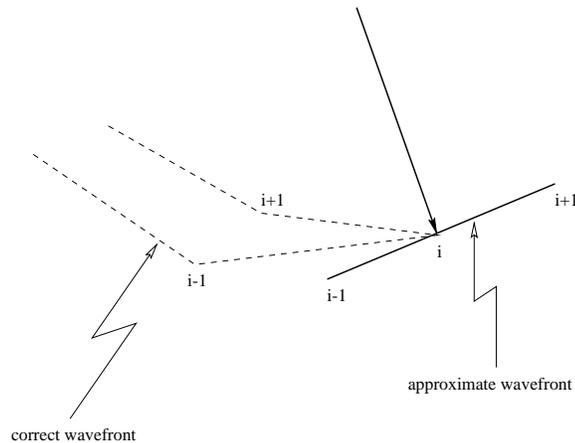


Figure 5: The centered finite-difference representation of the derivative along the wavefront cannot be used at the cusps. Instead, we can use a local approximation of the wavefront as a plane locally orthogonal to the ray arriving at the cusp. `paul2-hrt` [NR]



EXAMPLES

This section presents a number of results of traveltime computation using several models that exemplify the main features of the method.

The first example compares the results of the traveltime computation with those of the full wave-equation modeling for the 2-D salt dome model shown in Figure 6. Figure 7 is a snapshot of the wavefield at 1.23 s, superimposed on an outline of the velocity model. Figure 8 shows, in addition, the wavefront corresponding to the same propagation time (1.23s), and some of the rays derived from the computed wavefronts. The first arrivals of the wavefronts superimpose well on the similar events in the wavefield. The later arrivals also superimpose well on the corresponding events in the wavefield, though the sampling is a lot sparser. This is understandable, since in HWT the wavefronts are sampled evenly in the ray domain, but not in the physical domain. Because sampling in the two domains is related, a better sampling in the ray domain can generate more accurate sampling in the physical domain. However, sampling is dependent on the model; there is no guarantee of more accurate sampling for rays shot with a smaller angle step. A better idea is to dynamically modify the sampling on the wavefronts as is done in some of the wavefront construction methods (Vinje et al., 1993).

The next two examples are three-dimensional. In the first, I consider a strong negative Gaussian velocity anomaly of -2500 m/s placed in a constant velocity medium of 3000 m/s (Figure 9). The source is placed on the surface above the center of the anomaly. Figure 11 depicts the traveltimes obtained by HWT and interpolated on the Cartesian grid using a multi-

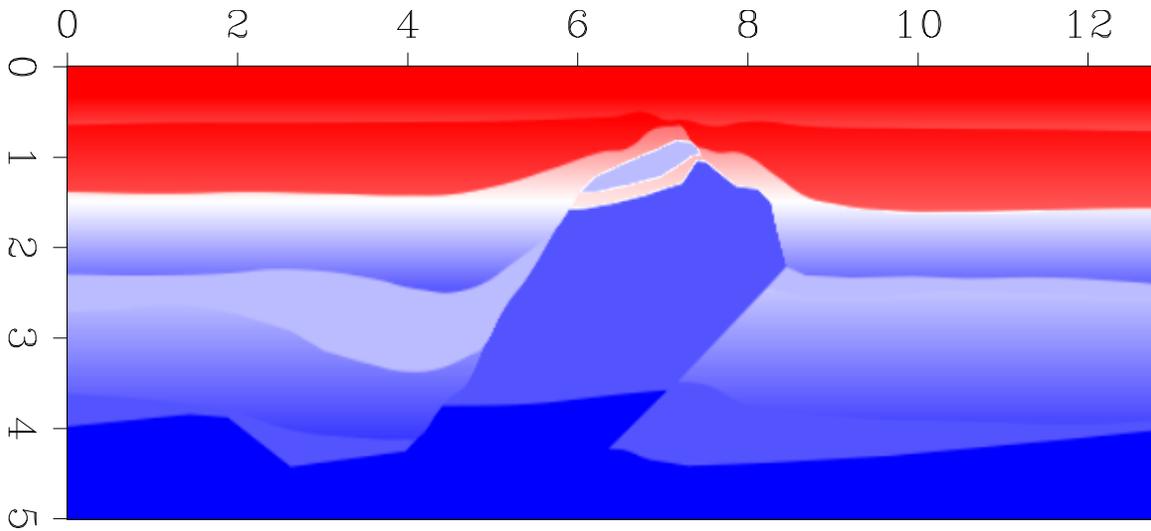


Figure 6: The Elf velocity model. `paul2-velfsa` [ER]

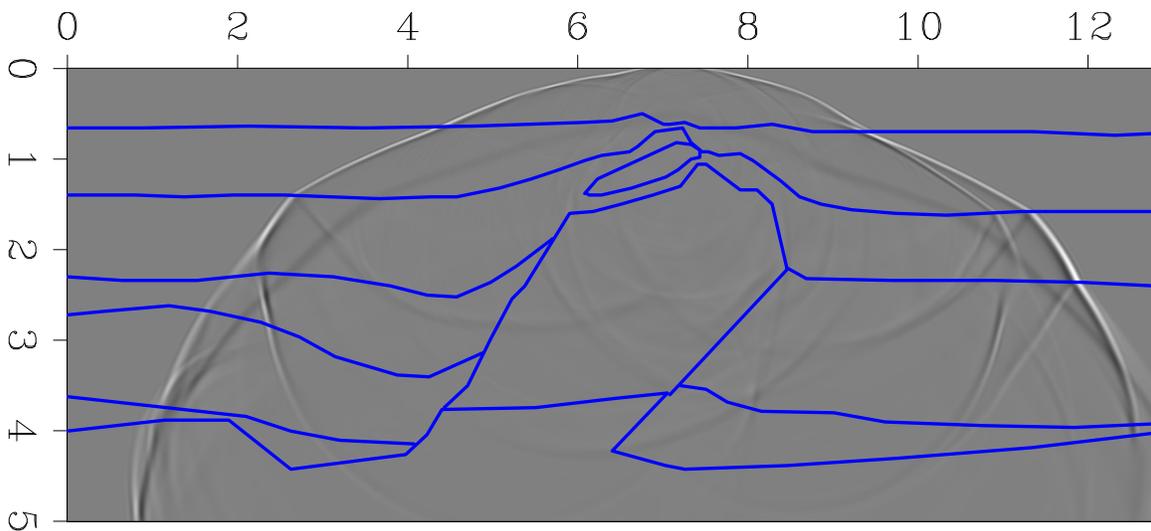


Figure 7: Wave-equation modeling on the Elf velocity model. The snapshot was taken at 1.23s. The outline on the model is superimposed on the wavefield. `paul2-melfsa` [ER]

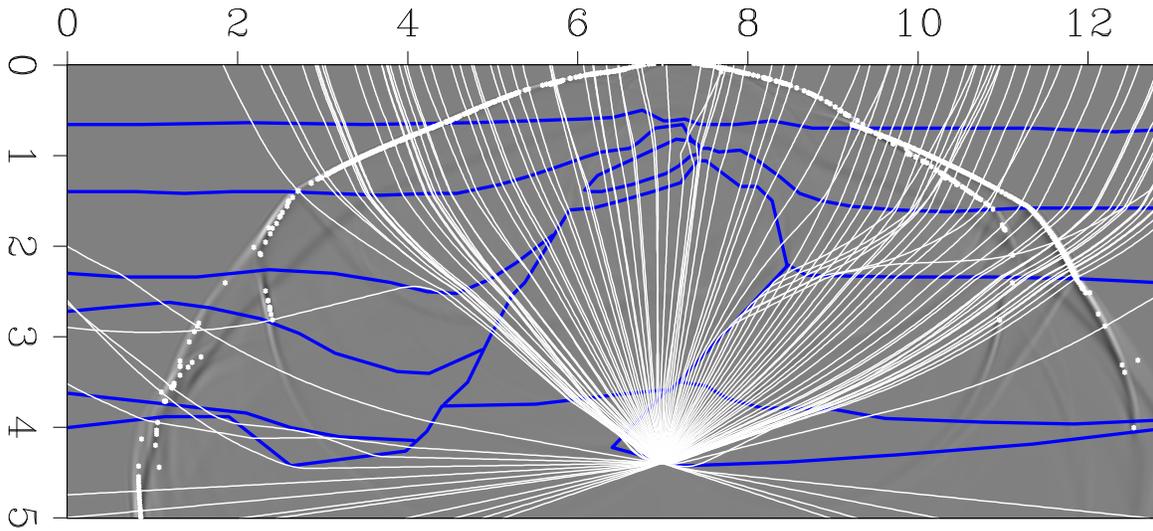


Figure 8: HWT wavefront superimposed on a wave-equation modeling snapshot at 1.23s. The portion of the wavefronts corresponding to the first-arrival matches well the first-arrival of the wavefield. Also, the later HWT arrivals match well with similar events of the wave-equation modeling shown in Figure 7. `paul2-welfsa` [ER]

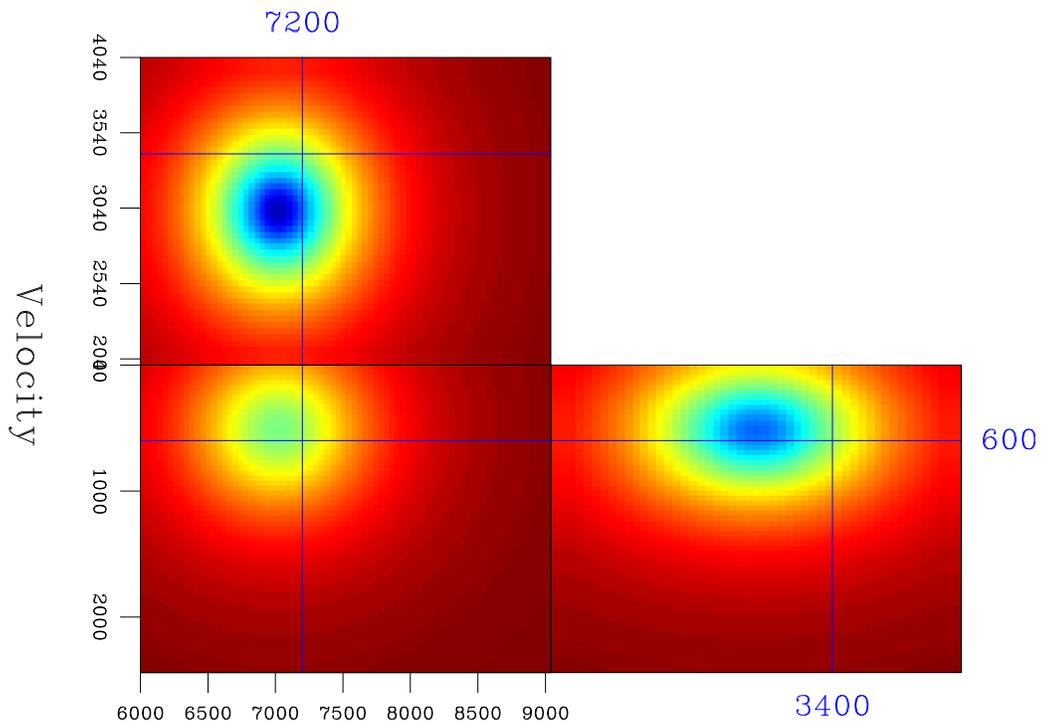


Figure 9: A 3-D Gaussian velocity anomaly. `paul2-velgaus` [ER]

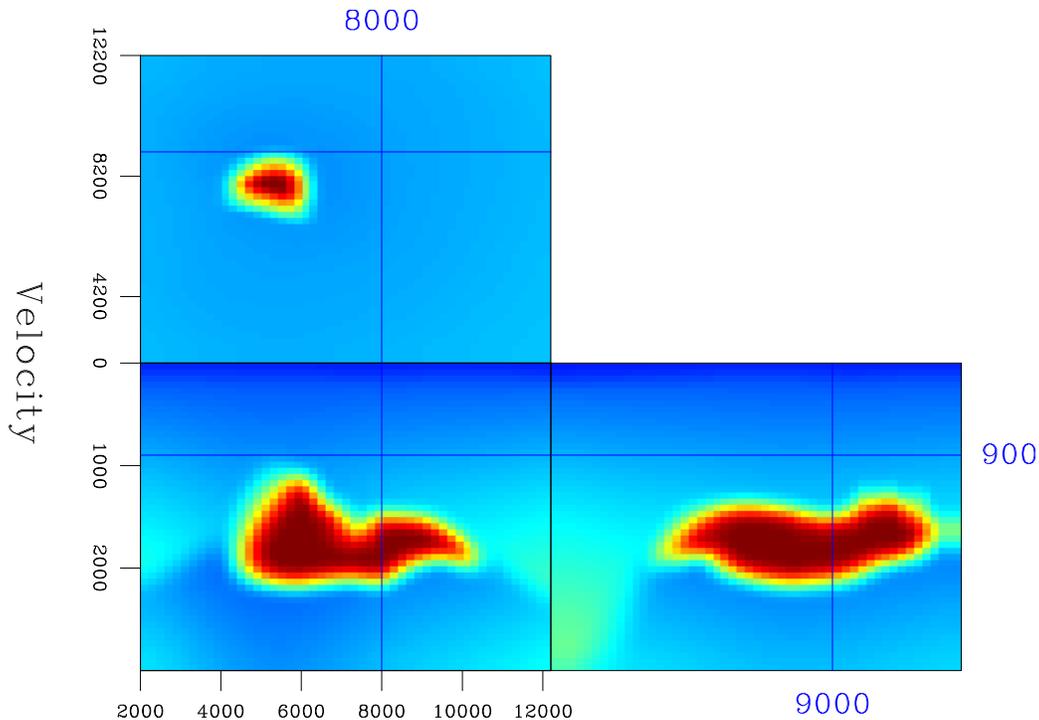


Figure 10: The 3-D SEG-EAGE salt velocity model. `paul2-velsalt` [ER]

valued traveltimes interpolation method (Sava and Biondi, 1997). From the multiple values of the traveltimes field, I have selected those that correspond to the minimum ray-length from the source (Nichols et al., 1998). For comparison, Figure 12 indicates the first-arrival traveltimes of a fast marching eikonal solver (FME) (Popovici and Sethian, 1997; Fomel, 1997). As expected, the traveltimes cubes in Figures 11 and 12 match well for the regions that correspond to the first-arrival wavefronts. However, the traveltimes cubes are completely different in the regions where the wavefronts triplicate, regions where the shortest ray does not correspond to the first arrival, but to a later one.

In the second 3-D example, I consider the SEG-EAGE salt model presented in Figure 10. Again, the HWT multiple-arrival traveltimes are interpolated on the Cartesian grid, and selected to have the minimum ray-length from the source, as we see in Figure 13. For comparison, the fast-marching eikonal results for the same source point and the same velocity model appear in Figure 14. The HWT traveltimes for the volume above the salt body correspond to the direct arrival, while the FME traveltimes correspond to the head waves from the top of the salt.

DISCUSSION

This section briefly compares Huygens wavefront tracing with the other major traveltimes computation methods: paraxial ray tracing (PRT) (Červený, 1987), eikonal solvers (ES) (Vidale,

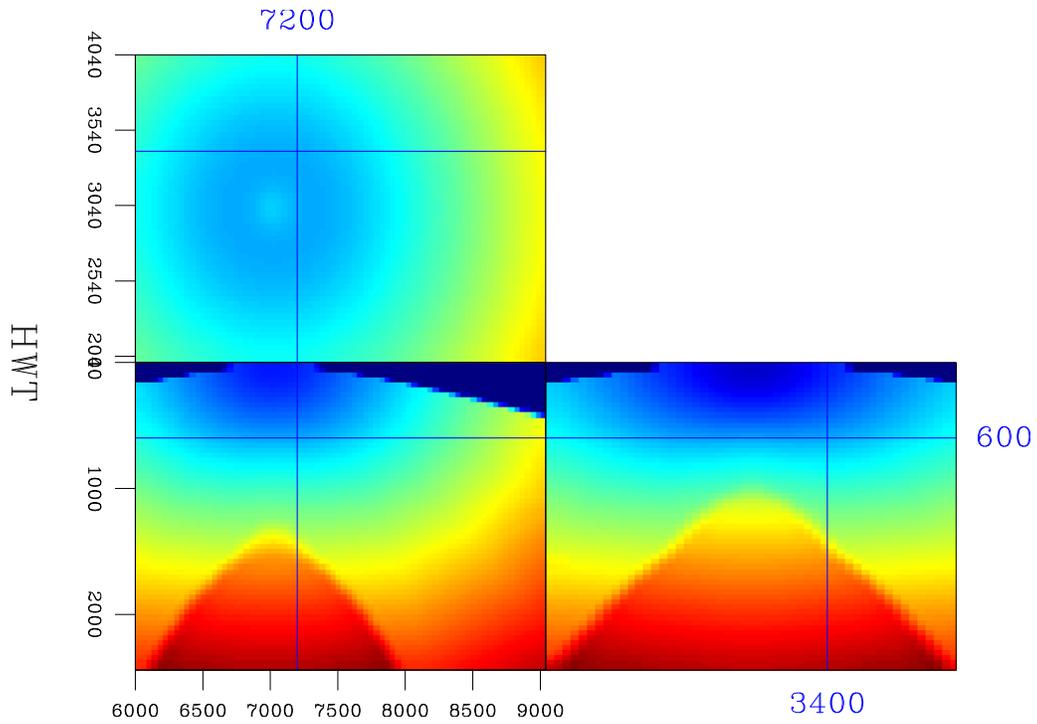


Figure 11: The 3-D Gaussian velocity anomaly. A traveltime cube obtained using HWT. The traveltimes are interpolated on the rectangular grid. The selected traveltimes correspond to the shortest rays (Nichols et al., 1998). `paul2-hwtgaus` [ER]

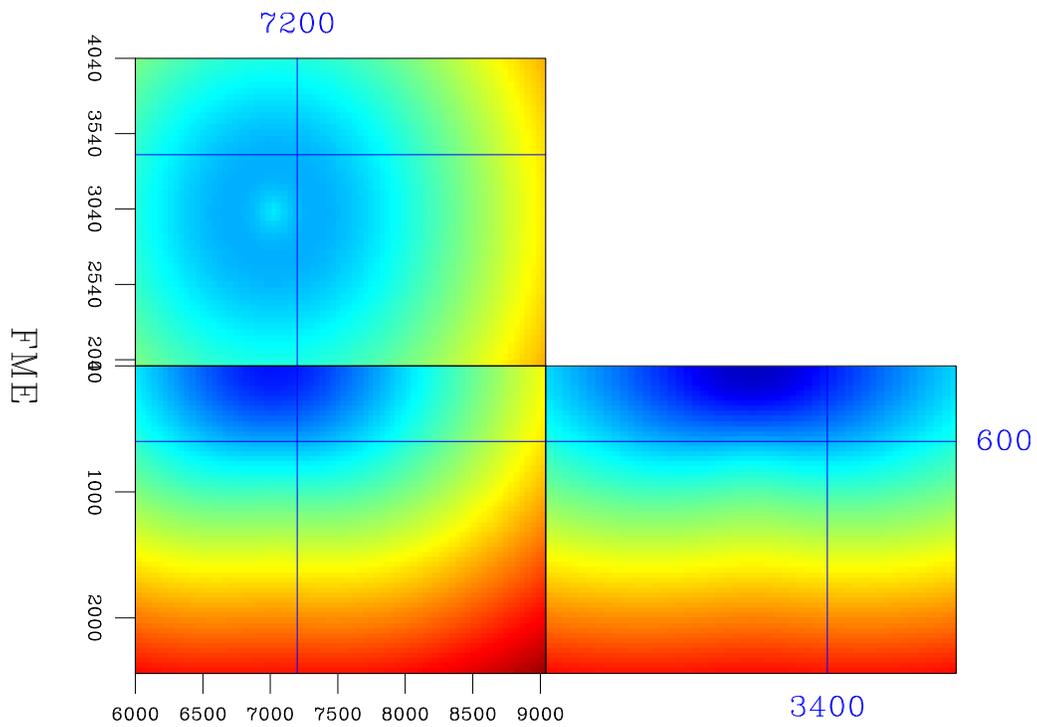


Figure 12: The 3-D Gaussian velocity anomaly. A traveltime cube obtained using a fast-marching first-arrival eikonal solver. `paul2-fmegaus` [ER]

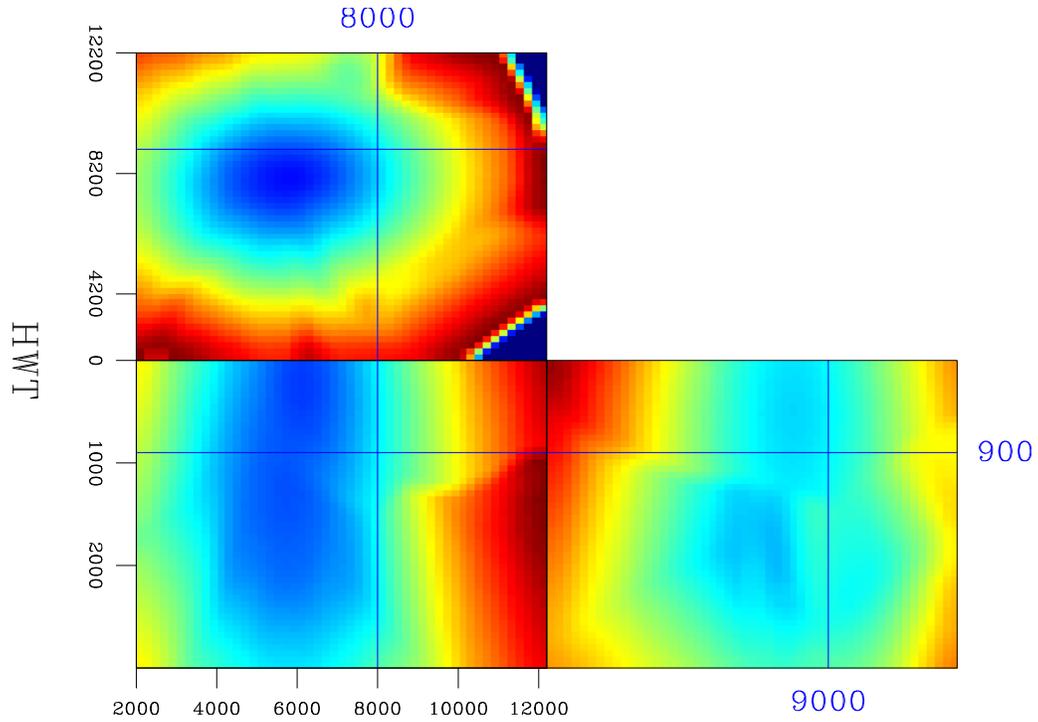


Figure 13: The SEG-EAGE salt model. A traveltime cube obtained using HWT. The traveltimes are interpolated on the rectangular grid. The selected traveltimes correspond to the shortest path rays. `paul2-hwtsalt` [ER]

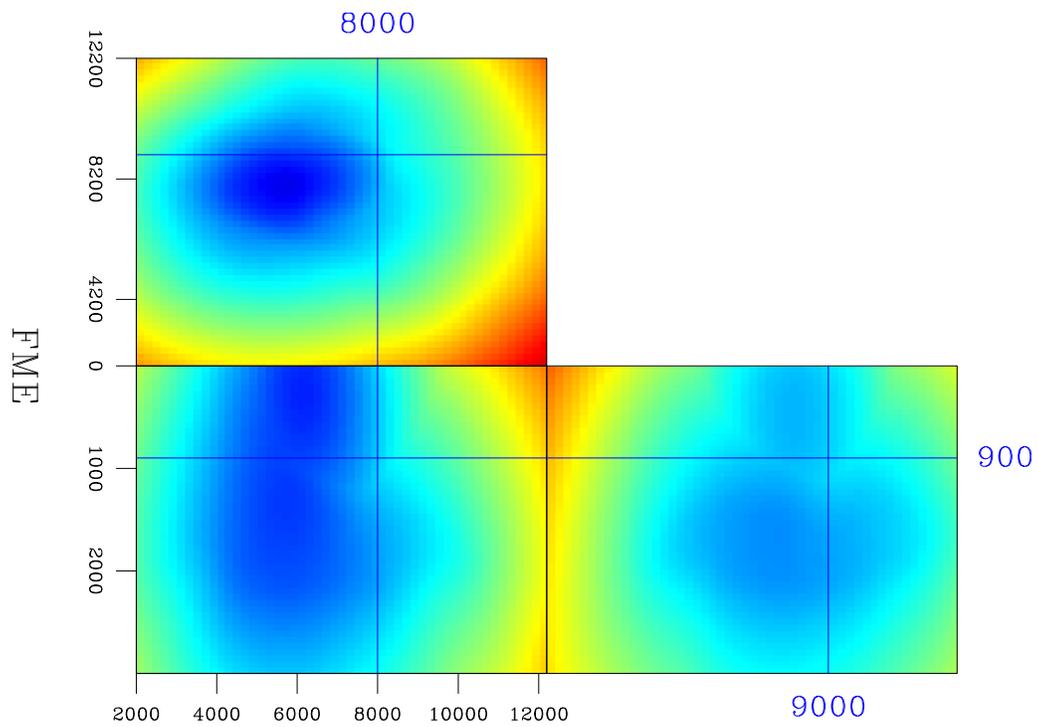


Figure 14: The SEG-EAGE salt model. A traveltime cube obtained using a fast-marching first-arrival eikonal solver. `paul2-fmesalt` [ER]

Table 1: Comparison of methods for travelttime computation

Wavefront Tracing finds the solution to a system of PDEs	Ray Tracing finds the solution to a system of ODEs
Wavefront Tracing gives the output in ray coordinates computes multiple arrivals	Eikonal Solvers gives the output in Cartesian coordinates computes one arrival
Wavefront Tracing finds a new wavefront by finite-differences	Wavefront Construction finds a new wavefront by ray tracing

1990; van Trier and Symes, 1991; Popovici and Sethian, 1997; Fomel, 1997), and wavefront construction (WC) (Vinje et al., 1993). Table 1 summarizes the comparison.

HWT has its output in ray coordinates, the same domain as PRT. However, PRT is done by solving a system of ordinary differential equations (ODE) in the physical domain, while in HWT the solution is obtained by solving a system of partial differential equations (PDE) using finite-differences in the ray coordinate domain.

Both HWT and ES are finite-difference methods. However, HWT represents a finite-difference method in the ray domain, while ES represent finite-difference methods in the Cartesian domain. Also, HWT generates all the arrivals, while the ES generate only one arrival, typically the first.

Finally, HWT is similar to WC in that both compute each wavefront from the preceding one. However, WC involves ray tracing from one wavefront to the next, while in HWT one wavefront is generated from the preceding by finite-differences in the ray domain.

CONCLUSIONS AND FUTURE WORK

My extension of Huygens wavefront tracing to three dimensions retains the benefits offered by the method in two dimensions, namely its stability in regions of high velocity contrast, its accuracy, and its computational efficiency.

As for any other travelttime method that involves computations in the ray domain, the interpolation to a Cartesian grid remains the major problem. In the current implementation, the interpolation is about one order of magnitude more expensive than the travelttime computation itself.

A possible direction for future work is to modify HWT to a grid-adaptive finite-difference

method. HWT requires computing derivatives along the wavefronts. As the distance between adjacent points on the wavefront increases, the accuracy of those derivatives decreases. One possible way to increase the accuracy is to adapt the sampling of the wavefront in the ray domain, as done with finite-difference methods in the physical domain (Symes et al., 1999).

ACKNOWLEDGEMENTS

I would like to thank Bill Symes for many useful discussions. I also wish to thank James Rickett for the support in generating the wave-equation modeling examples.

REFERENCES

- Červený, V., 1987, Ray tracing algorithms in three-dimensional laterally varying layered structures: *Seismic Tomography*, 99–134.
- Fomel, S., 1997, A variational formulation of the fast marching eikonal solver: *SEP-95*, 127–147.
- Nichols, D., Farmer, P., and Palacharla, G., 1998, Improving prestack imaging by using a new ray selection method: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1546–1549.
- Popovici, A. M., and Sethian, J., 1997, Three-dimensional traveltimes computation using the fast marching method: 67th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1778–1781.
- Sava, P., and Biondi, B., 1997, Multivalued traveltimes interpolation: *SEP-95*, 115–126.
- Sava, P. C., and Fomel, S., 1998, Huygens wavefront tracing: A robust alternative to ray tracing: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1961–1964.
- Symes, W. W., Belfi, C. S., and Qian, J., 1999, Adaptive grid eikonal solvers: 5th SIAM Conference on Mathematical and Computational Issues in the Geosciences, 90.
- van Trier, J., and Symes, W. W., 1991, Upwind finite-difference calculation of traveltimes: *Geophysics*, **56**, no. 6, 812–821.
- Vidale, J. E., 1990, Finite-difference calculation of traveltimes in three dimensions: *Geophysics*, **55**, no. 5, 521–526.
- Vinje, V., Iversen, E., and Gjoystdal, H., 1993, Traveltimes and amplitude estimation using wavefront construction: *Geophysics*, **58**, no. 8, 1157–1166.

An adaptive finite difference method for traveltimes and amplitude¹

Jianliang Qian and William W. Symes²

ABSTRACT

The eikonal equation with point source is difficult to solve with high order accuracy because of the singularity of the solution at the source. All the formally high order schemes turn out to be first order accurate without special treatment of this singularity. Adaptive upwind finite difference methods based on high order ENO (Essentially NonOscillatory) Runge-Kutta difference schemes for the paraxial eikonal equation overcome this difficulty. The method controls error by automatic grid refinement and coarsening based on an *a posteriori* error estimation. It achieves prescribed accuracy at far lower cost than fixed grid methods. Reliable auxiliary quantities, such as take-off angle and geometrical spreading factor, are by-products.

INTRODUCTION

Many finite difference methods have been introduced to compute the traveltimes for isotropic media directly on a regular grid (Reshef and Kosloff, 1986; Vidale, 1988; van Trier and Symes, 1991; Schneider et al., 1992; Qin et al., 1992; Schneider, 1995; El-Mageed, 1996; El-Mageed et al., 1997; Fomel, 1997; Popovici and Sethian, 1997).

The traveltimes field is mostly smooth, and the use of upwind differencing (in all of the cited methods) confines the errors due to singularities which develop away from the source point. The source point itself is, however, also a singularity. The truncation error of a p th order method is dominated by the product of $(p + 1)$ st derivatives of the time field and the $(p + 1)$ st power of the step(s). The $(p + 1)$ st derivatives of the time field, however, go like the $(-p + 1)$ th power of the distance to the source. Therefore, near the source — when the distance is on the order of the step — the truncation error is quadratic in the step, i.e., first order. This inaccuracy spreads throughout the computation, and renders all higher order methods first-order convergent unless the scheme is modified near the source. The issue is not academic: the first-order error is sizeable, as we shall show. Moreover, it prevents reliable computation of auxiliary quantities such as takeoff angle and amplitude.

In this paper, we show how to use adaptive gridding concepts commonplace in the numerical solution of ordinary differential equations to resolve this difficulty. This work refines

¹This paper will also appear in TRIP 1999 Annual Report.

²email: jlqian@caam.rice.edu, symes@caam.rice.edu

and extends the method introduced in Belfi and Symes (1998). The efficiencies achieved by adaptive gridding are considerable — usually more than an order of magnitude gain in computation time for problems of typical exploration size. We also obtain dramatic improvements in the accuracy of takeoff angle computations and, therefore, for other geometrical acoustics quantities as well.

PARAXIAL EIKONAL EQUATION

Denote by (x_s, z_s) the coordinates of a source point, and by (x, z) the coordinates of a general point in the subsurface. The first arrival traveltime field $\tau(x, z; x_s, z_s)$ is the viscosity solution of the eikonal equation

$$\left(\frac{\partial \tau}{\partial x}\right)^2 + \left(\frac{\partial \tau}{\partial z}\right)^2 = s^2(x, z) \quad (1)$$

with the initial condition

$$\lim \left(\frac{\tau(x, z; x_s, z_s)}{\sqrt{(x - x_s)^2 + (z - z_s)^2}} - \frac{1}{v(x, z)} \right) = 0$$

as $(x, z) \rightarrow (x_s, z_s)$, where v is the velocity, $s = \frac{1}{v}$ is the slowness (Lions, 1982).

In some seismic applications, the traveltime field is only needed in regions where

$$\frac{\partial \tau}{\partial z} \geq s \cos \theta_{\max} > 0,$$

i.e., along downgoing, first-arriving rays making an angle $\leq \theta_{\max} < \frac{\pi}{2}$ with the vertical.

To enforce this condition, we modify the eikonal equation as an evolutionary equation in depth, as suggested by Gray and May (1994):

$$\frac{\partial \tau}{\partial z} = H\left(\frac{\partial \tau}{\partial x}\right) = \sqrt{\text{smmax}\left(s^2 - \left(\frac{\partial \tau}{\partial x}\right)^2, s^2 \cos^2 \theta_{\max}\right)}, \quad (2)$$

where smmax is a smoothed max function:

$$\text{smmax}(x, a) = \begin{cases} \frac{1}{2}a & \text{if } x < 0, \\ \frac{1}{2}a + 2\frac{x^4}{a^3}\left(1 - \frac{4}{5}\frac{x}{a}\right) & \text{if } 0 \leq x < \frac{a}{2}, \\ x + 2\frac{(x-a)^4}{a^3}\left(1 + \frac{4}{5}\frac{x-a}{a}\right) & \text{if } \frac{a}{2} \leq x < a, \\ x & \text{if } x \geq a. \end{cases}$$

This equation defines a stable nonlinear evolution in z , suitable for explicit finite difference discretization. The solution τ is identical to the solution of the eikonal equation provided that the ray makes an angle $\leq \theta_{\max} < \frac{\pi}{2}$ with the vertical.

THE ADVECTION EQUATION FOR TAKE-OFF ANGLE

The amplitude satisfies the zeroth order transport equation (Červený et al., 1977):

$$\nabla\tau \cdot \nabla A + \frac{1}{2}A\nabla^2\tau = 0. \quad (3)$$

If the traveltimes field has been found by solving the eikonal equation, then equation (3) is a first order advection equation. However, we see that the Laplacian of traveltimes field is involved in the transport equation, which implies that we need a third order accurate traveltimes field to get a first order accurate amplitude field (Symes, 1995; El-Mageed, 1996; El-Mageed et al., 1997). To avoid this complexity, we use another approach to compute the amplitude.

In 2D isotropic media, the amplitude satisfies (Červený et al., 1977; Friedlander, 1958)

$$\begin{aligned} A &= \frac{v}{2\pi\sqrt{2}\sqrt{|J|}} \\ &= \frac{v}{2\pi\sqrt{2}}\sqrt{\nabla\tau \times \nabla\phi}, \end{aligned}$$

where $J(x, z; x_s, z_s)$ is the Jacobian of the transformation from Cartesian coordinates (x, z) to ray coordinates (τ, ϕ) , τ is the traveltimes, $\phi = \phi(x, z; x_s, z_s)$ is the take-off angle from source point (x_s, z_s) to a general point (x, z) in the subsurface:

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial x}{\partial \tau} & \frac{\partial z}{\partial \tau} \\ \frac{\partial x}{\partial \phi} & \frac{\partial z}{\partial \phi} \end{vmatrix} \\ &= \begin{vmatrix} \frac{\partial \tau}{\partial x} & \frac{\partial \tau}{\partial z} \\ \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial z} \end{vmatrix}^{-1} \\ &= \frac{1}{\nabla\tau \times \nabla\phi}, \end{aligned}$$

where $\nabla\phi$ and $\nabla\tau$ are the gradients of take-off angle and traveltimes, respectively.

Since the take-off angle ϕ is constant along any ray,

$$\nabla\tau \cdot \nabla\phi = \frac{\partial \tau}{\partial x} \frac{\partial \phi}{\partial x} + \frac{\partial \tau}{\partial z} \frac{\partial \phi}{\partial z} = 0. \quad (4)$$

That is, the gradient $(\frac{\partial \tau}{\partial x}, \frac{\partial \tau}{\partial z})$ is the wavefront normal which is tangential to the ray; the gradient $(\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial z})$ is tangential to the wavefront.

However, the gradient of the take-off angle depends on the second order derivative of traveltimes, so that we need third order accurate traveltimes to get a first order accurate gradient of take-off angle. Zhang (1993) used this equation in polar coordinates to compute the geometrical spreading factor, but his computation of the take-off angle was based on the first order traveltimes field. Consequently, the gradient of take-off angle computed by his scheme was inaccurate. Vidale (1990) encountered a similar difficulty.

ENO AND WENO FOR EIKONAL EQUATIONS

There is a large number of competing finite difference and related schemes for the solution of the eikonal equation suggested in the literature; we have chosen to use the essentially nonoscillatory (“ENO”) schemes of Osher and Sethian (1988) and Osher and Shu (1991) and the related weighed ENO (WENO) schemes (Liu et al., 1994; Shu, 1997; Jiang and Peng, 1997) for the following reasons: (1) stable schemes of arbitrarily high order accuracy exist, permitting accurate solutions on coarse grids (which is critical to the mesh refinement or coarsening); (2) versions exist in any dimension so that we can extend our methodology to the three-dimensional case straightforwardly (El-Mageed, 1996; El-Mageed et al., 1997; Qian and Symes, 1998).

Our adaptive scheme is based on the 2nd and 3rd order WENO difference schemes introduced by Jiang and Peng (1997). These are in turn extensions of 2nd and 3rd order ENO difference schemes, which we present first.

For a function f of the space variable (x, z) in the computational domain, we write

$$\begin{aligned} f_i^k &= f(x_i, z_k), \\ (x_i, z_k) &= (x_{\min} + (i-1)\Delta x, z_{\min} + (k-1)\Delta z). \end{aligned}$$

Let $\tau_i^k = \tau(x_i, z_k; x_s, z_s)$ and define the forward D^+ and backward D^- finite difference operators

$$D_x^\pm \tau_i^k = \frac{\pm[\tau_{i\pm 1}^k - \tau_i^k]}{\Delta x}.$$

The second and third order ENO refinements of $D_x^\pm \tau$ are

$$\begin{aligned} D_x^{\pm, 2} \tau &= D_x^\pm \tau \mp \frac{1}{2} \Delta x m(D_x^\pm D_x^\pm \tau, D_x^- D_x^+ \tau), \\ D_x^{\pm, 3} \tau &= D_x^{\pm, 2} \tau - \frac{1}{6} (\Delta x)^2 m(D_x^\pm D_x^\pm D_x^\pm \tau, D_x^+ D_x^+ D_x^- \tau, D_x^+ D_x^- D_x^- \tau), \end{aligned}$$

where

$$m(x, y) = \min(\max(x, 0), \max(y, 0)) + \max(\min(x, 0), \min(y, 0)).$$

The upwind ENO approximations for $\frac{\partial \tau}{\partial x}$ are

$$\widehat{D}_x^n \tau = \text{modmax}(\max(D_x^{-, n} \tau, 0), \min(D_x^{+, n} \tau, 0)),$$

for $n = 2, 3$, where modmax function returns the larger value in modulus.

The second order and third order ENO Runge-Kutta steps are

$$\begin{aligned} \delta_2^1 \tau &= \Delta z H(\widehat{D}_x^2 \tau), \\ \delta_2^2 \tau &= \frac{1}{2} (\delta_2^1 \tau + \Delta z H(\widehat{D}_x^2 (\tau + \delta_2^1 \tau))), \end{aligned} \tag{5}$$

and

$$\begin{aligned}
\delta_3^1 \tau &= \Delta z H(\widehat{D}_x^3 \tau), \\
\delta_3^2 \tau &= \frac{1}{4} (3\delta_3^1 \tau + \Delta z H(\widehat{D}_x^3 (\tau + \delta_3^1 \tau))), \\
\delta_3^3 \tau &= \frac{1}{3} (2\delta_3^2 \tau + 2\Delta z H(\widehat{D}_x^3 (\tau + \delta_3^2 \tau))), .
\end{aligned} \tag{6}$$

The depth step Δz must satisfy the stability condition:

$$\Delta z \leq \Delta z_{cfl} = \frac{0.9 \Delta x}{\tan(\theta_{max})}.$$

We have typically chosen $\Delta z = 0.9 * \Delta z_{cfl}$.

The n th order scheme is then

$$\tau^{k+1} = \tau^k + \delta_n^k \tau^k \tag{7}$$

for $k = 0, 1, 2, \dots, n = 2, 3$.

However, we have observed that the gradient of the take-off angle based on the third order ENO traveltimes is too noisy to give us a smooth amplitude function. To alleviate this phenomenon, instead of ENO 3rd order refinements, we use WENO 5th order refinement (Jiang and Peng, 1997) to compute $D_x^\pm \tau$ in the third order Runge-Kutta step, which gives us a smooth amplitude field.

The WENO 5th order schemes for $D_x^\pm \tau_i$ are

$$\begin{aligned}
D_x^{\pm,5} \tau_i &= \frac{1}{12} (-D_x^+ \tau_{i-2} + 7D_x^+ \tau_{i-1} + 7D_x^+ \tau_i - D_x^+ \tau_{i+1}) \\
&\pm \Delta x \Phi^{WENO} (D_x^- D_x^+ \tau_{i\pm 2}, D_x^- D_x^+ \tau_{i\pm 1}, D_x^- D_x^+ \tau_i, D_x^- D_x^+ \tau_{i\mp 1}),
\end{aligned}$$

where

$$\Phi^{WENO}(a, b, c, d) = \frac{1}{3} \omega_0 (a - 2b + c) + \frac{1}{6} (\omega_2 - \frac{1}{2}) (b - 2c + d)$$

with weights defined as

$$\begin{aligned}
w_0 &= \frac{\alpha_0}{\alpha_0 + \alpha_1 + \alpha_2}, w_2 = \frac{\alpha_2}{\alpha_0 + \alpha_1 + \alpha_2}, \\
\alpha_0 &= \frac{1}{(\delta + \beta_0)^2}, \alpha_1 = \frac{1}{(\delta + \beta_1)^2}, \alpha_2 = \frac{1}{(\delta + \beta_2)^2}, \\
\beta_0 &= 13(a - b)^2 + 3(a - 3b)^2, \\
\beta_1 &= 13(b - c)^2 + 3(b + c)^2, \\
\beta_2 &= 13(c - d)^2 + 3(3c - d)^2.
\end{aligned}$$

In the denominators above, we added a small positive number δ to avoid dividing by zero.

WENO FOR ADVECTION EQUATIONS

Recall the advection equation for the take-off angle,

$$\frac{\partial \tau}{\partial x} \frac{\partial \phi}{\partial x} + \frac{\partial \tau}{\partial z} \frac{\partial \phi}{\partial z} = 0.$$

To match with the evolutionary form of the eikonal equation in depth, we formulate the advection equation as an evolution equation in depth as well, i.e.,

$$\frac{\partial \phi}{\partial z} = -\left(\frac{\partial \tau}{\partial z}\right)^{-1} \frac{\partial \tau}{\partial x} \frac{\partial \phi}{\partial x}. \quad (8)$$

To take advantage of the accuracy of traveltime provided by an ENO (or WENO) Runge-Kutta third order scheme for eikonal equation, we embed the third order scheme for equation (8) into the third order scheme for the eikonal equation.

To be more precise, we introduce the approximations for x-derivative and z-derivative of τ in the above advection equation:

$$\frac{\partial \tau}{\partial x} \approx \widehat{D}_x^5 \tau = \text{modmax}(\max(D_x^{-,5} \tau, 0), \min(D_x^{+,5} \tau, 0)),$$

$$\frac{\partial \tau}{\partial z} \approx H(\widehat{D}_x^5 \tau).$$

However, we must be careful in defining the upwind ENO difference approximation for $\frac{\partial \phi}{\partial x}$ because the ENO choice of stencil is too sensitive to the zeros of solution (Liu et al., 1994; Jiang and Shu, 1996). Instead, we will use a weighted ENO (WENO) scheme to approximate the derivative. Because the coefficient of the discretized advection equation has only second order accuracy, which is computed from the eikonal equation by the third order scheme, we use a third order WENO scheme to approximate the derivative $\frac{\partial \phi}{\partial x}$. The third order WENO scheme is based on the second order ENO stencils, so it does not give rise to any new complexities in the coding.

The third order upwind WENO approximations to $\frac{\partial \phi}{\partial x}$ (Jiang and Peng, 1997) are

$$\begin{aligned} \left(\frac{\partial \phi}{\partial x}\right)_i^k &\approx (D_x^{-,w,3} \phi)_i^k = \frac{1}{2}(D_x^+ \phi_{i-1}^k + D_x^+ \phi_i^k) - \frac{w_-}{2}(D_x^+ \phi_{i-2}^k - 2D_x^+ \phi_{i-1}^k + D_x^+ \phi_i^k), \\ \left(\frac{\partial \phi}{\partial x}\right)_i^k &\approx (D_x^{+,w,3} \phi)_i^k = \frac{1}{2}(D_x^+ \phi_{i-1}^k + D_x^+ \phi_i^k) - \frac{w_+}{2}(D_x^+ \phi_{i+1}^k - 2D_x^+ \phi_i^k + D_x^+ \phi_{i-1}^k), \end{aligned}$$

where

$$\begin{aligned} w_- &= \frac{1}{1+2r_-^2}, r_- = \frac{\delta + (D^- D^- \phi_{i-1}^k)^2}{\delta + (D^- D^+ \phi_i^k)^2}, \\ w_+ &= \frac{1}{1+2r_+^2}, r_+ = \frac{\delta + (D^+ D^+ \phi_i^k)^2}{\delta + (D^- D^+ \phi_i^k)^2}, \end{aligned}$$

and δ is a small positive constant to prevent the denominators from becoming zero.

Now we define the upwind WENO difference for $\frac{\partial \phi}{\partial x}$, which corresponds to the upwind direction of $\frac{\partial \tau}{\partial x}$ as follows:

$$D_x^{up} \phi_i^k = \begin{cases} D_x^{-w,3} \phi_i^k & \text{if } \widehat{D}_x^5 \tau_i^k \geq 0, \\ D_x^{+w,3} \phi_i^k & \text{else.} \end{cases}$$

Finally, we can formulate the third order WENO Runge-Kutta scheme for the advection equation as

$$\begin{aligned} \delta_3^1 \phi &= \Delta z \Psi(\widehat{D}_x^5 \tau, \phi), \\ \delta_3^2 \phi &= \frac{1}{4} (3\delta_3^1 \phi + \Delta z \Psi(\widehat{D}_x^5(\tau + \delta_3^1 \tau), \phi + \delta_3^1 \phi)), \\ \delta_3^3 \phi &= \frac{1}{3} (2\delta_3^2 \phi + 2\Delta z \Psi(\widehat{D}_x^5(\tau + \delta_3^2 \tau), \phi + \delta_3^2 \phi)), \end{aligned} \quad (9)$$

where

$$\Psi(\tau, \phi) = -\frac{\widehat{D}_x^5 \tau D_x^{up} \phi}{H(\widehat{D}_x^5 \tau)}.$$

The 3rd order scheme for ϕ is then

$$\phi^{k+1} = \phi^k + \delta_3^3 \phi^k \quad (10)$$

for $k = 0, 1, 2, \dots$

But this scheme is really a second order scheme because the coefficients have only second order accuracy.

ADAPTIVE GRID METHOD AND ITS IMPLEMENTATION

The adaptive multigrid method has been widely used for the accurate and efficient solution of PDE (Berger and Olinger, 1984; Berger and LeVeque, 1997), which uses an adaptive local error estimate to refine or coarsen the computational grid near the singularity or discontinuity of solution. For initial value problems of ODE, most state-of-the-art software uses adaptive timestepping algorithms where the timesteps are inductively chosen so that some estimate of the local (one step) error at each step is less than some quantity related to the user-defined tolerance.

The essential principle is simple. It is based on a hierarchy of difference schemes of various orders. Presumably a higher order step is more accurate than a lower order step, so the higher order step can serve as an *ersatz* for the exact solution of the differential equation with the same data. Therefore, one can combine the step computations of two different orders to obtain a so-called *a posteriori* estimate of the truncation error for the lower order step. Since

the lower and higher order truncation errors stand in a known asymptotic relation, this permits an estimate of the higher order truncation error as well. The asymptotic form of the truncation error then permits prediction of a step that will result in a lower order truncation error less than a user-specified tolerance. So long as the steps are selected to maintain this local error, standard theory predicts that the higher order global error, i.e., the actual error in the solution computed using the higher order scheme, will be proportional to the user-specified tolerance.

This straightforward idea is embedded in most modern software packages for solutions of ordinary differential equations (Gear, 1971). Its use for partial differential equations is a little more complicated because it is usually necessary to adjust the grid of the non-evolution variables along with the evolution step. The solution of the (paraxial) eikonal equation changes in a sufficiently predictable way to make grid adjustment practical.

To initialize our algorithm, the user supplies a local error tolerance ϵ ; σ_1 and σ_2 are two user-defined positive functions of ϵ which are used to control the coarsening and refinement, for example, we can take $\sigma_1 = 0.1\epsilon$ and $\sigma_2 = \epsilon$. We use the 2nd and 3rd order eikonal solvers (equations (5) and (6)), and estimate the truncation error of the 2nd order scheme at the k th step as $e_2^k = \delta_2^2 \tau^k - \delta_3^3 \tau^k$. So long as $\sigma_1(\epsilon) \leq e_2^k \leq \sigma_2(\epsilon)$, we simply proceed to the next step; it is well known and explained in the references already cited that efficient adaptive stepping requires rather loose control of the local error. When $e_2^k < \sigma_1(\epsilon)$, we increase the step by a factor of two, i.e., $\Delta z \leftarrow 2\Delta z$, and recompute e_2^k . Similarly, when $e_2^k > \sigma_2(\epsilon)$, we decrease the step by a factor of two. As soon as the local error is once again within the tolerance interval we continue depth-stepping. A very important point: we retain the 3rd order (a more accurate one) computation of τ at the end of each depth step, discarding the 2nd order computation, which is used only in step control.

The usual step adjustment in ODE solvers would change Δz by a factor computed from the asymptotic form of the truncation error. This is impractical for a PDE application, because it would require an arbitrary adjustment of the spatial grid (i.e., the x -grid in the eikonal scheme) and, therefore, expensive interpolation. Scaling Δz by a factor of two, however, implies that stability may be maintained by scaling Δx by the same factor. For coarsening, this means simply throwing out every other grid point, i.e., no interpolation at all, which dramatically reduces the floating point operations required. Since the typical behaviour of the traveltimes field is to become smoother as one moves away from the source, the truncation errors tend in general to decrease. Therefore, most of the grid adjustments are coarsenings and very little or no interpolation is required.

One final detail must be supplied to produce a working algorithm. Since the traveltimes field is nonsmooth at the source point, the truncation error analysis on which the adaptive step selection criterion is based is not valid there. So it is necessary to produce a smooth initial traveltimes field. We do this by estimating the largest $z_{\text{init}} > 0$, at which the constant velocity traveltimes is in error by less than $\sigma_2(\epsilon)$. Details of the z_{init} calculation are given in (Belfi and Symes, 1998). Having initialized τ at z_{init} , the algorithm invokes adaptive gridding. Since z_{init} is quite small, τ changes rapidly, resulting in a large number of grid refinements at the outset. However, no interpolation is performed, as τ is given analytically on $z = z_{\text{init}}$. This initial very fine grid is rapidly coarsened as the depth stepping proceeds.

In our current implementation, we maintain a data structure for the computational grid that is independent of the output grid; the desired quantities are calculated on the computational grid and interpolated back to the output grid.

A simplified algorithm framework is as follows:

- Input: $\epsilon, z^0, \Delta z^0, \Delta x^0, n = 0, maxref$.
- Marching step:
 - do while $z^n \leq \text{target depth}$,
 - $z^n \leftarrow z^n + \Delta z^n, n \leftarrow n + 1$,
 - compute $e_2^k(\tau^n, \Delta z^n)$;
 - if $e_2^k(\tau^n, \Delta z^n) \leq \sigma_1(\epsilon)$ and $k \geq 0$,
 - * $\Delta z^n \leftarrow 2\Delta z^n$,
 - * $\Delta x^n \leftarrow 2\Delta x^n$,
 - * $k \leftarrow k - 1$.
 - else if $e_2^k(\tau^n, \Delta z^n) \geq \sigma_2(\epsilon)$ and $k \leq maxref$,
 - * $\Delta z^n \leftarrow \Delta z^n / 2$,
 - * $\Delta x^n \leftarrow \Delta x^n / 2$,
 - * $k \leftarrow k + 1$.
 - end if
 - end do

NUMERICAL EXPERIMENTS

In the first example, we test our method on a constant velocity model ($v = 1\text{km/s}$) with two-dimensional geometry $\{(x, z) : -0.5 \leq x \leq 0.5, 0 \leq z \leq 1.0\}$.

In this case, all the desired quantities have an obvious analytical solution to compare against the computed solutions. We compare the results obtained by two approaches, the first approach is computing all the quantities on the fixed output grid which reduces to be a first order method, the second is computing all the quantities by adaptive gridding. The output grid is 51×51 with $\Delta x = \Delta z = 0.02$. For adaptive grid, $maxref$ is set to be 5 with the coarsest grid 17×17 , and the error tolerance is set to be 0.0001.

Figures 1 and 2 show the traveltimes contours by the two approaches, from which we cannot see any differences between two approaches. Similarly, we cannot see any differences in accuracy between two approaches from take-off angle, as shown in figures 3 and 4. This is due to visual limitations of graphics.

Figures 5 and 6 are contours of τ_x computed by two approaches. We can see that τ_x by the fixed grid approach is oscillating, but τ_x by the adaptive grid traveltimes solver is convergent. Similar conclusions can be drawn for τ_z , as shown in figures 7 and 8.

Now we come to take-off angle derivatives. Figure 9 and 10 are contours of ϕ_x by two approaches. Because the coefficients in the advection equation for take-off angle depend on the traveltimes gradient, any first order traveltimes solver results in inaccurate $\nabla\tau$ which leads to the divergence of ϕ_x , as shown in figure 9. However, the adaptive gridding approach gives us accurate traveltimes gradients, which leads to the convergent ϕ_x , as shown in figure 10. Similar conclusions can be drawn for ϕ_z , as shown in figure 11 and 12.

To illustrate further the difference of accuracy between two approaches, figures 13 and 14 show the error distributions of ϕ_x and ϕ_z .

Finally, the amplitudes based on the gradients of traveltimes and take-off angle computed by the two approaches are shown in figures 15 and 16, one is divergent by the fixed grid approach, another accurate by the adaptive gridding approach.

In the second example, we use our new eikonal and amplitude solver in 2D Kirchhoff prestack migration and inversion. Figure 17 shows the impulse response by inversion with an ENO 3rd order eikonal solver. Figure 18 shows the impulse response by inversion with a WENO 5th order eikonal solver. By figures 17 and 18, we can see that the WENO 5th order solver does give us a more smooth amplitude than ENO 3rd solver.

Figure 1: Traveltimes for constant velocity model: fixed grid
[jianliang1-comtau.fir](#) [CR]

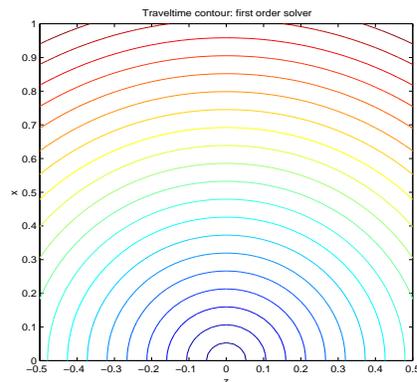


Figure 2: Traveltimes for constant velocity model: adaptive grid
[jianliang1-comtau.ada](#) [CR]

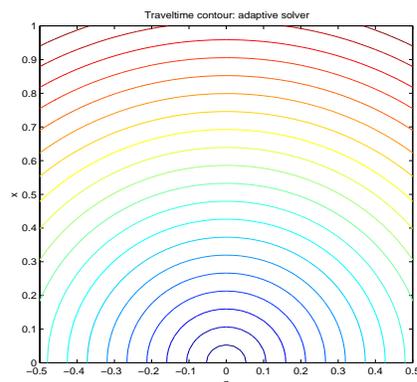


Figure 3: Take-off angle ϕ for constant velocity model: fixed grid
 [jianliang1-cometa.fir] [CR]

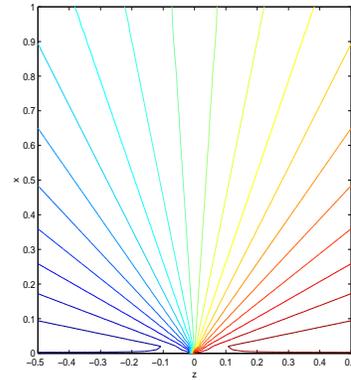


Figure 4: Take-off angle ϕ for constant velocity model: adaptive grid
 [jianliang1-cometa.ada] [CR]

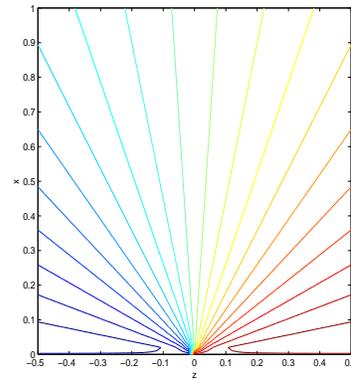


Figure 5: Traveltime x derivative τ_x for constant velocity model: fixed grid
 [jianliang1-comtaux.fir] [CR]

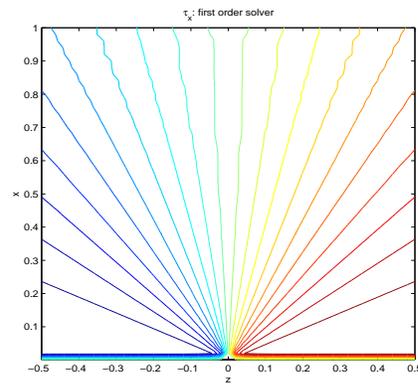


Figure 6: Traveltime x derivative τ_x for constant velocity model: adaptive grid
 [jianliang1-comtaux.ada] [CR]

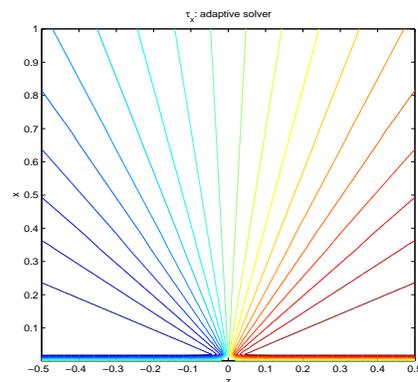


Figure 7: Traveltime z derivative τ_z for constant velocity model: fixed grid `jianliang1-comtauz.fir` [CR]

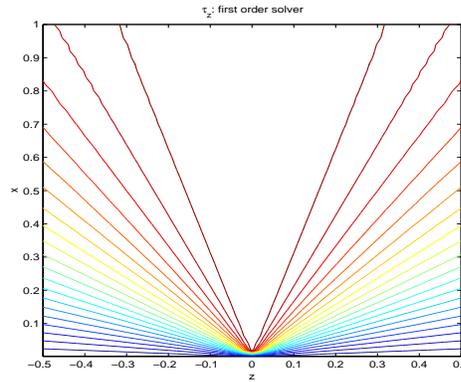


Figure 8: Traveltime z derivative τ_z for constant velocity model: adaptive grid `jianliang1-comtauz.ada` [CR]

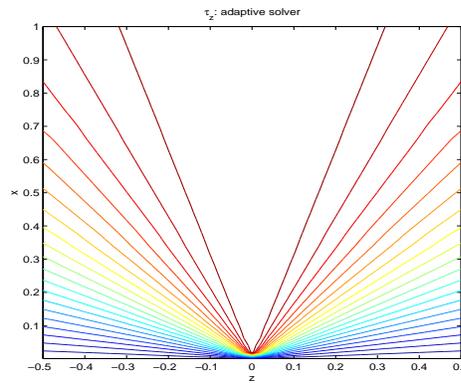


Figure 9: Take-off angle x derivative ϕ_x at $z = 1$ for constant velocity model: fixed grid `jianliang1-angxder.fir` [CR]

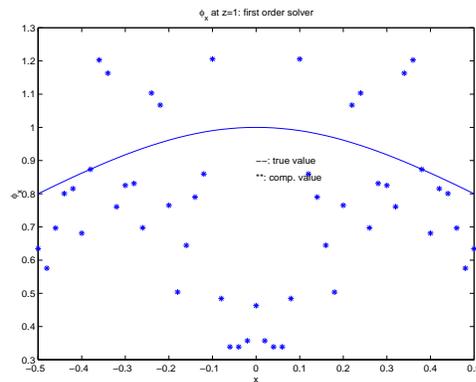


Figure 10: Take-off angle x derivative ϕ_x at $z = 1$ for constant velocity model: adaptive grid `jianliang1-angxder.ada` [CR]

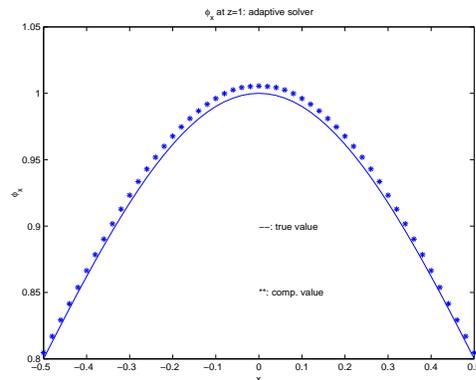


Figure 11: Take-off angle z derivative ϕ_z at $z = 1$ for constant velocity model: fixed grid
 [jianliang1-angzder.fir] [CR]

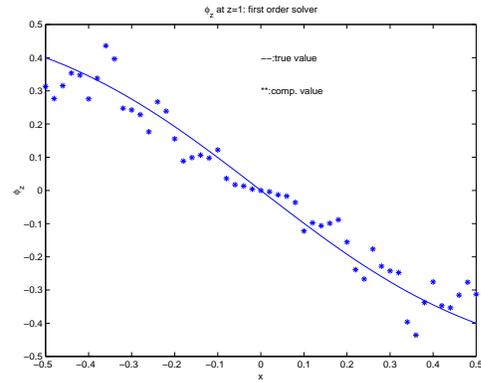


Figure 12: Take-off angle z derivative ϕ_z at $z = 1$ for constant velocity model: adaptive grid
 [jianliang1-angzder.ada] [CR]

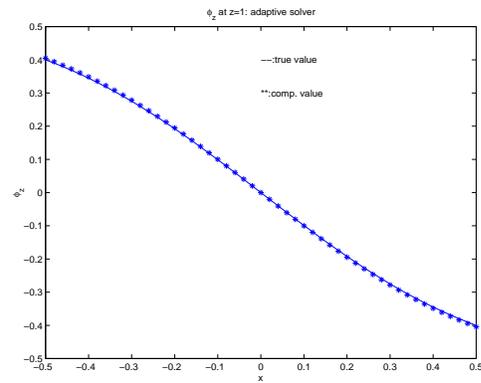


Figure 13: Error distribution in take-off angle derivatives ϕ_x and ϕ_z : fixed grid
 [jianliang1-angxzerr.fir] [CR]

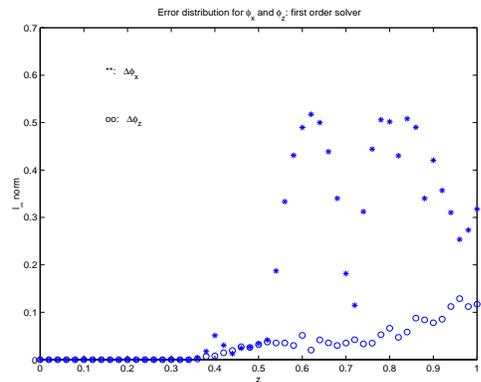


Figure 14: Error distribution in take-off angle derivatives ϕ_x and ϕ_z : adaptive grid
 [jianliang1-angxzerr.ada] [CR]

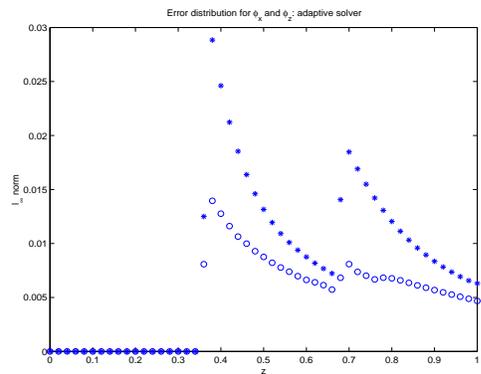


Figure 15: Amplitude for constant velocity model: fixed grid
`jianliang1-amp.fir` [CR]

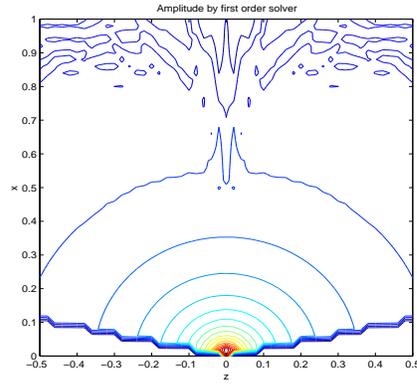


Figure 16: Amplitude for constant velocity model: adaptive grid
`jianliang1-amp.ada` [CR]

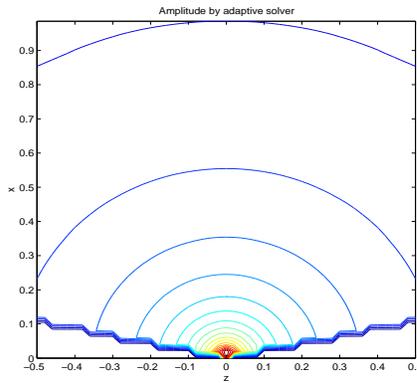


Figure 17: The impulse response by inversion with ENO 3rd order adaptive eikonal solver
`jianliang1-eno600.ada` [NR]

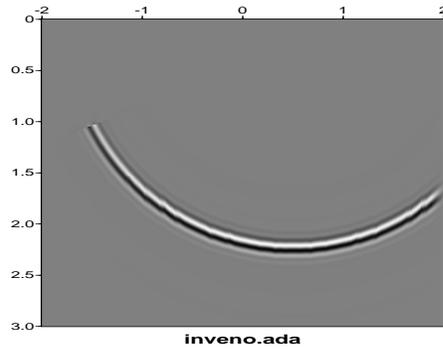
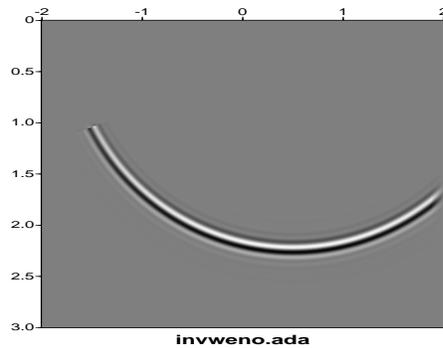


Figure 18: The impulse response by inversion with WENO 5th order adaptive eikonal solver
`jianliang1-weno600.ada` [NR]



CONCLUSIONS

In this paper we formulated a paraxial eikonal equation with depth as evolution direction. Then we presented high order ENO difference schemes to solve the eikonal equation for traveltime and the advection equation for take-off angle. To deal with the singularity of point source, we proposed a new adaptive traveltime eikonal solver and detailed the implementation. Numerical experiments showed that the new method is not only accurate but also gives us efficiency gain of more than an order of magnitude in computational time. The extension to 3D isotropic media is straightforward.

ACKNOWLEDGEMENTS

The authors thank the National Science Foundation, the Office of Naval Research, the U. S. Department of Energy, and the sponsors of The Rice Inversion Project for support of this work. The authors performed part of the research reported here while guests of the Stanford Exploration Project; JLQ and WWS thank SEP members and director Prof. Jon Claerbout for their hospitality and many stimulating discussions. JLQ thanks Jim Berryman, Nizar Chemingui and James Rickett for proofreading the paper.

REFERENCES

- Belfi, C., and Symes, W. W., 1998, An adaptive ENO algorithm for the eikonal equation: Annual Report, The Rice Inversion Project, Rice University.
- Berger, M., and LeVeque, R. J., 1997, Adaptive mesh refinement using wave-propagation algorithms for hyperbolic systems: Preprint, Courant Institute.
- Berger, M., and Olinger, J., 1984, Adaptive mesh refinement for hyperbolic partial differential equations: *J. Comput. Phys.*, **53**, 484–512.
- Červený, V., Molotkov, I. A., and Pšenčík, I., 1977, Ray method in seismology: Univerzita Karlova press.
- El-Mageed, M., Kim, S., and Symes, W. W., 1997, 3-D kirchhoff migration using finite difference traveltimes and amplitudes: Annual Report, The Rice Inversion Project, Rice University.
- El-Mageed, M. A., 1996, 3D first arrival traveltimes and amplitudes via eikonal and transport finite differences solvers: Ph.D. thesis, Department of Computational and Applied mathematics, Rice University, Houston, TX77251-1892.
- Fomel, S., 1997, A variational formulation of the fast marching eikonal solver: SEP95, Stanford Exploration Project, Stanford University.
- Friedlander, F., 1958, Sound pulses: Cambridge University Press.

- Gear, C. W., 1971, Numerical initial value problems in ordinary differential equations: Englewood Cliffs, N. J.: Prentice-Hall.
- Gray, S., and May, W., 1994, Kirchhoff migration using eikonal equation traveltimes: *Geophysics*, **59**, 810–817.
- Jiang, G. S., and Peng, D. P., 1997, Weighted ENO schemes for Hamilton-Jacobi equations: CAM97-29, UCLA.
- Jiang, G. S., and Shu, C. W., 1996, Efficient implementation of weighted ENO schemes: *J. Comput. Phys.*, **126**, 202–228.
- Lions, P. L., 1982, Generalized solutions of Hamilton-Jacobi equations: Pitman Advanced Publishing Program.
- Liu, X. D., Osher, S. J., and Chan, T., 1994, Weighted essentially nonoscillatory schemes: *J. Comput. Phys.*, **115**, 200–212.
- Osher, S. J., and Sethian, J., 1988, Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations: *J. Comput. Phys.*, **79**, 12–49.
- Osher, S., and Shu, C. W., 1991, High-order essentially nonoscillatory schemes for Hamilton-Jacobi equations: *SIAM J. Num. Anal.*, **28**, 907–922.
- Popovici, A. M., and Sethian, J., 1997, Three-dimensional traveltime computation using the fast marching method: 67th SEG Annual Internat.Mtg. Expanded Abstracts, 1778–1781.
- Qian, J. L., and Symes, W. W., 1998, Paraxial eikonal equations and upwind finite difference schemes for traveltime and amplitude in inhomogeneous anisotropic media: Mathematical formulations: Annual Report, The Rice Inversion Project, Rice University.
- Qin, F., Luo, Y., Olsen, K. B., Cai, W., and Schuster, G. T., 1992, Finite difference solution of the eikonal equation along expanding wavefronts: *Geophysics*, **57**, 478–487.
- Reshef, M., and Kosloff, D., 1986, Migration of common shot gathers: *Geophysics*, **51**, 324–331.
- Schneider, W. A. J., Ranzinger, K., Balch, A., and Kruse, C., 1992, A dynamic programming approach to first arrival traveltime computation in media with arbitrarily distributed velocities: *Geophysics*, **57**, 39–50.
- Schneider, W. A. J., 1995, Robust and efficient upwind finite-difference traveltime calculations in three dimensions: *Geophysics*, **60**, 1108–1117.
- Shu, C., 1997, Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws: ICASE No.97-65, NASA Langley Research Center.
- Symes, W. W., 1995, Mathematics of reflection seismology: Annual Report, The Rice Inversion Project, Rice University.

- van Trier, J., and Symes, W. W., 1991, Upwind finite-difference calculation of traveltimes: *Geophysics*, **56**, 812–821.
- Vidale, J. E., and Houston, H., 1990, Rapid calculation of seismic amplitudes: *Geophysics*, **55**, 1504–1507.
- Vidale, J., 1988, Finite-difference calculation of travel times: *Bull., Seis. Soc. Am.*, **78**, 2062–2076.
- Zhang, L., 1993, Imaging by the wavefront propagation method: Ph.D. thesis, Stanford University, Stanford, CA94305.

Short Note

A second-order fast marching eikonal solver

James Rickett and Sergey Fomel¹

INTRODUCTION

The fast marching method (Sethian, 1996) is widely used for solving the eikonal equation in Cartesian coordinates. The method's principal advantages are: stability, computational efficiency, and algorithmic simplicity. Within geophysics, fast marching traveltimes (Popovici and Sethian, 1997) may be used for 3-D depth migration or velocity analysis.

Unfortunately, first-order implementations lead to inaccuracies in computed traveltimes, which may lead to poor image focusing for migration applications. In addition, first-order traveltimes are not accurate enough for reliable amplitude calculations. This has led to the development of the fast marching method on non-Cartesian (Alkhalifah and Fomel, 1997; Sun and Fomel, 1998), and even unstructured (Fomel, 1997) grids. These non-Cartesian formulations reduce inaccuracies, while retaining the fast marching method's characteristic stability and efficiency. Unfortunately, the cost is the loss of algorithmic simplicity.

We implement a second-order fast marching eikonal solver, which reduces inaccuracies while retaining stability, efficiency *and* simplicity.

FAST MARCHING AND THE EIKONAL EQUATION

Under a high frequency approximation, propagating wavefronts may be described by the eikonal equation,

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

where t is the traveltimes, s is the slowness, and x , y and z represent the spatial Cartesian coordinates.

The fast marching method solves equation (1) by directly mimicking the advancing wavefront. Every point on the computational grid is classified into three groups: points behind the wavefront, whose traveltimes are known and fixed; points on the wavefront, whose traveltimes

¹email: james@sep.stanford.edu, sergey@sep.stanford.edu

have been calculated, but are not yet fixed; and points ahead of the wavefront. The algorithm then proceeds as follows:

1. Choose the point on the wavefront with the smallest traveltime.
2. Fix this traveltime.
3. Advance the wavefront, so that this point is behind it, and adjacent points are either on the wavefront or behind it.
4. Update traveltimes for adjacent points on the wavefront by solving equation (1) numerically.
5. Repeat until every point is behind the wavefront.

The update procedure (step 4.) requires the solution of the following quadratic equation for t ,

$$\begin{aligned} & \max(D_{ijk}^{-x}t, 0)^2 + \min(D_{ijk}^{+x}t, 0)^2 + \\ & \max(D_{ijk}^{-y}t, 0)^2 + \min(D_{ijk}^{+y}t, 0)^2 + \\ & \max(D_{ijk}^{-z}t, 0)^2 + \min(D_{ijk}^{+z}t, 0)^2 = s_{ijk} \end{aligned} \quad (2)$$

where D_{ijk}^{-x} is a backward x difference operator at grid point, ijk , D_{ijk}^{+x} is a forward x operator, and finite-difference operators in y and z are defined similarly. The roots of the quadratic equation, $at^2 + bt + c = 0$, can be calculated explicitly as

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (3)$$

Solving equation (2) amounts to accumulating coefficients a , b and c from its non-zero terms, and evaluating t with equation (3).

If we choose a two-point finite-difference operator, such as

$$D_{ijk}^{-x}t = \frac{t_{ijk} - t_{(i-1)jk}}{\Delta x} \quad (4)$$

$$\text{then } (D_{ijk}^{-x}t)^2 = \alpha t_{ijk}^2 + \beta t_{ijk} + \gamma \quad (5)$$

where $\alpha = \frac{1}{\Delta x^2}$, $\beta = -2t_{(i-1)jk}\alpha$ and $\gamma = t_{(i-1)jk}^2\alpha$. Coefficients a , b and c can now be calculated from $a = \sum_l \alpha_l$, $b = \sum_l \beta_l$, and $c = \sum_l \gamma_l - s^2$, where the summation index, l , refers to the six terms in equation (2) subject to the various min/max conditions.

This two-point stencil, however, is only accurate to first-order. If instead we choose a suitable three-point finite-difference stencil, we may expect the method to have second-order accuracy. For example, the second-order upwind stencil,

$$D_{ijk}^{-x}t = \frac{3t_{ijk} - 4t_{(i-1)jk} + t_{(i-2)jk}}{2\Delta x} \quad (6)$$

$$\text{gives } (D_{ijk}^{-x}t)^2 = \alpha' t_{ijk}^2 + \beta' t_{ijk} + \gamma' \quad (7)$$

$$\text{where this time } \alpha' = \frac{9}{4\Delta x^2},$$

$$\beta' = \frac{-3(4t_{(i-1)jk} - t_{(i-2)jk})}{2\Delta x^2} = -2\alpha' t'_{(i-1)jk},$$

$$\gamma' = \frac{(4t_{(i-1)jk} - t_{(i-2)jk})^2}{4\Delta x^2} = \alpha' t'^2_{(i-1)jk},$$

$$\text{and } t'_{(i-1)jk} = \frac{1}{3}(4t_{(i-1)jk} - t_{(i-2)jk}).$$

Coefficients, a , b and c can be accumulated from α' , β' and γ' as before, and if the traveltimes, $t_{(i-2)jk}$ is not available, first-order values may be substituted.

ACCURACY

Figure 1 shows traveltimes contour maps computed with the first and second-order fast marching methods on a sparse (20×20) grid. The large errors for waves propagating at 45° to the grid are visibly reduced by the second-order formulation.

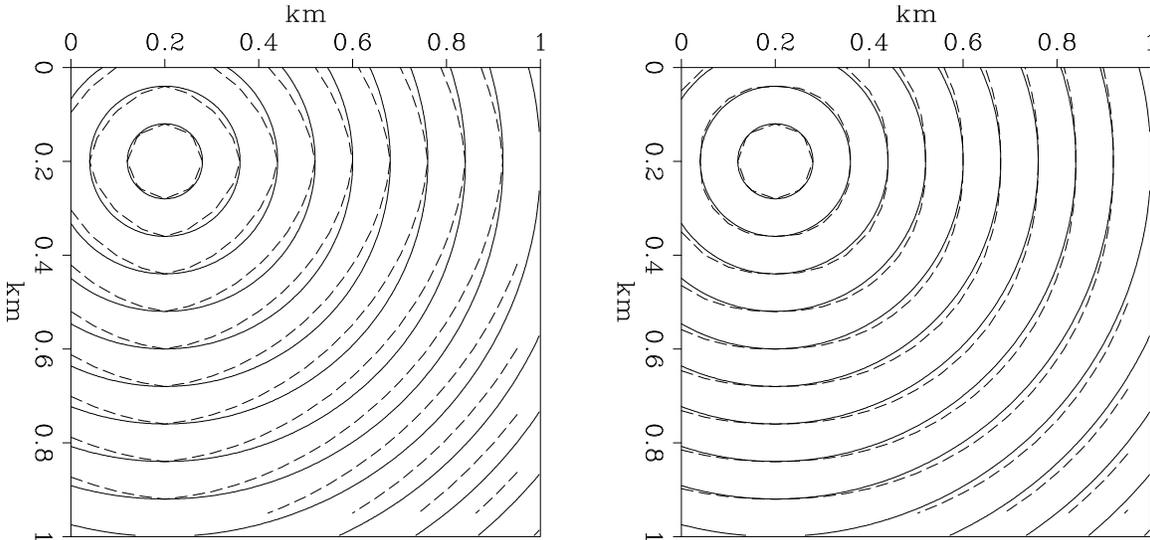


Figure 1: Traveltimes contours in a constant velocity medium. The solid line shows the exact result. The dashed line shows the first-order (left panel) and second-order (right panel) fast marching results, calculated on a 20×20 grid. james2-circles [ER]

Figure 2 shows the average error as a function of grid spacing for the first and second-order solvers. Not only is the second-order formulation more accurate at large grid spacing, but its accuracy increases more rapidly as grid spacing decreases. Theory predicts the log – log plots of average error against grid spacing to be a linear function with gradient of one for first-order methods, and two for second order methods. In practice, the fast marching results come

very close to these criteria up to the limits of machine precision. Figure 2 demonstrates the superiority of the second-order fast marching formulation.

It is worth noting, at this point, that special treatment is required at the source location, since the singularity in wavefront curvature will cause numerical errors to propagate into the traveltimes solution. We surround the source with a constant velocity box, within which we calculate traveltimes by ray-tracing. Errors are inversely proportional to the radius of this box. Therefore, if the radius of the box decrease with grid spacing, errors will increase linearly, reducing the accuracy of the method to first-order. For full second-order accuracy, the box size should be independent of grid spacing.

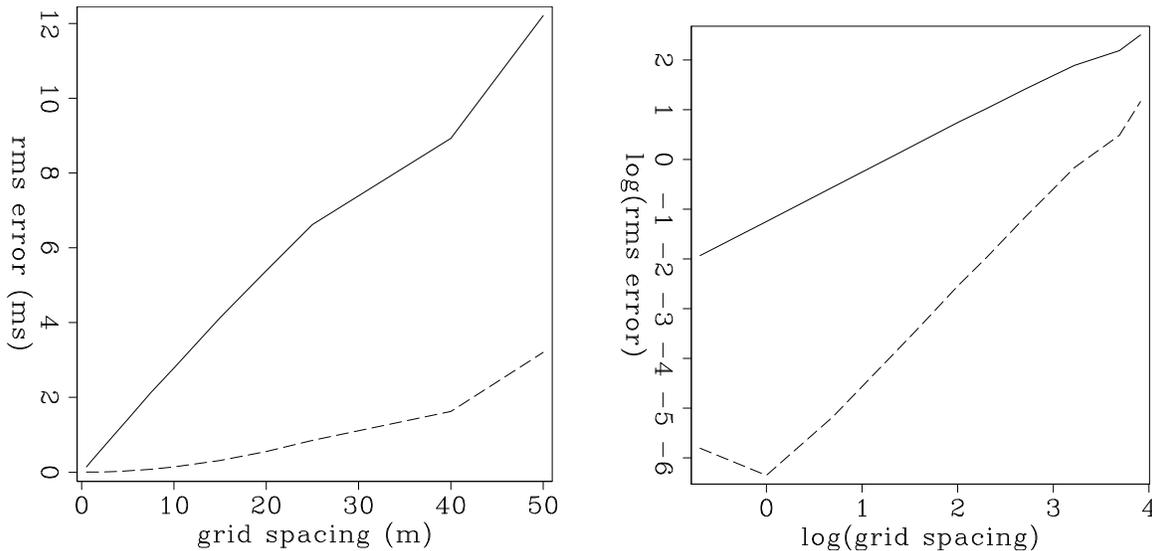


Figure 2: Average error against grid spacing for a constant velocity model. The solid line corresponds to the first-order eikonal solver, and the dashed line corresponds to the second-order solver. The left panel has linear axes, whereas the right panel is a log – log plot. james2-error [ER]

COMPUTATIONAL COST

The leading term in the computational cost of the fast marching algorithm comes from the first step: choosing the point on the wavefront with the smallest traveltimes. Consequently, the cost should not depend strongly on the order of the finite-difference stencil, but rather the sort algorithm used. Heap sorting has a cost of $O(\log N)$, and so in principle, with this algorithm, the fast marching method has a cost of $O(N \log N)$.

The left panel of Figure 4 shows a plot of CPU time against N for the same models as Figure 2. The time shown is elapsed (wall clock) time on a 300 MHz Pentium II. For the largest model computed here, the second-order code takes 11% longer to run than the first-order code, and this percentage decreases as N increases.

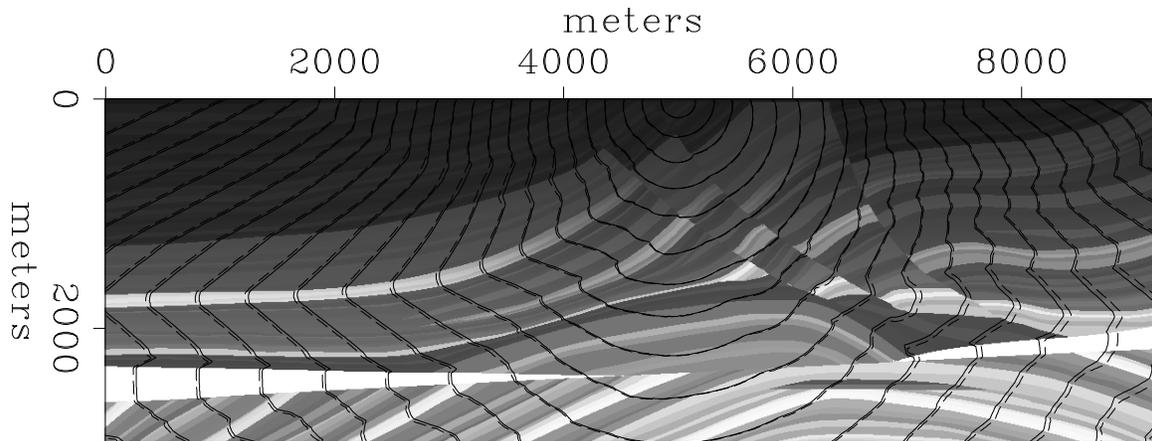


Figure 3: Traveltime contours calculated through the Marmousi velocity model sampled at 4 m. Solid line shows first-order results, and dashed line shows second-order results.

`james2-marmousi` [ER]

Because $\log N$ grows slowly compared to N , the plot of CPU time against N is dominated by the linear term. The right panel in Figure 4 addresses this issue by showing CPU time divided by N versus N . On this graph, the $\log N$ behaviour is clearly visible.

CONCLUSIONS

We have shown that a second-order implementation of the fast marching eikonal solver produces traveltimes with a much higher accuracy than the first-order implementation. What is more, the additional accuracy is achieved at only a marginal increase in cost.

This second-order implementation should become the standard method for computing first-arrival traveltimes within SEP.

REFERENCES

- Alkhalifah, T., and Fomel, S., 1997, Implementing the fast marching eikonal solver: Spherical versus cartesian coordinates: SEP-95, 149–171.
- Fomel, S., 1997, A variational formulation of the fast marching eikonal solver: SEP-95, 127–147.
- Popovici, A. M., and Sethian, J. A., 1997, Three-dimensional traveltime computation using the fast marching method: 67th Ann. Internat. Meeting, Soc. Expl. Geophys., 1778–1781.
- Sethian, J. A., 1996, Level set methods: Evolving interfaces in geometry, fluid mechanics, computer vision, and materials science: Cambridge University Press.

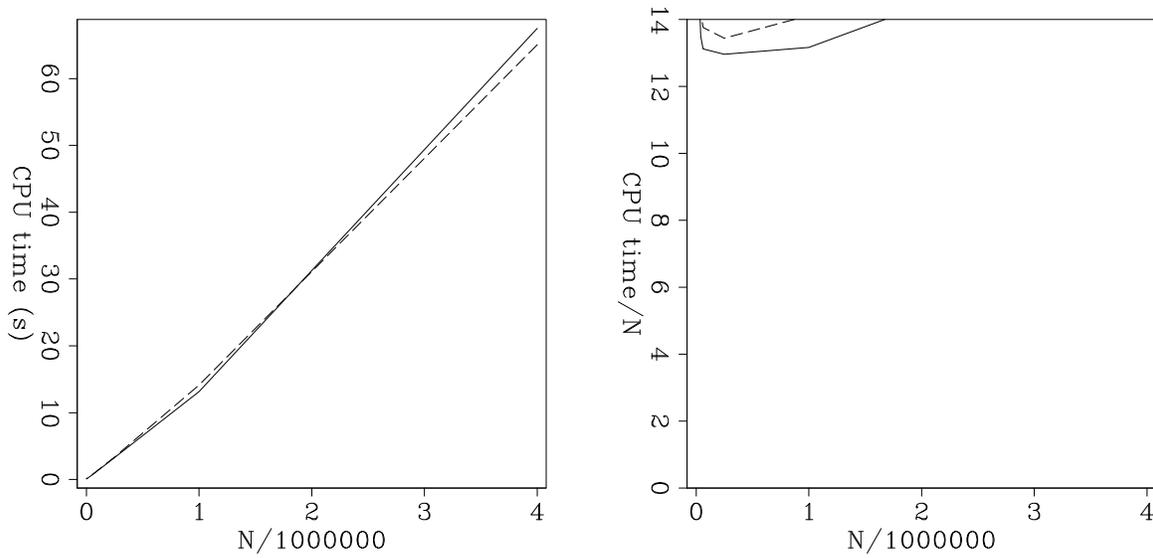


Figure 4: Elapsed CPU time vs. the number of grid points, N , for first-order (solid line) and second order (dashed line) eikonal solvers. Left panel shows CPU time vs N . Right panel shows CPU time/ N vs N . james2-times [CR]

Sun, Y., and Fomel, S., 1998, Fast-marching eikonal solver in the tetragonal coordinates: SEP-97, 241–250.

Robust and stable velocity analysis using the Huber function

Antoine Guitton and William W. Symes¹

ABSTRACT

The *Huber function* is one of several robust error measures which interpolates between smooth (l^2) treatment of small residuals and robust (l^1) treatment of large residuals. Since the Huber function is differentiable, it may be minimized reliably with a standard gradient-based optimizer. Tests with a linear inverse problem for velocity analysis, using both synthetic and field data, suggest that (1) the Huber function gives far more robust model estimates than does least squares, (2) its minimization using a standard quasi-Newton method is comparable in computational cost to least squares estimation using conjugate gradient iteration, and (3) the result of Huber data fitting is stable over a wide range of choices for $l^2 \rightarrow l^1$ threshold and total number of quasi-Newton steps.

INTRODUCTION

Robust error measures such as the l^1 norm have found a number of uses in geophysics. As measures of data misfit, they show considerably less sensitivity to large measurement errors than does the mean square (l^2) measure. Since geophysical inverse problems are generally ill-posed, relatively noise insensitive misfit measures can yield far more stable estimates of Earth parameters than does the mean square measure (Claerbout and Muir, 1973; Taylor et al., 1979; Chapman and Barrodale, 1983; Scales and Gersztenkorn, 1988; Scales et al., 1988). This insensitivity to large noise has a statistical interpretation: robust measures are related to long-tailed density functions in the same way that the mean square is related to the (short-tailed) Gaussian (Tarantola, 1987).

A simple choice of robust measure is the l^1 norm: denoting the residual (misfit) components by r_i , $i = 1, \dots, N$, l^1 norm of the residual vector is $\sum_{i=1}^N |r_i|$. This function is not smooth: it is singular where any residual component vanishes. As a result, numerical minimization is difficult. Various approaches based for example on a linear programming viewpoint (Barrodale and Roberts, 1980) or iterative smoothing (Scales et al., 1988), have been used with success but require considerable tuning. Moreover, the singularity implies that small residuals are “taken as seriously” as large residuals, which may not be appropriate in all circumstances.

These drawbacks of the l^1 norm have led to various proposals which combine robust treatment of large residuals with Gaussian treatment of small residuals. In the work reported here,

¹email: antoine@sep.stanford.edu, symes@caam.rice.edu

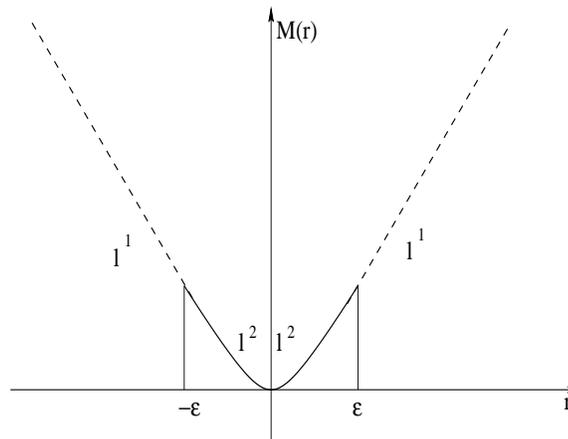
we use a hybrid l^1 - l^2 error measure proposed by Huber (Huber, 1973):

$$M_\epsilon(r) = \begin{cases} \frac{r^2}{2\epsilon}, & 0 \leq |r| \leq \epsilon \\ |r| - \frac{\epsilon}{2}, & \epsilon < |r| \end{cases}$$

We will call $\sum_{i=1}^N M_\epsilon(r_i)$ the *Huber misfit function*, or Huber function for short (Figure 1). Note that the Huber function is smooth near zero residual, and weights small residuals by mean square. It is reasonable to suppose that the Huber function is easier to minimize than l^1 while still robust against large residuals.

This paper describes the application of the Huber misfit function to velocity analysis. Estimation of RMS velocity (or slowness) can be posed as a linear inverse problem through the *velocity transform* described in the next section. Definition of the misfit *via* the Huber function (or any other robust error measure) results in a nonlinear optimization problem for the velocity model. This nonlinearity would seem to compare unfavorably with the least squares (l^2) treatment of the same problem, which leads to a linear system (the normal equation) and so can be solved by efficient iterative methods such as conjugate gradient. We show that use of an appropriate nonlinear optimization method gives a Huber-based solution with comparable efficiency to that of conjugate gradient least squares solution. Thus the noise rejection properties of the Huber misfit function come at no appreciable premium in computational effort. In the work reported here we have used a version of the Limited Memory BFGS algorithm

Figure 1: Error measure proposed by Huber (Huber, 1973). The upper part above ϵ is the l^1 norm while the lower part is the l^2 norm. antoine1-huber
[NR]



(Nocedal, 1980) as implemented in the Hilbert Class Library (Gockenbach et al., 1999). Other nonlinear iterative optimizers could be used; we have solved the same examples with nonlinear conjugate gradient methods (Fletcher, 1980) and obtained comparable results. We note that specially adapted Huber minimizers have been suggested (Ekblom and Madsen, 1989). One of our questions in beginning this work was whether a standard quasi-Newton method, as opposed to a special solver, would perform satisfactorily in Huber estimation.

The second section of the paper explains the velocity transform and formulates a linear inverse problem for velocity analysis. The third and fourth sections present synthetic and field data examples.

APPLICATION TO VELOCITY ESTIMATION

The *velocity domain* representation of seismic data is an alternative to the standard CMP presentation. Transformation of CMP data into the velocity domain (producing a velocity *model* or *panel* of the data) exhibits clearly the moveout inherent in the data and therefore, forms a convenient basis for velocity analysis as a linear inverse problem. The velocity transform \mathbf{A} from the model space (velocity domain) into the data space (CMP gathers) stretches the velocities back in the offset plane (superposition of hyperbolas) whereas the adjoint operation (\mathbf{A}') squeezes the data (summation over hyperbolas):

$$\mathbf{A} = \mathbf{H}\mathbf{S},$$

with

$$\mathbf{S}\mathbf{m}(t, x) = \sum_s \frac{t}{\tau} w(s, x, \tau) m(\tau, s) \Big|_{\tau = \sqrt{t^2 - s^2 x^2}}$$

$$\mathbf{A}' = \mathbf{S}'\mathbf{H}',$$

with

$$\mathbf{S}'d(\tau, s) = \sum_x w(s, x, \tau) d(t, x) \Big|_{t = \sqrt{\tau^2 + s^2 x^2}}$$

where $w(s, x, \tau)$ is a weighting function, \mathbf{H} is a filter that we define later. \mathbf{A} is related to the velocity stack as defined by Taner and Koehler (1969).

The problem is: given a CMP gather can we find a velocity panel which synthesizes it *via* \mathbf{A} ? In equations, given data \mathbf{d} , we want to solve for model \mathbf{m} :

$$\mathbf{A}\mathbf{m} \approx \mathbf{d}.$$

A simple way to solve this problem is to find a model \mathbf{m} that minimizes the *mean square* misfit

$$(\mathbf{A}\mathbf{m} - \mathbf{d})'(\mathbf{A}\mathbf{m} - \mathbf{d}).$$

This optimization problem is equivalent to the linear system (“normal equations”)

$$\mathbf{A}'\mathbf{A}\mathbf{m} = \mathbf{A}'\mathbf{d}.$$

This system is easy to solve if $\mathbf{A}'\mathbf{A} \approx \mathbf{I}$, i.e if \mathbf{A} is close to unitary: then $\mathbf{m} = \mathbf{A}'\mathbf{d}$. In general, \mathbf{A} is far from an unitary operator for many reasons. However, the choice of a weighting function compensates to some extent for geometrical spreading and other effects (Claerbout and Black, 1997):

$$w(s, x, \tau) = \frac{1}{\sqrt{(\tau^2 + s^2 x^2)^{1/2}}} \frac{\tau}{\sqrt{\tau^2 + s^2 x^2}} \sqrt{x s}.$$

The summation in the velocity space boosts low frequencies. Claerbout and Black (1997) suggest that a good choice of filter \mathbf{H} is a half derivative operator ($\sqrt{i\omega}$). These choices for \mathbf{H} and $w(s, x, \tau)$ bring \mathbf{A} closer to being an unitary operator.

Since the data is noisy, the modeling operator is not unitary and the numbers of equations and unknowns may be large, an iterative data-fitting approach seems reasonable:

$$\min_{\mathbf{m}}(E(\mathbf{A}\mathbf{m} - \mathbf{d}))$$

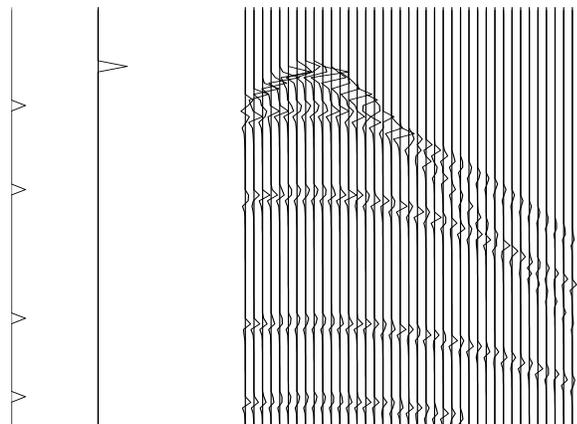
where \mathbf{m} is the model, \mathbf{d} , the data we want to fit, \mathbf{A} the modeling operator, and E a misfit measurement function we have to choose. We have already presented one possibility, namely that E is the l^2 norm (least squares inversion). A convenient iterative method in this case is to solve the normal equation using conjugate gradient iteration. We refer to this approach as “CG” or “ l^2 ”. An alternative approach is to take for E the Huber function introduced in the first section. With this Huber misfit measure, the velocity transform inverse problem is no longer equivalent to a linear system. We choose to solve it using a general-purpose nonlinear optimizer, as mentioned above, rather than one of several special-purpose methods invented for this type of problem. We refer to this approach as “Huber” or the “Huber solver”.

The next two parts of this paper compare the performance of the CG algorithm to Huber in the velocity analysis application.

SYNTHETIC DATA TESTS

To compare the Huber function with the least squares measure, we generate a synthetic CMP gather (Figure 2) that we perturb by introducing: (1) missing traces, (2) a low velocity aliased plane wave, and (3) some sparsely distributed spiky noisy events. These data sets constitute the input for the iterative schemes ($\epsilon = 0.01$ for each result). The panels display the model space on the left (after 20 iterations), and the data space on the right. The bottom right panels show the modeling of the last velocity result. All these results (Figures 4, 6, 8) prove the following: outcome of the Huber solver is insensitive to spiky events, like a pure l^1 norm misfit function. The outcome of the missing data problem was probably less predictable, but again, Huber copes more easily with the inconsistency introduced in the data.

Figure 2: Left: ideal velocity panel. Right: data model.
antoine1-datamodel [CR]



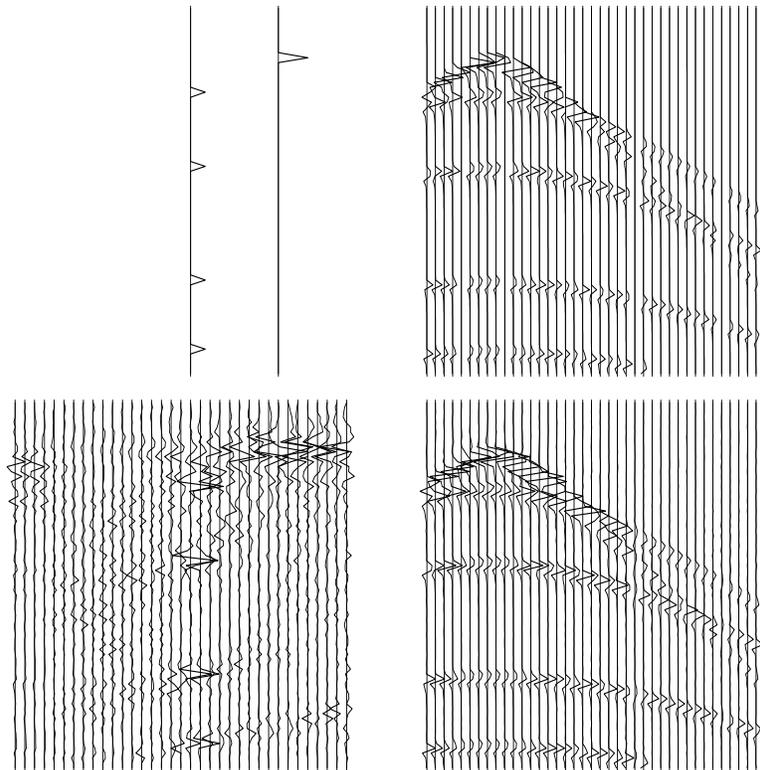


Figure 3: CG result with missing data. antoine1-vel-miss2g [CR]

FIELD DATA EXAMPLES

In this section we compare the CG to the Huber function for two CMP gathers. We divide this section in three: first, we show the inverted velocity panels and modeled data obtained for both the CG and Huber solver. The second and third section address the stability problems and computational efficiency.

Velocity analysis on field data

Figure 9 displays different interesting features: the first CMP (A) shows low velocity events (ground roll or guided waves) and time shifts near offset 2km while the second example (B) shows bad traces with high amplitudes. The same clip has been applied to each figure. Note that the offsets are probably wrong since we have reflected waves traveling at about 10 km/s for Input A! The velocities are computed after 5 iterations only, and $\epsilon = 0.001$ for both gathers. Figure 10 shows the velocity analysis using Huber and CG on a first data set. It appears that the CG result has artifacts at times above 1.5 sec. We see some horizontal stripes that make reliable interpretation difficult. In contrast, the Huber result displays a focused velocity corridor. Some low frequency events do appear in the upper right part of the panel but they do not interfere with the main fairway. It is interesting to notice that Huber separates the low-frequency low-velocity noise from the signal whereas the l^2 measure spreads it along the velocity axis. If

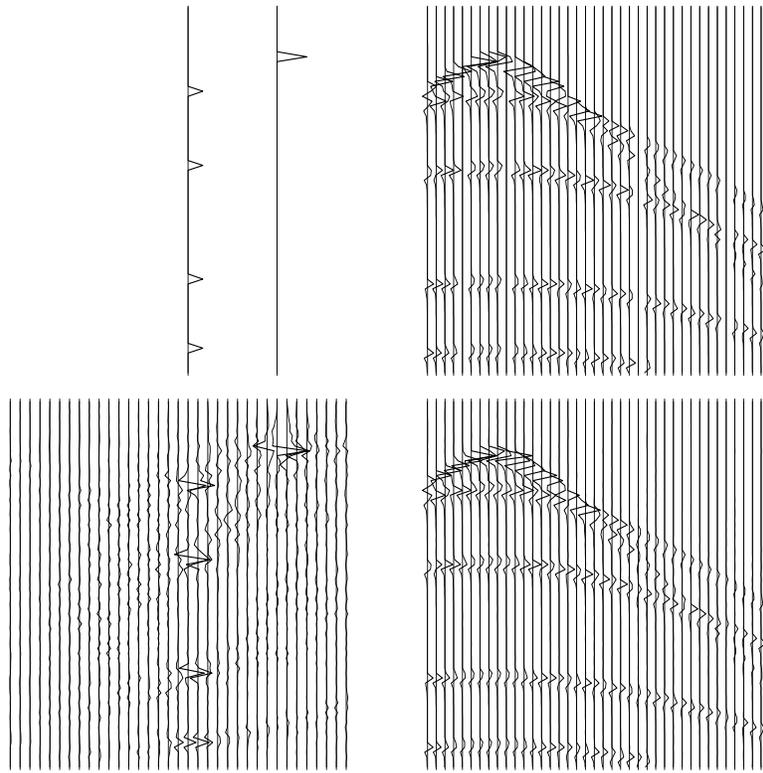


Figure 4: Huber result with missing data. `antoine1-vel-miss3h` [CR]

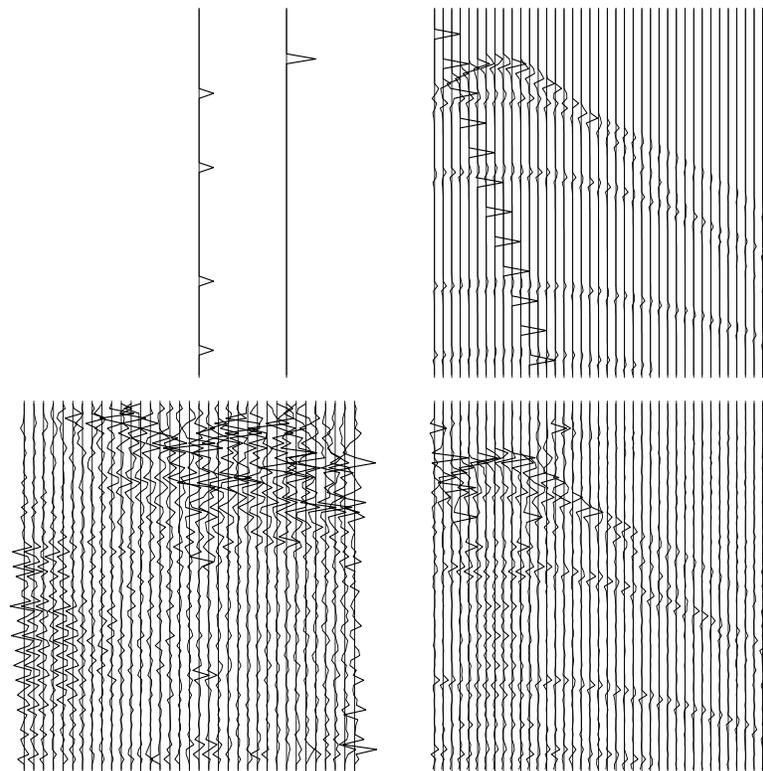


Figure 5: CG result with a slow plane wave. `antoine1-vel-surf1g` [CR]

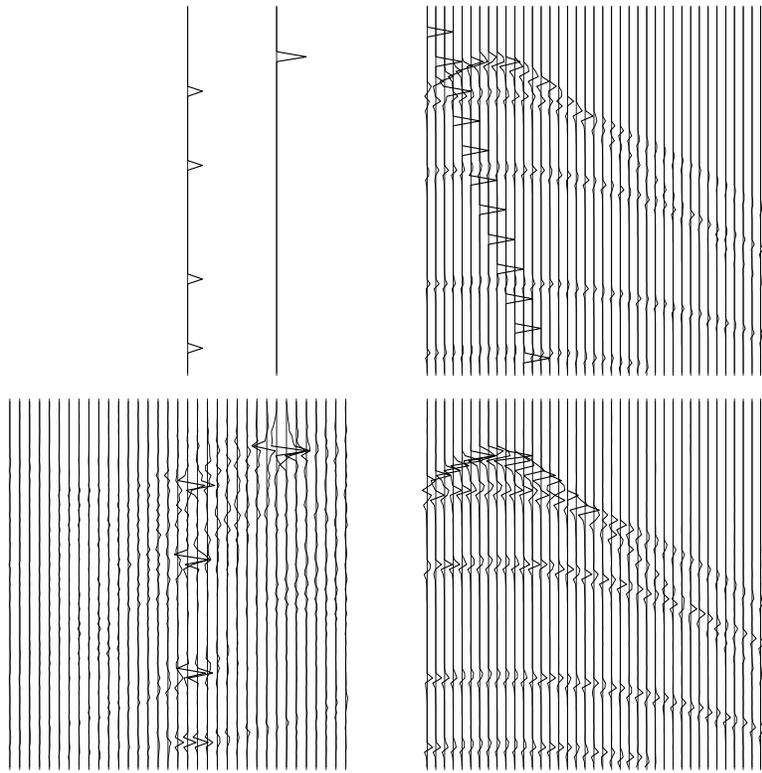


Figure 6: Huber result with a slow plane wave. `antoine1-vel-surf2h` [CR]

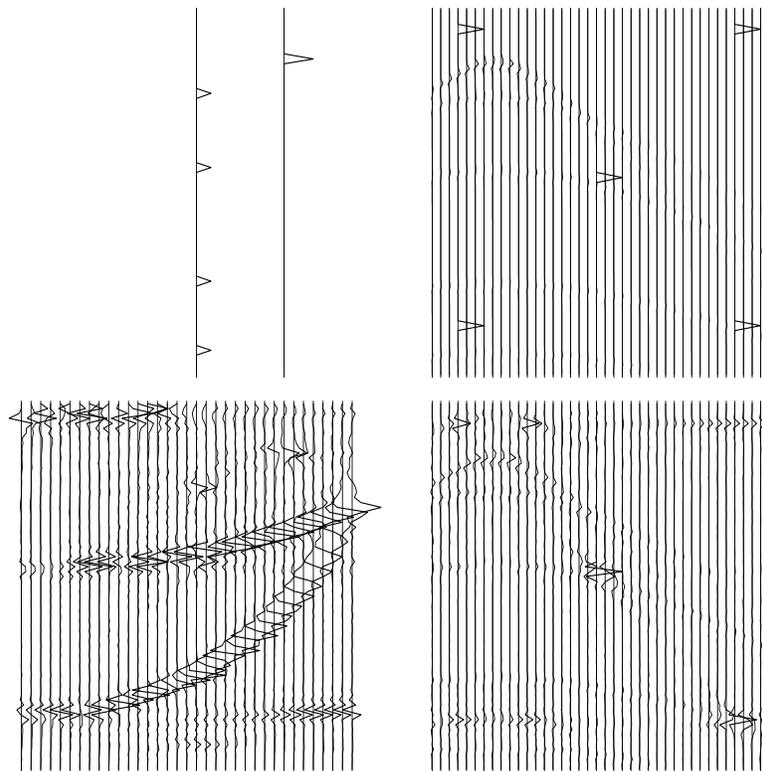


Figure 7: CG result with spiky events. `antoine1-vel-spiky1g` [CR]

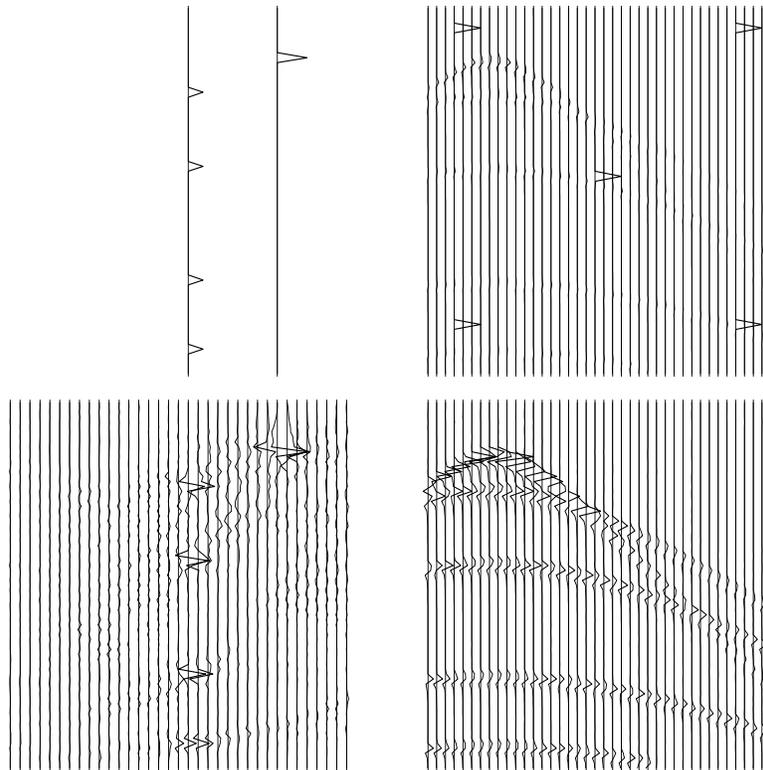


Figure 8: Huber result with spiky events. antoine1-vel-spiky2h [CR]

we now model those results back into the data-space, we obtain Figure 11. We notice that CG does not do a good job in estimating the upper part. Furthermore, some high frequency noise appears. The Huber result however looks close to the original data. In particular, the upper part is well estimated and no harmful artifacts are visible. The anomalous high amplitude first trace does not affect the final results either.

Let's now look at the results for Input *B*. Figure 12 show the strength of the Huber solver compared to CG. The CG velocity panel shows horizontal stripes in the velocity scan making any reliable picking quite impossible. The Huber velocity panel displays a focused bended corridor with low noise amplitude. If we now model those results back into the data-space, we get Figure 13. Those sections support the same observations as previous, and as expected, Huber allows to recover more accurately the original data. These results are quite encouraging and give us a flavor of what could be accomplished with the Huber norm. The same conclusion applies on the field data as on the synthetics: the Huber function is robust.

Is Huber a stable solver ?

For Huber to be a serious competitor to the l^2 norm and the CG, it needs to be stable with respect to the number of iterations and to the Huber threshold, ϵ . This is what we investigate in this section. Figure 14 displays two velocity panels using 5 and 70 CG iterations on input *A*.

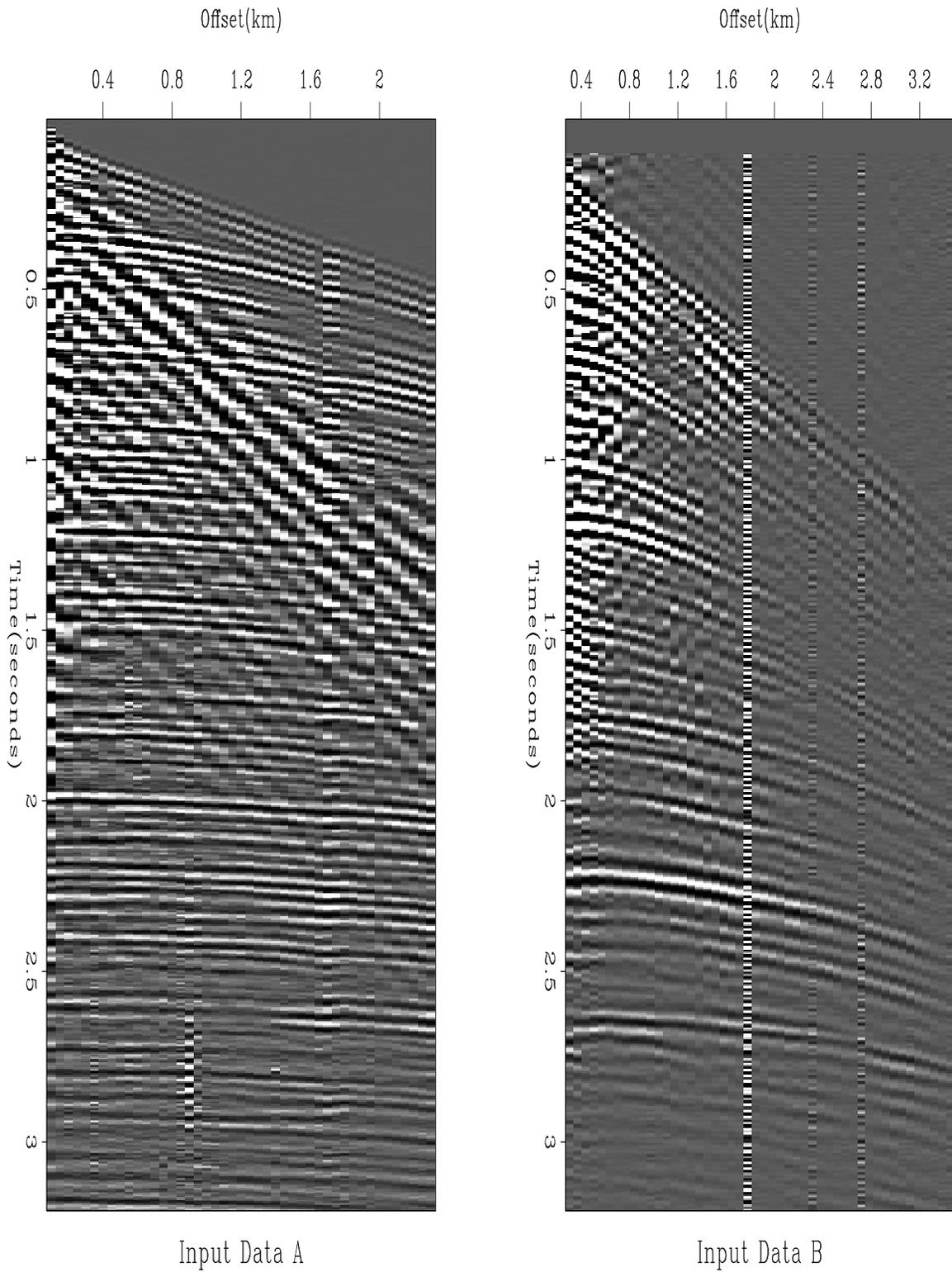


Figure 9: Input data- 2 CMP gathers. `antoine1-datamodelreal` [NR]

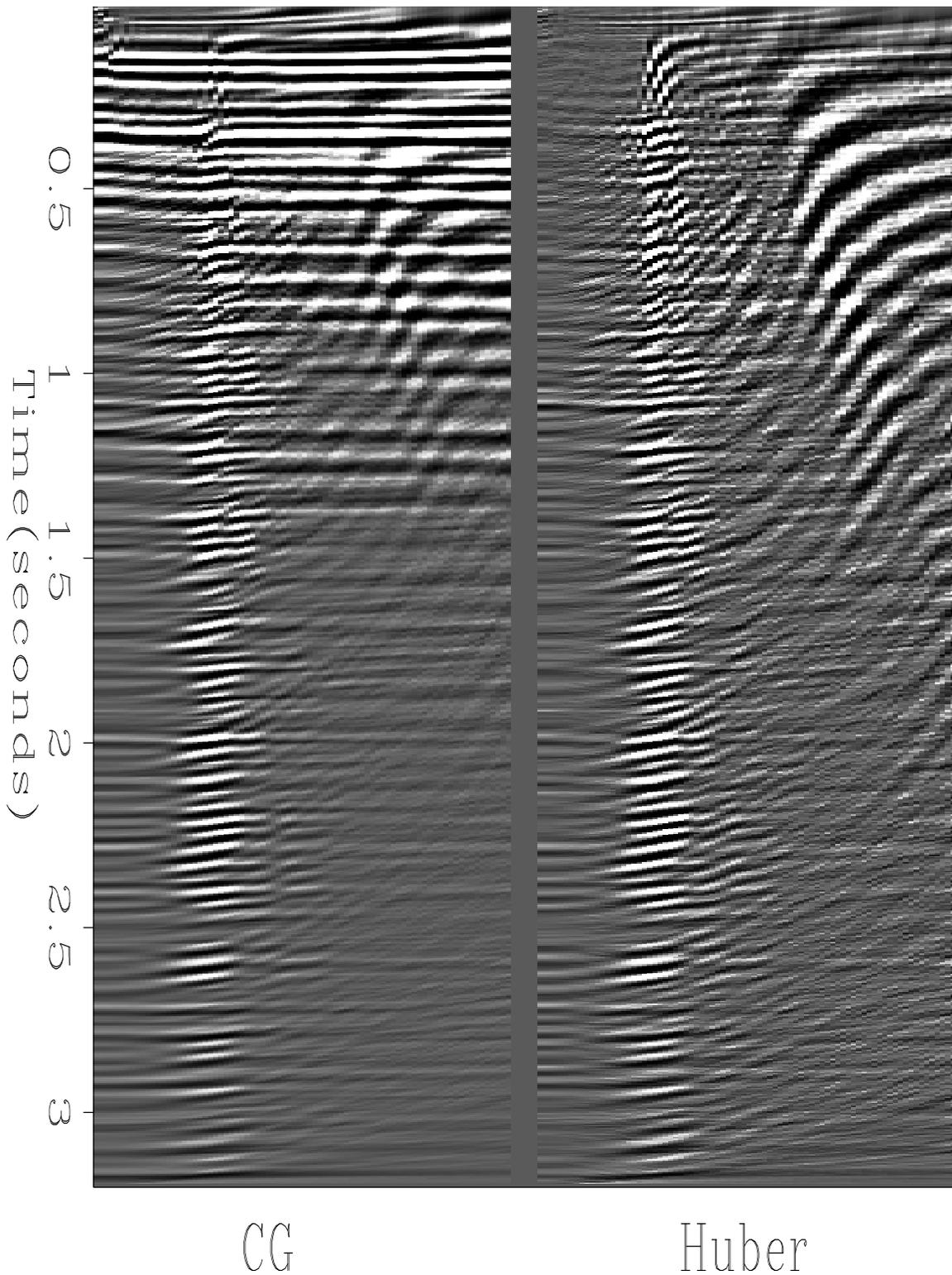


Figure 10: Velocity panel using CG and Huber solver: the left panel shows some stability problems above 1.5 sec. The Huber solver result shows a focused velocity corridor on the left and some artifacts on the right. Those artifacts are well separated from the fairway, however. [antoine1-modelwz08](#) [CR]

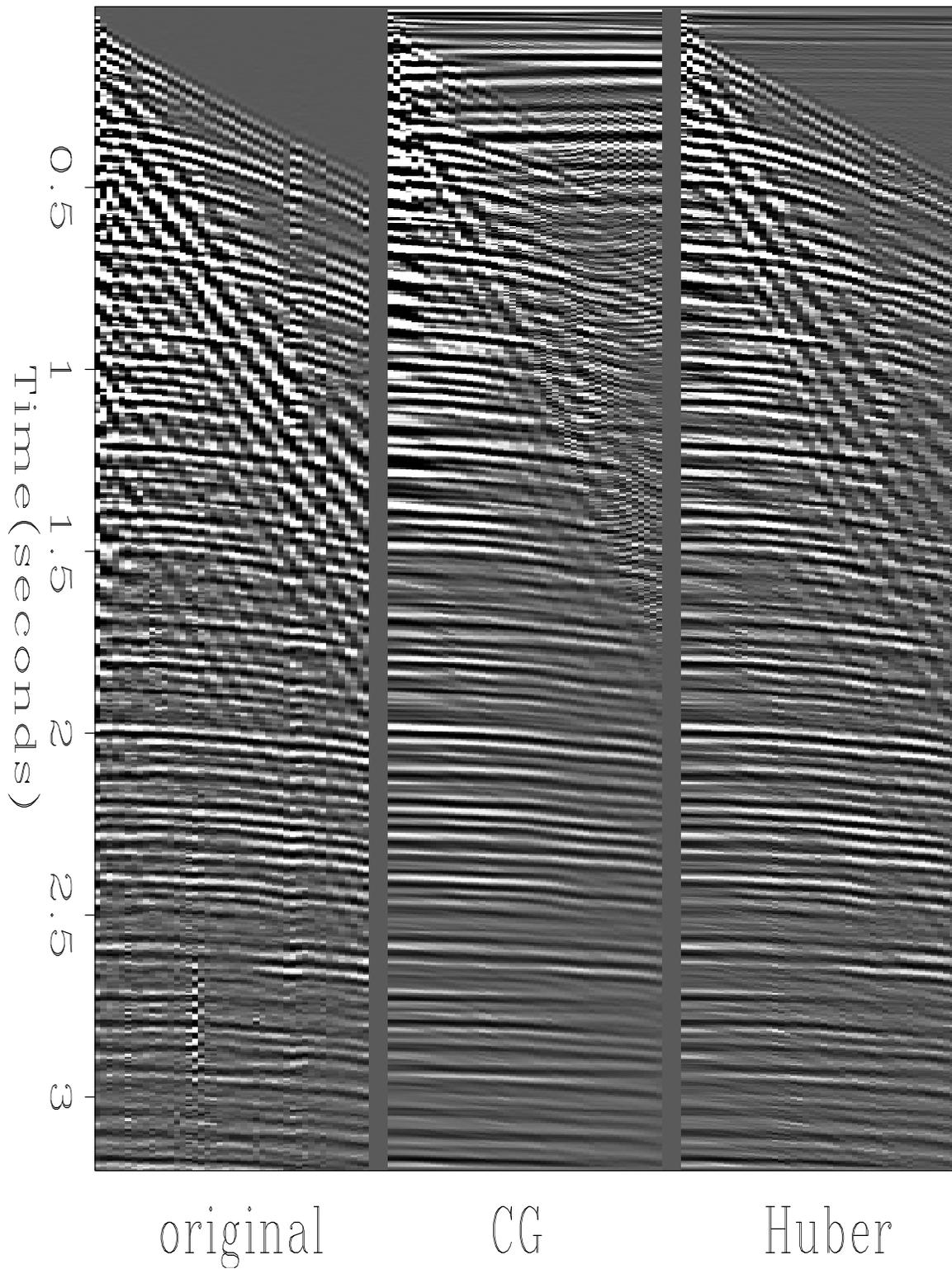


Figure 11: Data modeling after iteratively improved velocity model. antoine1-datawz08
[CR]

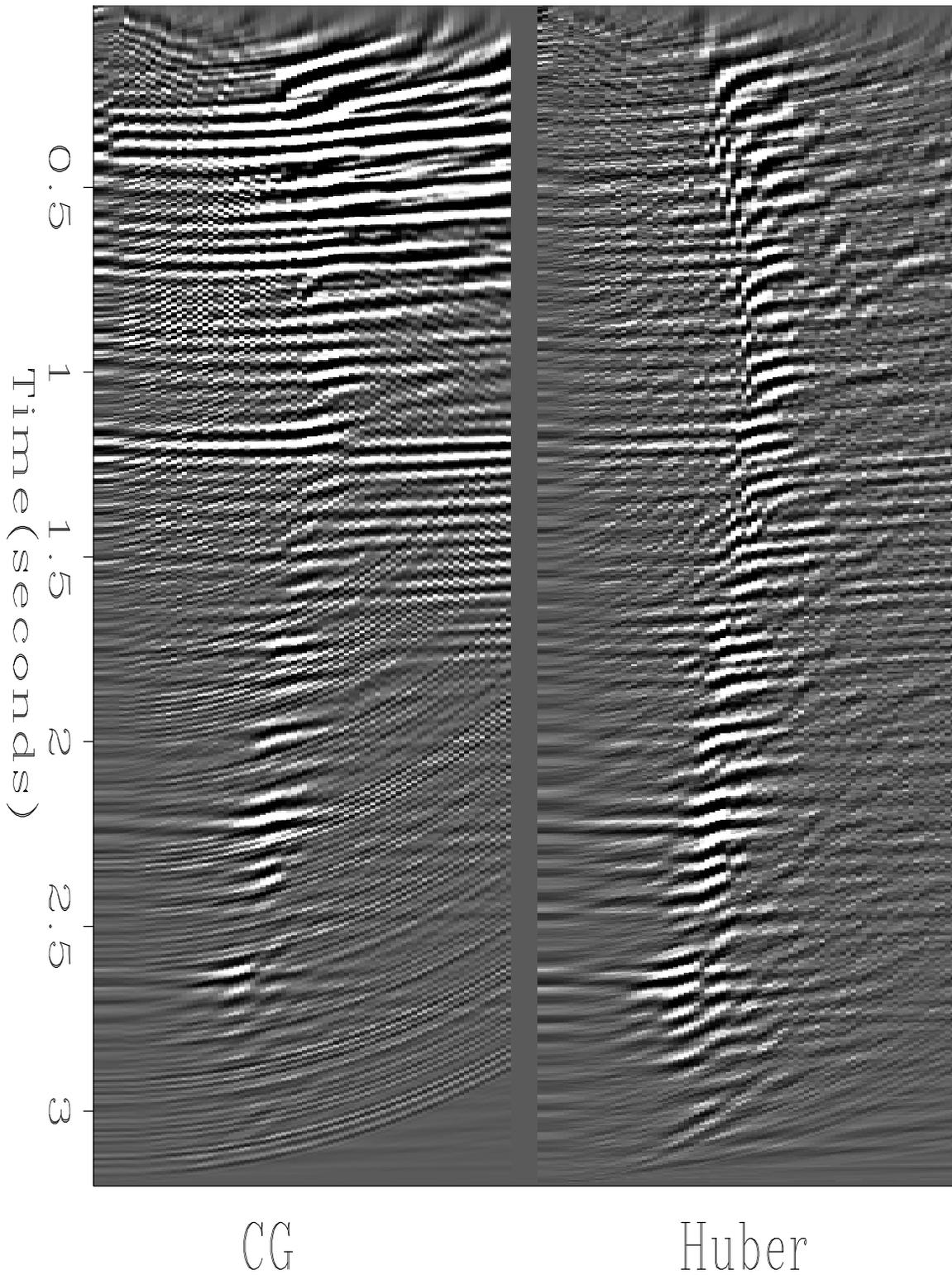


Figure 12: Velocity model using CG and Huber solver. [antoine1-modelwz11](#) [CR]

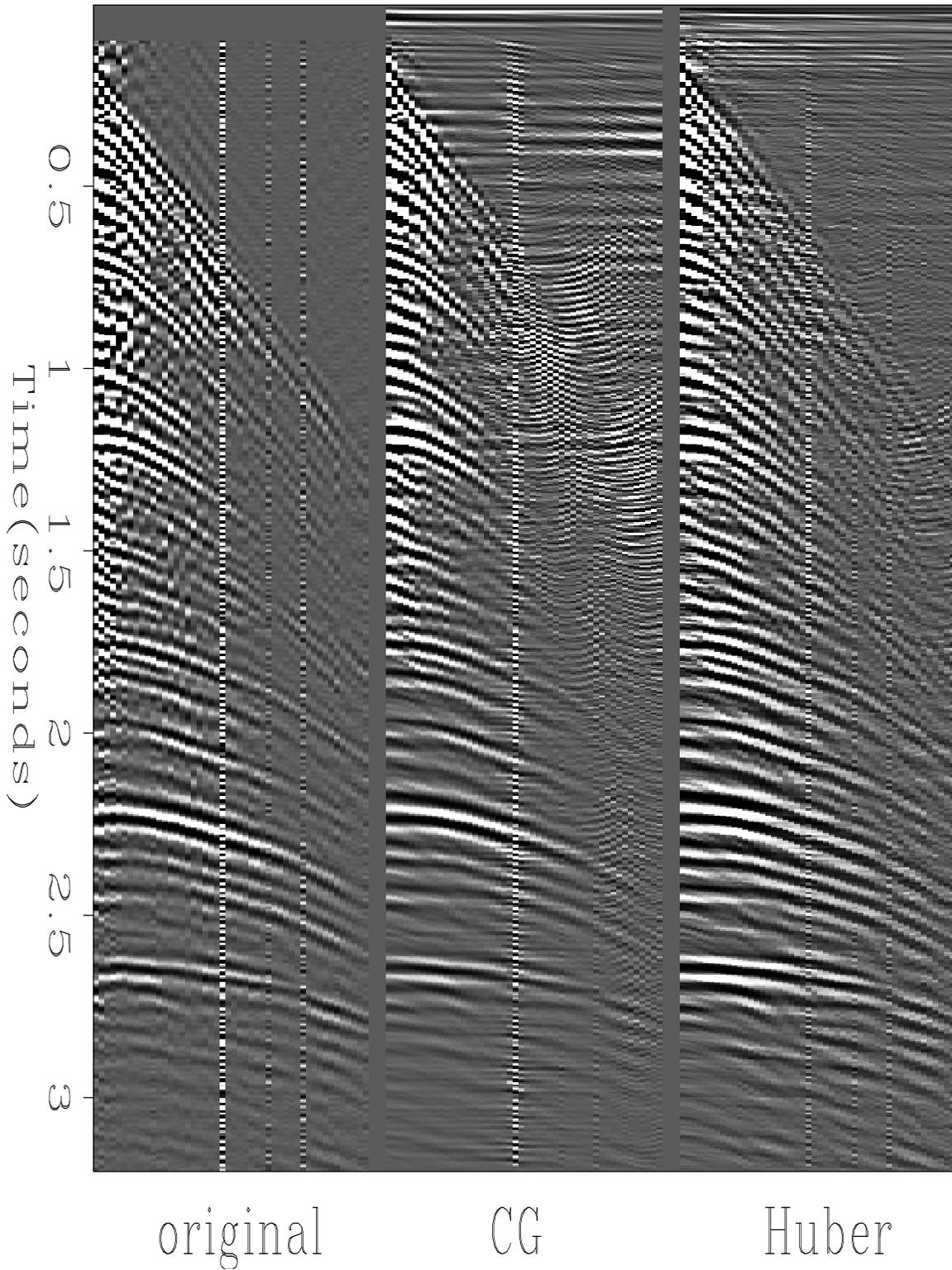


Figure 13: Data modeling after iteratively improved velocity model. antoine1-datawz11
[CR]

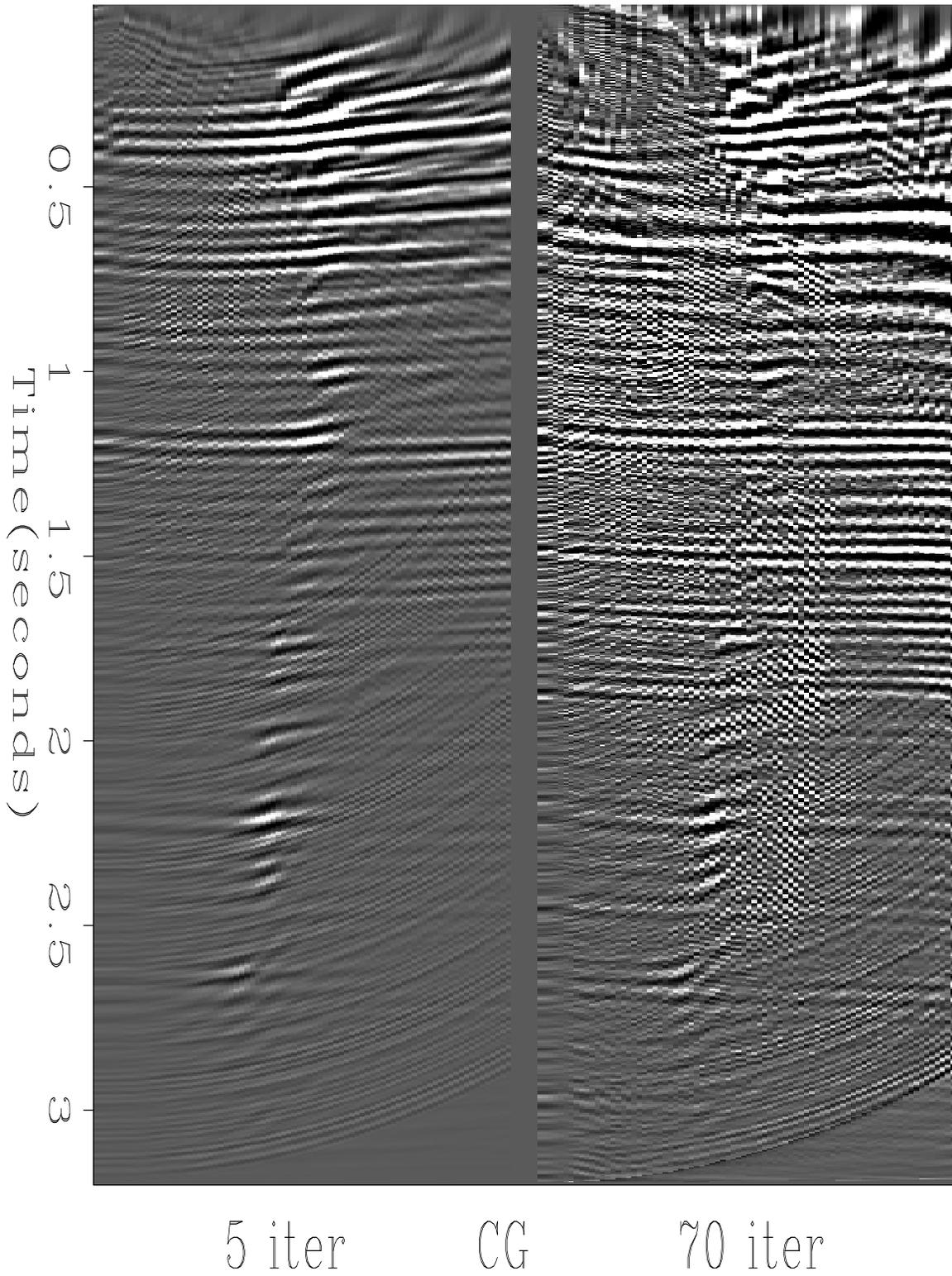


Figure 14: Input *B*: velocity panel using 5 and 70 CG iterations. antoine1-compmodel11g
[NR]

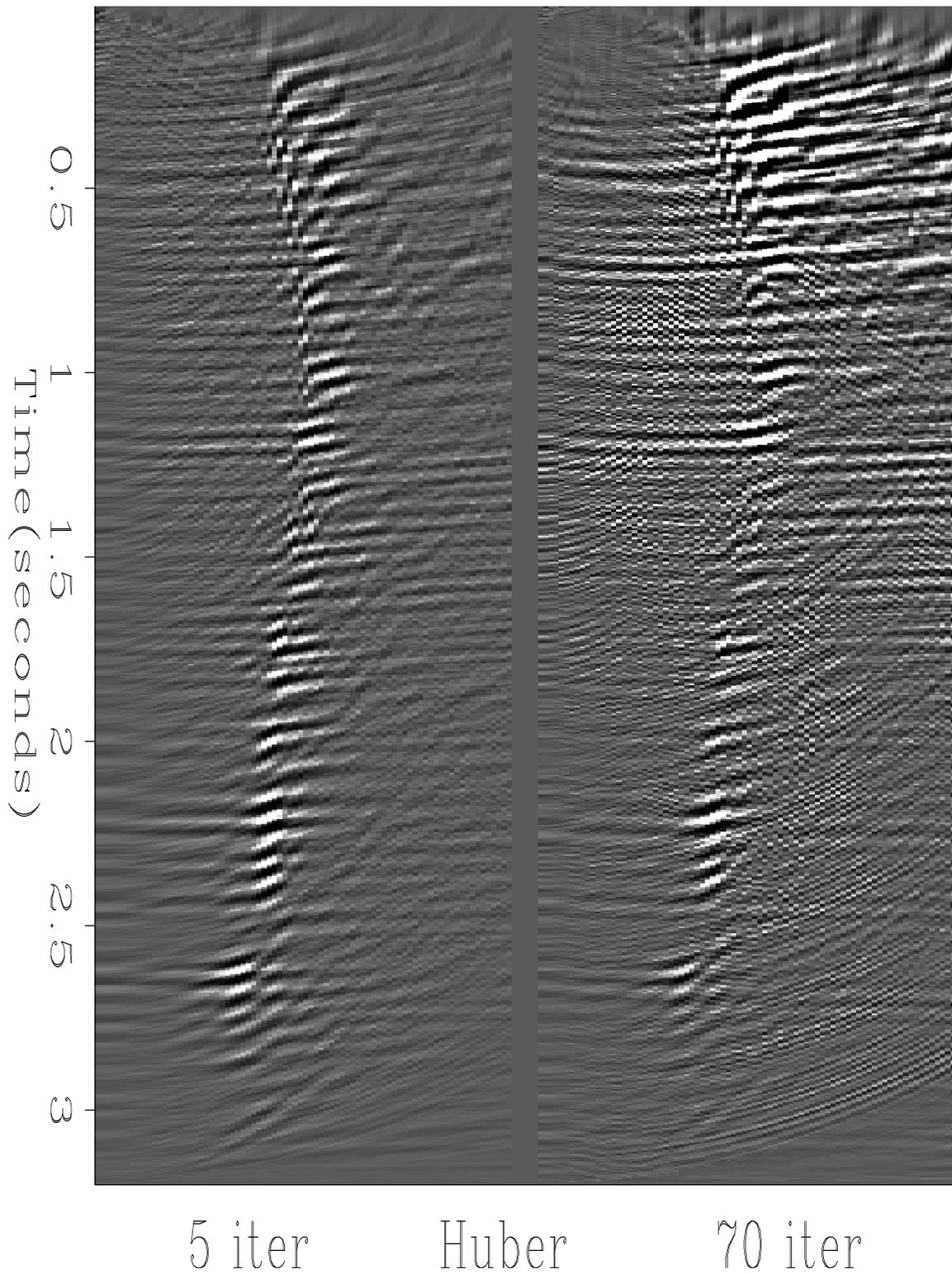


Figure 15: Input *B*: velocity panel using 5 and 70 Huber iterations. antoine1-compmodel11h
[NR]

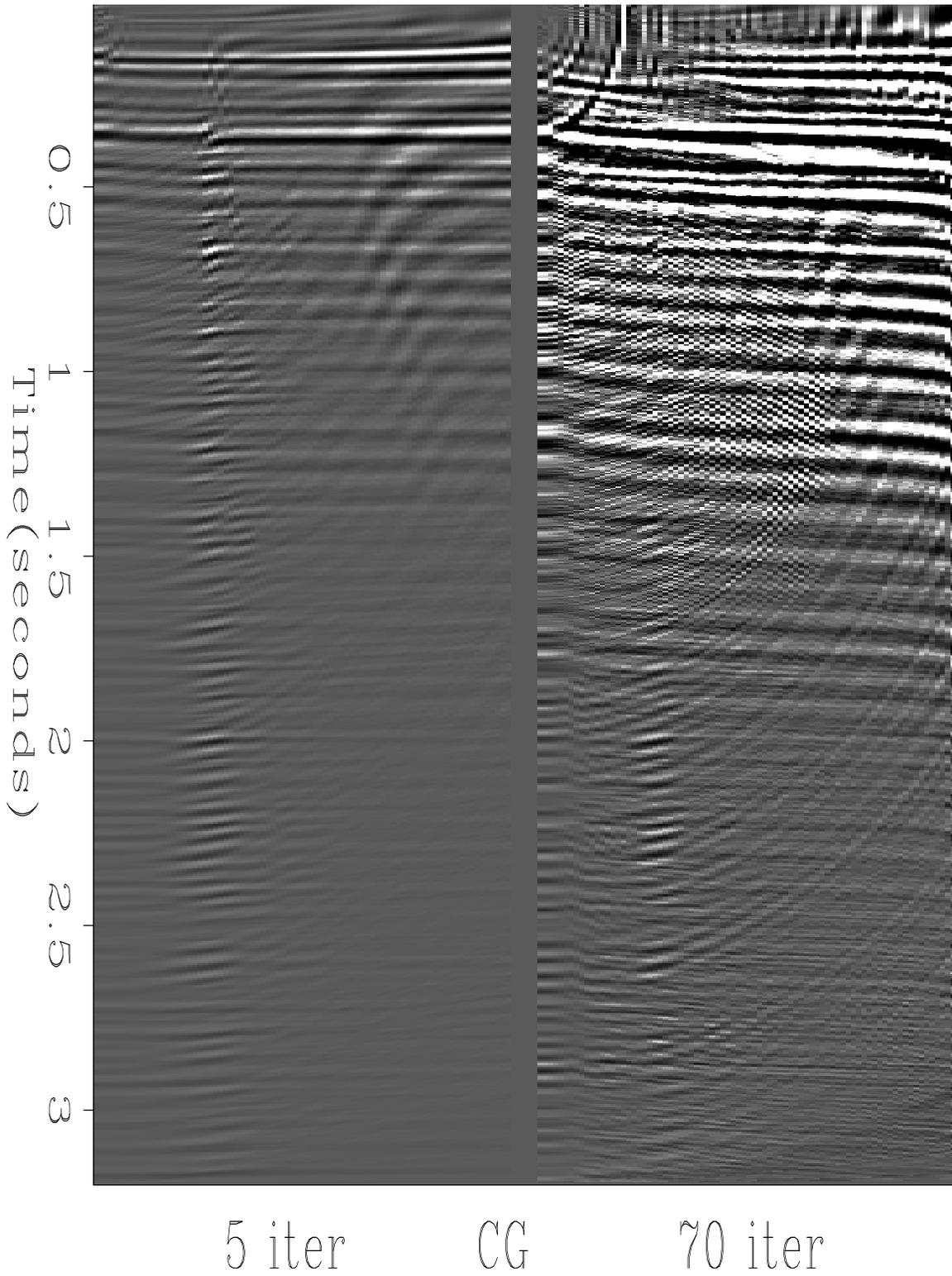


Figure 16: Input A: velocity panel using 5 and 70 CG iterations. antoine1-compmodel08g
[NR]

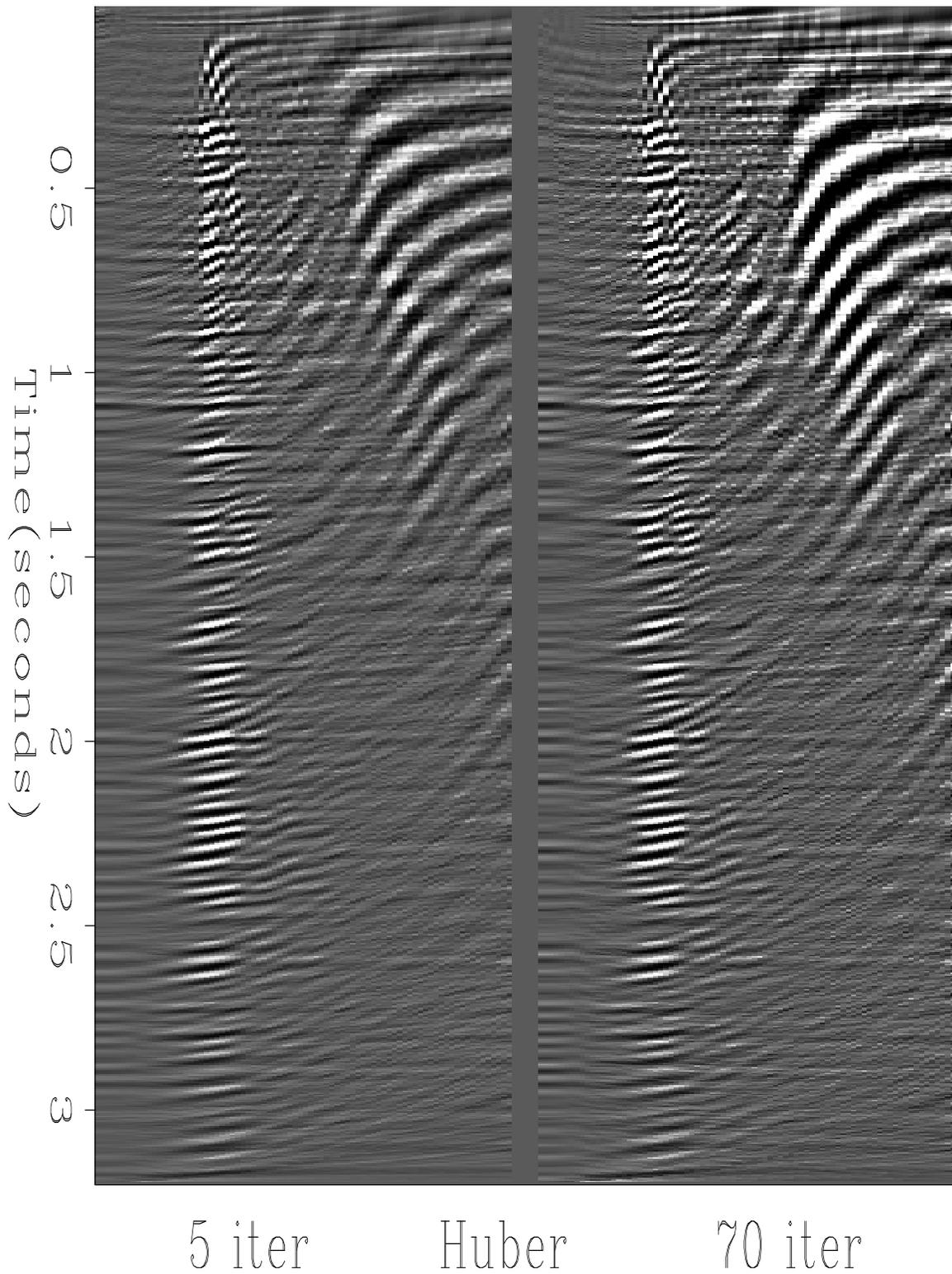


Figure 17: Input A: velocity panel using 5 and 70 Huber iterations. antoine1-compmodel08h
[NR]

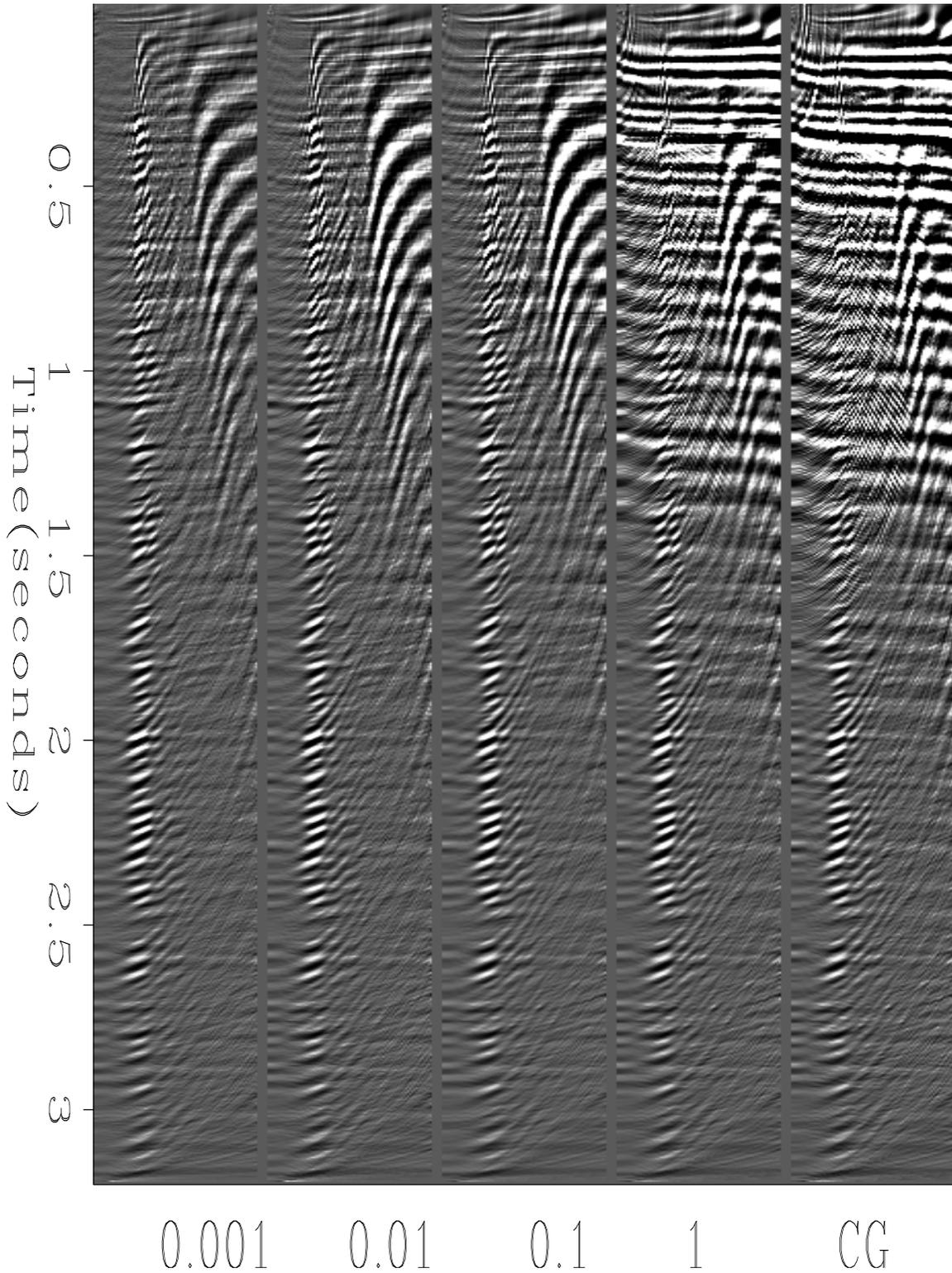


Figure 18: Input A: velocity panel for different Huber thresholds. The last panel is obtained using the CG. [antoine1-treshcomp08](#) [NR]

The right panel shows one of the well known characteristics of the conjugate gradient: after a relative small number of iterations, the algorithm starts to invert noise in the data space, making the model space particularly messy and uninterpretable. In contrast, the Huber results are fairly stable (Figure 15) since we can pick a reasonable velocity function without bothering with the noise. Some artifacts appear, however, which may be squelched by moving up or down the threshold. The next result is by far the most favorable to Huber (Input B). Figure 16 shows two velocity functions after 5 and 70 iterations using conjugate gradient. The noise level is so strong after 70 iterations that we cannot distinguish any coherent feature. Once again, the Huber solver gives a better velocity panel (Figure 17). This is a major improvement to the performance of CG.

Testing the Huber response for different thresholds is another important issue. Remember that this threshold determines the border between the l^1 and l^2 norms. Ideally, we would like not to have to specify this parameter *a priori*, but rather have the algorithm adaptively estimate it from the data. It's interesting, however, to test the system's sensitivity to the threshold. Figure 18 shows a comparison for ϵ ranging from 0.001 to 1 for 20 iterations. The last panel is the CG result. As expected, Huber starts to behave like an l^2 norm for large ϵ . Nonetheless, for $\epsilon = 0.001$ to $\epsilon = 0.1$, the velocity panels are fairly comparable giving us a wide range of possibilities in the choice of the threshold.

Computational efficiency

Despite its general design intended for arbitrary optimization problem, the Limited Memory BFGS "off the shelf" code was successful in minimizing the Huber misfit function. The results achieved by the two algorithms are significantly different so that a direct comparison of cost is difficult. The Huber/BFGS combination tends to require about twice as much CPU time per iteration as the Least squares/CG.

CONCLUSION

Since geophysical inverse problems are often ill-posed due to the presence of inconsistent data, high amplitude anomalies and outliers, relative insensitivity to noise is a desirable characteristic of an inversion method. The Huber function is a compromise misfit measure between l^1 and l^2 norms, not only boasting robustness in the presence of noise and outlier effects like l^1 measures, but also smoothness for small residuals characteristic of l^2 measures. The transition between the two norms is governed by a free parameter, the Huber threshold ϵ .

The Huber solver is fairly stable with respect to two major choices: the number of iterations and ϵ . The most striking result arises when we increase the number of iterations: while the l^2 result explodes, the Huber result looks stable. In addition, we may choose a threshold within a large range without degrading the estimated velocity model (once ϵ is small enough). We did not apply any regularization on the least squares method: it would make l^2 less noise-sensitive but requires either a regularization weight or a noise level estimate and results are rather sensitive to these. The Huber function also requires an estimate for the parameter ϵ , but

the results seem not to depend strongly on its choice. Furthermore, the Huber function gives better results than the l^2 when applied to velocity analysis showing its robustness to outlier effects. A data-dependent criterion for choosing the Huber threshold may prove fruitful, i.e., “treat $X\%$ of the data as Gaussian in the small residuals”, where X is specified interactively by the end-user.

These results encourage the use of the Huber function whenever the data are contaminated with noise and, as a robust and stable measure, to replace the l^2 norm in many geophysical applications.

ACKNOWLEDGEMENTS

The authors are grateful to Louis Vaillant and Jon Claerbout for their help and fruitful suggestions.

REFERENCES

- Barrodale, I., and Roberts, F. D. K., 1980, Algorithm 552 : Solution of the constrained l_1 linear approximation problem: *ACM Transactions on Mathematical Software*, **6**, 231–235.
- Chapman, N. R., and Barrodale, I., 1983, Deconvolution of marine seismic data using the l_1 norm: *Geophys. J. Roy. Astr. Soc.*, **72**, 93–100.
- Claerbout, J. F., and Black, J. L., 1997, Basic earth imaging: Class notes, <http://sepwww.stanford.edu/sep/prof/index.html>.
- Claerbout, J. F., and Muir, F., 1973, Robust modeling with erratic data: *Geophysics*, **38**, 820–844.
- Eklblom, H., and Madsen, K., 1989, Algorithms for non-linear huber estimation: *BIT*, **29**, 60–76.
- Fletcher, R., 1980, *Practical methods of optimization, I: Unconstrained Optimization*: Wiley & Sons, New York.
- Gockenbach, M. S., Petro, M., and Symes, W. W., 1999, C++ classes for linking optimization with complex simulations: *ACM Transactions on Mathematical Software*, in press.
- Huber, P. J., 1973, Robust regression: Asymptotics, conjectures, and Monte Carlo: *Ann. Statist.*, **1**, 799–821.
- Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: *Mathematics of Computation*, **95**, 339–353.
- Scales, J. A., and Gersztenkorn, A., 1988, Robust methods in inverse theory: *Inverse Problems*, **4**, 1071–1091.

Scales, J. A., Gersztenkorn, A., and Treitel, S., 1988, Fast lp solution of large, sparse, linear systems: application to seismic travel time tomography: *J. Comp. Phys.*, **75**, 314–333.

Taner, M., and Koehler, F., 1969, Velocity spectra: digital computer derivation and application of velocity functions: *Geophysics*, **34**, 859–881.

Tarantola, A., 1987, *Inverse problem theory*: Elsevier.

Taylor, H. L., Banks, S. C., and McCoy, J. F., 1979, Deconvolution with the l1 norm: *Geophysics*, **44**, 39–52.

Extremal regularization¹

William W. Symes²

ABSTRACT

Extremal regularization finds a model fitting the data to a specified tolerance, and additionally minimizing an auxiliary criterion. It provides relative model/data space weights when no statistical information about the model or data is available other than an estimate of noise level. A version of the Moré-Hebden algorithm using conjugate gradients to solve the various linear systems implements extremal regularization for large scale inverse problems. A deconvolution application illustrates the possibilities and pitfalls of extremal regularization in the linear case.

INTRODUCTION

Many important inverse problems are ill-posed: precise fit to data is either impossible or produces a model estimate so sensitive to data error as to obscure physically important model features. Since no intrinsic sample-level distinction between signal and noise exists in general, solution of such problems requires specification or estimation of data noise level, or acceptable degree of data misfit. Even so, the set of “feasible” models (fitting the data to within the prescribed tolerance or noise level) is very large. Finding a model representing the information content of the data then requires additional information.

Under some circumstances, Bayesian estimation theory provides a computational prescription for selecting a maximum likelihood model which represents the information inherent in the data and computing its *a posteriori* variance. When the modeling operator is linear, the data statistics are known and Gaussian, and signal and noise are known to be statistically independent, noise variance is the only additional parameter required to set up a linear system for the maximum likelihood estimator. However if these statistical hypotheses are not satisfied or if the modeling operator is nonlinear, Bayesian theory does not give an explicit prescription for selecting a representative model.

This paper explores an alternative selection principle, which I call *extremal regularization*. Extremal regularization does not require the extensive statistical assumptions of the Bayesian theory. It selects from the feasible set a model minimizing some auxiliary model property. Use of extremal regularization requires (1) a choice of auxiliary model property to extremize, (2) choices of norms to measure the data misfit and auxiliary model property, (3) knowledge

¹This report will also appear in the TRIP 1999 Annual Report

²email: symes@caam.rice.edu

of the data noise level, and (4) an algorithm for finding an extremal solution. Depending on application, (1), (2), and (3) involve greater or lesser degrees of arbitrariness. When the modeling operator is linear and the auxiliary property is quadratic in the model, extremal regularization amounts to minimization of a quadratic function subject to a quadratic constraint. In effect such an algorithm finds the penalty parameter or relative weight between data and model spaces as a function of the prescribed noised level. Note that the noise level has a much more obvious intuitive or physical meaning than the penalty parameter, though it is not always easy to determine from available data.

This notion is not new in geophysics. For example D. D. Jackson proposed similar ideas more than 20 years ago (Jackson, 1973, 1976). From (Jackson, 1979):

There are some who hold the recalcitrant point of view that the normal Backus-Gilbert resolving kernels tend to present results in too abstract a fashion, but that the use of *priori* data makes any error estimate rather arbitrary. For these, the only satisfactory evidence on which to base physical conclusions is a catalogue of models which fit the data well, are physically plausible, and contain among them models which have the maximum and minimum presence of some hypothetical feature. I must admit to having strong sympathies for this point of view.

Jackson extremized linear functions of the model subject to prescribed data misfit. These extrema represent the ends of model error bars.

This “recalcitrant” point of view is also natural when there is some nonlinear auxiliary quantity that should be minimized - or ideally even zeroed out - by virtue of fundamental model requirements of the model. This is the case for example in Claerbout’s proposal for signal extraction *via* Jensen inequalities (Claerbout, 1998, 1992) and for the extended version of differential semblance optimization for velocities (Gockenbach and Symes, 1997). In other settings, for example the deconvolution problem used as an example in this report, extremizing an auxiliary quantity serves merely to pick out a “simplest” solution amongst many.

The Moré-Hebden algorithm finds the relative weight between data and model spaces by applying Newton’s method to the so called secular equation. The secular equation requires that the norm of the auxiliary model property be equal to its prescribed value. Since this norm will change as you change the relative weight between data misfit and auxiliary model property, the secular equation determines the weight. This idea is much used in numerical optimization, where quadratically constrained quadratic minimization goes under the name “trust region problem” (Dennis and Schnabel, 1983). The published versions of Moré-Hebden (Moré, 1977; Hebden, 1973), also (Björk, 1997), pp. 205-6, have typically used LU decompositions to solve the linear systems required by Newton’s method, so are limited to small and medium scale problems. This report describes a version appropriate for large scale problems, using conjugate gradient iteration. The presence of this “inner” iteration and the necessary lack of precision in the solution of the Newton system has interesting consequences for convergence of the algorithm.

This report presents extremal regularization of linear inverse problems in the form of the quadratically constrained quadratic minimization problem solved by the Moré-Hebden algo-

rithm. Examples based on ill-posed 1D deconvolution illustrate the extremal regularization concept and the behaviour of the algorithm.

QUADRATICALLY CONSTRAINED QUADRATIC MINIMIZATION

A mathematical statement of the extremal regularization problem is (equivalent to)

$$\min_x (Rx)^T Rx \text{ subject to } (Ax - d)^T (Ax - d) \leq \sigma^2 d^T d$$

Here A is the modeling operator, x is the model vector, d is the data, and R is the regularizing operator. The noise level σ is *relative*, as that is usually the most useful way to pose noise estimates. Thus solution of this problem demands quadratically constrained quadratic minimization.

The solution minimizes whatever quality is represented by R , subject to fitting the data to a relative error level σ . The first order necessary conditions of optimality state that the solution satisfies

$$\begin{aligned} \lambda A^T (Ax - d) + R^T Rx &= 0 \\ \lambda [(Ax - d)^T (Ax - d) - \sigma^2 d^T d] &= 0 \end{aligned}$$

The first condition states parallelism of the gradients of the constraint and objective functions. The second implies that either the constraint is satisfied as an equality - i.e. the solution is on the boundary of the set of constraint-satisfying models - or else the Lagrange multiplier λ vanishes, which means that the most regular solution actually has a smaller residual than assumed - i.e. σ is larger than the actual noise level.

The first condition is the familiar normal equation of the unconstrained problem

$$\min_x \lambda (Ax - d)^T (Ax - d) + (Rx)^T Rx$$

or

$$\min_x (Ax - d)^T (Ax - d) + \epsilon^2 (Rx)^T Rx$$

where $\epsilon = \lambda^{-\frac{1}{2}}$ is the “notoriously elusive” relative weight between model space (really constraint space) and data space.

The point of this paper, and the basis of the Moré-Hebden algorithm, is that the first order conditions make the ϵ a function of the assumed noise level σ . Whenever σ can be estimated directly, this relationship provide a method of estimating ϵ .

ESTIMATING THE REGULARIZATION PARAMETER FROM THE NOISE LEVEL

For arbitrary $\lambda > 0$, denote by $x(\lambda)$ the solution of the normal equations

$$\lambda A^T (Ax(\lambda) - d) + R^T Rx(\lambda) = 0$$

Set

$$\phi(\lambda) = \sqrt{(Ax(\lambda) - d)^T (Ax(\lambda) - d)}$$

The *secular equation* is

$$\phi(\lambda) = \sigma \sqrt{d^T d}$$

and its solution, if it has one, gives the correct value of the Lagrange multiplier λ .

The Moré-Hebden algorithm takes its cue from the simplest possible case: x and d are one-dimensional, and A and R are scalars. In that very special case,

$$x(\lambda) = \frac{\lambda A d}{\lambda A^2 + R^2}$$

hence

$$\phi(\lambda) = \frac{R^2 d}{\lambda A^2 + R^2}$$

i.e. the reciprocal of ϕ is a linear function of λ . This suggests that Newton's method is more likely to converge quickly when applied to

$$\psi(\lambda) \equiv \frac{1}{\phi(\lambda)} = \frac{1}{\sigma \sqrt{d^T d}}$$

and that is exactly what the Moré-Hebden algorithm does.

The iteration proceeds as follows:

- initialize λ somehow
- until convergence do: replace λ by

$$\lambda - \frac{\psi(\lambda) - \frac{1}{\sigma \sqrt{d^T d}}}{\psi'(\lambda)}$$

in which ψ' stands for the derivative of ψ , which you compute like so:

$$\psi'(\lambda) = -\frac{\phi'(\lambda)}{\phi^2(\lambda)}$$

Now

$$\phi'(\lambda) = \phi^{-1}(\lambda) (Ax'(\lambda))^T (Ax(\lambda) - d)$$

From the normal equations,

$$(\lambda A^T A + R^T R)x'(\lambda) = A^T (d - Ax(\lambda))$$

so

$$\phi'(\lambda) = \phi^{-3}(\lambda) (A^T (Ax(\lambda) - d))^T (\lambda A^T A + R^T R)^{-1} (A^T (Ax(\lambda) - d))$$

Putting all of this together, one obtains the following algorithm for updating λ :

- solve the normal equations for x , compute the residual ϕ
- compute the normal residual $g = A^T(Ax - d)$
- solve the auxiliary system $(\lambda A^T A + R^T R)s = g$ for s
- compute $\phi' = g^T s / \phi$
- replace λ by

$$\lambda + \frac{\phi}{\phi'} \left(1 - \frac{\phi}{\sigma \sqrt{d^T d}} \right)$$

The first and third steps involve solution of linear systems, which in geophysical applications may be very large. Therefore, in contrast to conventional implementations of this algorithm, I use *conjugate gradient iteration* (Björk, 1997) to compute the solutions of these systems. As one might expect, the error reduction attained by these inner iterations affects the overall convergence rate of the algorithm.

A final detail: since $\lambda = \epsilon^{-2}$ must remain positive, I have replaced any large decrease implied by the above formula by a bisection strategy. Since $\phi' < 0$, as soon as λ is too small (which forces the weight onto the regularization term and increases the residual), the algorithm produces regular increases in λ and converges very rapidly, usually in one or two steps, so long as the normal equations are solved successfully. This is not always the case, but failure to converge rapidly appears to signal large data components associated with very small eigenvalues and is a sure sign that the noise level estimate σ has been chosen too small.

DECONVOLUTION EXAMPLES

The operator A is 1D convolution of a source pulse w with the input time series x . The data d is this convolution plus the noise series n . The regularization operator R is taken to be the identity I for all examples presented here.

All of the examples in this section will concern the source pulse (w) depicted in Figure 1, which is a 15 Hz Ricker wavelet sampled at 4 ms.

The data space consists of time series of length 1001, sample rate 4 ms. The noise free data is the convolution of the wavelet with a spike located at 1 s, see Figure 2.

The noise strength for the first set of experiments is 0.5. The noise is concentrated in the pulse passband, as it is the convolution of a pseudorandom sequence with the pulse, followed by scaling. Thus a signal of reasonable size fits the noisy data (Figure 3) very precisely.

The deconvolutions (signal estimations) resulting from underestimating, correctly estimating, and overestimating the noise level appear as Figure 4, Figure 5, and Figure 6. The estimated noise levels are 10%, 50%, and 80% respectively. In the notation of the last section, $\sigma = 0.1, 0.5, \text{ and } 0.8$ respectively. There is no particular identifiable virtue of one result over the other, which reinforces my contention that in order to solve one of these problems, you

Figure 1: 15 Hz Ricker wavelet used in deconvolution experiments.

bill1-fig1 [ER]

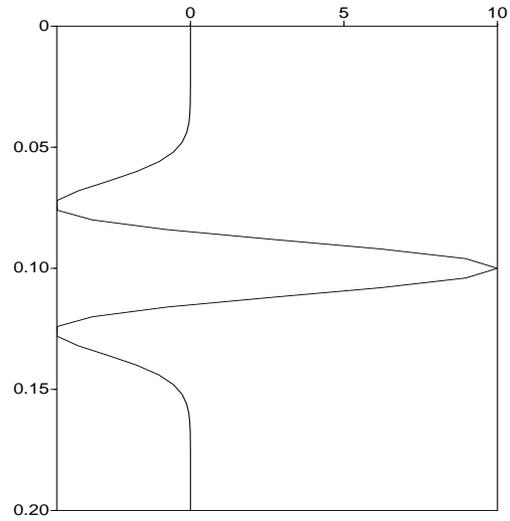


Figure 2: Noise free data. bill1-fig2 [ER]

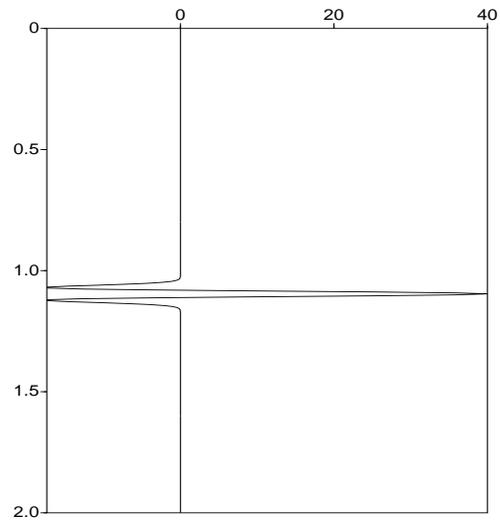
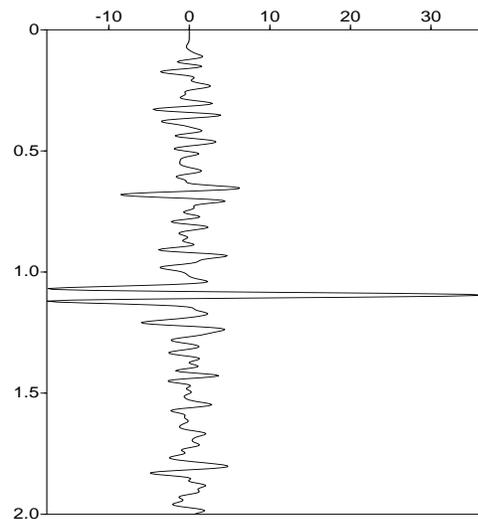
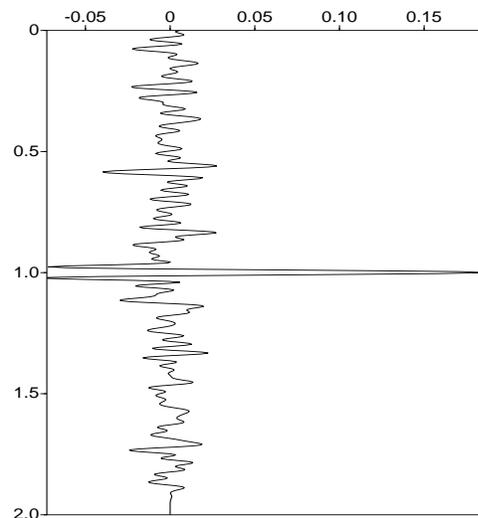


Figure 3: Noisy data: 50% RMS filtered noise. `bill1-fig3` [ER]



must have an independent means of estimating noise level: neither the data nor the results of the signal estimations reveal the signal/noise dichotomy.

Figure 4: Signal estimate: target noise level 10%, filtered noise. `bill1-fig4` [ER]



Note that even for the correctly estimated noise level, namely $\sigma = 0.5$, you do not recover the isolated spike. The discrepancy is partly due to the less than perfect linear system solves *via* conjugate gradient iteration, but also to the nature of the problem: it is actually possible to achieve the same fit error as that provided by the noise free data with a slightly smaller signal, by fitting the signal less and the noise more. That's because signal and noise are not entirely orthogonal (and they rarely are, so you're going to have to live with this "crosstalk" imperfection!).

Figure 5: Signal estimate: target noise level 50%, filtered noise.
bill1-fig5 [ER]

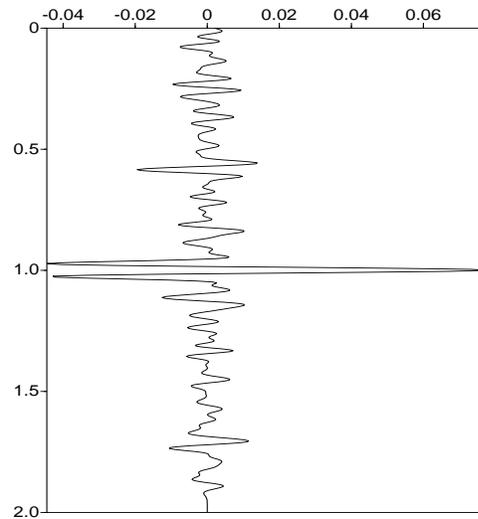
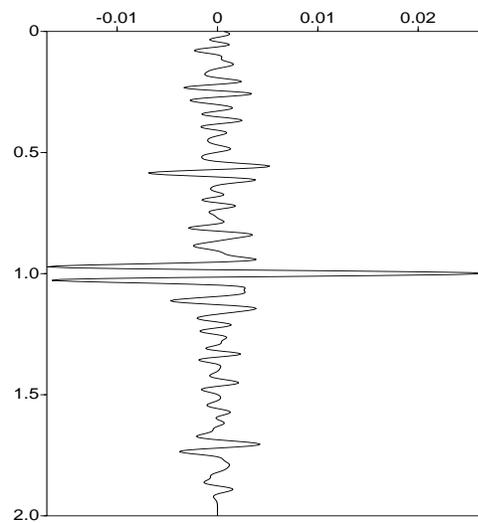


Figure 6: Signal estimate: target noise level 80%, filtered noise.
bill1-fig6 [ER]



The relation between the noise level or fit error and the penalty parameter ϵ really is obscure, as the following results suggest:

- $\sigma = 0.1 \Rightarrow \epsilon = 48.8958$
- $\sigma = 0.5 \Rightarrow \epsilon = 213.011$
- $\sigma = 0.8 \Rightarrow \epsilon = 470.256$

I would not have guessed the precise values of these ϵ s - would you have done? On the other hand the trend is exactly as you would expect: as you permit more misfit, you are able to make the auxiliary quantity (the model L^2 norm in this case) smaller, corresponding to a larger ϵ .

The second set of experiments uses the same noise free data contaminated with unfiltered noise at the 50% level (Figure 7). As the data now contain much out of passband energy, a perfect fit is no longer achievable.

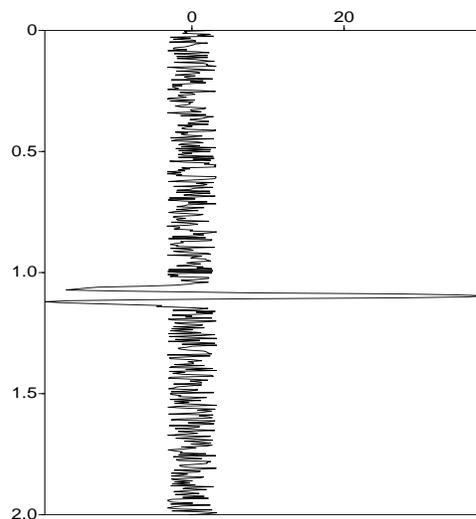


Figure 7: Noisy data: 50% RMS unfiltered noise. [bill1-fig7](#) [ER]

Estimating the noise level at $\sigma = 0.1, 0.5$, and 0.8 as before gives the signals depicted in Figure 8, Figure 9, and Figure 10 respectively. The first of these fit errors is impossible to achieve by means of the conjugate gradient algorithm at least with any reasonable number of iterations. The solution simply grows without bound, as one would expect, and retains almost no character of the target model (Figure 8). The correct estimate $\sigma = 0.5$ on the other hand gives you a reasonable estimate of the signal (Figure 9), with a bandlimited version of the spike dominating the series.

Again, the precise values of ϵ are inscrutable:

- $\sigma = 0.1 \Rightarrow \epsilon = 2.04543e-11$
- $\sigma = 0.5 \Rightarrow \epsilon = 130.219$

Figure 8: Signal estimate: target noise level 10%, unfiltered noise.
bill1-fig8 [ER]

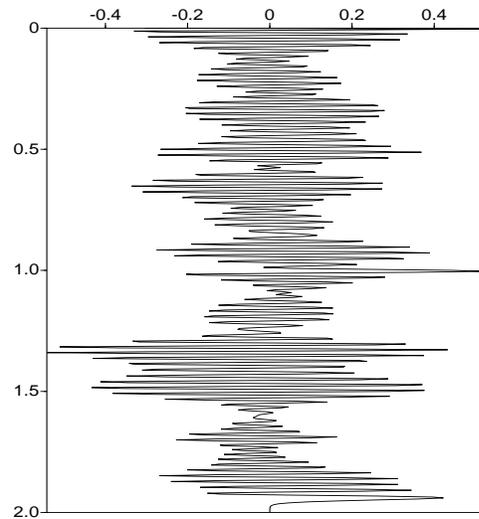
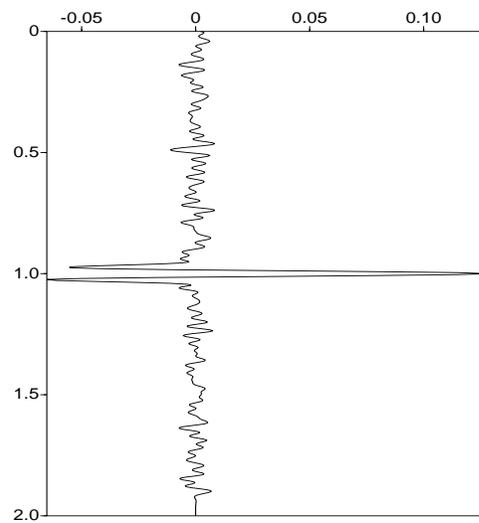


Figure 9: Signal estimate: target noise level 50%, unfiltered noise.
bill1-fig9 [ER]



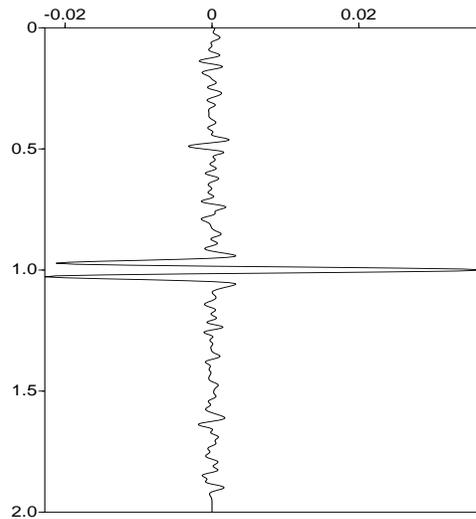


Figure 10: Signal estimate: target noise level 80%, unfiltered noise.

bill1-fig10 [ER]

- $\sigma = 0.8 \Rightarrow \epsilon = 399.49$

The trend is even more marked here. The large out-of-band components in the data are essentially impossible to fit. So when you ask for a rather precise fit - 10% error - the weight on the model space decreases throughout the iteration, apparently with no end in sight. The value in the table above was the result of 10 Moré-Hebden iterations, and ϵ diminished by an order of magnitude or so each iteration. As soon as the level of fit permits you to discard the out of band components (that's what happened at $\sigma = 0.5$), the desired fit actually occurs and a reasonable value of ϵ results.

Clearly, prior knowledge of a reasonable model size would enable you to guess σ in this example. However then you have to know the size of the model! This may be no more obvious than the size of the noise. This observation reinforces my contention that solution of problems like these demands that you know *something* in addition to the data samples - there is no “born yesterday” bootstrapping into a signal - noise distinction.

CONCLUSION

Extremal regularization appears to be practical for large scale problems, as the Moré-Hebden algorithm with conjugate gradient inner solves either converges in a reasonable number of steps or doesn't converge when the constraint (target noise level) forces too many small singular values into the act. All of these terms are relative - small, doesn't converge, etc. Modulo floating point arithmetic, the algorithm will *always* work if enough effort is expended. The issue of course is reasonable level of effort, and that is in some sense a translation of the concept of “noise level” - it's the misfit between the data and what you can achieve with an easily computable model, no more.

Thus extremal regularization as implemented in this report appears to give a reasonable approach to relative weighting in model and data space when an independent estimate of noise level is somehow available. This is the case for example in the examples mentioned in the introduction. Maybe quiet parts of seismic traces furnish pure noise series which might give a usable estimate of noise level - provided that the modeling operator is sophisticated enough to fit the rest!

ACKNOWLEDGEMENT

The author performed part of the work reported here while a guest of the Stanford Exploration Project. The author thanks the Director of SEP, Prof. Jon Claerbout, and the other project members for their hospitality and for the stimulating intellectual atmosphere of SEP.

This work was partially supported by the National Science Foundation under grant numbers DMS 9404283 and DMS 9627355; the Office of Naval Research under grant number N00014-96-1-0156; the U. S. Department of Energy under grant number DE-FG07-97 ER14827; and The Rice Inversion Project. TRIP Sponsors for 1998 are Amerada Hess, Amoco Production Co., ARCO Oil and Gas, Conoco Inc., Discovery Bay, Exxon Production Research Co., Mobil Research and Development Corp., and Shell International Research.

REFERENCES

- Björk, A., 1997, Numerical methods for least squares problems: Society for Industrial and Applied Mathematics, Philadelphia.
- Claerbout, J. F., 1992, Earth sounding analysis: Processing versus inversion: Blackwell, Boston.
- Claerbout, J., S/n segregation: The best ratio between data fitting and model regularization: Gain control:, Technical report, Stanford Exploration Project, 1998.
- Dennis Jr., J., and Schnabel, R., 1983, Numerical methods for unconstrained optimization and nonlinear equations: Prentice-Hall, Englewood Cliffs.
- Gockenbach, M., and Symes, W. W., 1997, Duality for inverse problems in wave propagation *in* Biegler, L., Coleman, T., Santosa, F., and Conn, A., Eds., Large Scale Optimization:: Springer Verlag.
- Hebden, M. D., An algorithm for minimization using exact second derivatives:, Technical Report TP515, A.E.R.E., Harwell, 1973.
- Jackson, D. D., 1973, Marginal solutions to quasi-linear inverse problems in geophysics: the edgehog method: Geophysical Journal of the Royal Astronomical Society, **35**, 121–136.
- Jackson, D. D., 1976, Most squares inversion: J. Geophys. Research, **81**, 1027–1030.

Jackson, D. D., 1979, The use of *a priori* data to resolve nonuniqueness in linear inversion: Geophysical Journal of the Royal Astronomical Society, **57**, 137–157.

Moré, J. J., 1977, The levenberg-marquardt algorithm: implementation and theory *in* Watson, G. A., Ed., Numerical Analysis:: Springer-Verlag, 630 ff.

APPENDIX: WORKING WITH THE EXAMPLES IN THIS REPORT

This report, together with its associated files, constitutes a reproducible research document. Makefiles tie together the various components - text, code source, data, and postscript figures. The principal make rules are the SEP standards: build, view, clean, burn. In this section, I will assume familiarity with the Stanford Exploration Project reproducible research concept, which guides the structure of this document. This code behind this document is an application of the Hilbert Class Library. So the first thing you need to do is to make HCL available. If HCL is already installed on your system, you do this by adding a line to your `.cshrc` file or appropriate component file. Otherwise you must install HCL first. The easiest way to do this is to download it from the TRIP web page:

`http://www.trip.caam.rice.edu`

and follow the installation instructions. The code also depends on the SU/SEG Y vector class package `sVector`, which therefore must also be installed. It will be part of the next HCL release. **NB:** At Rice/TRIP and Stanford/SEP, no installation is necessary: the packages are already installed. Simply add the following lines to your shell environment files:

- at TRIP, add to your `.cshrc`:

```
setenv HCLROOT /import/masc39c/symes/hclr0.9
setenv KBDAROOT /import/masc39c/symes/kbda
setenv QPROOT <path to the root directory of this package>
```

- at SEP, add to your `Setup/cshrc.generic`:

```
setenv HCLROOT /jon/symes/hclr0.9
setenv KBDAROOT /jon/symes/kbda
setenv QPROOT <path to the root directory of this package>
```

The “root directory of this package” referenced in these instructions is the directory you create by downloading the `tar` file containing this report. In so doing, you create a directory tree with root named `qqqm`. This is the “root” in question. All pathnames in the following discussion are relative to this root. HCL includes a set of make rules which evolved from the SEP rule set as it was about two years ago. I am sure that SEP’s rules have also evolved, and differently.

The makefiles for this report are all output of the HCL makefile autowriter, `maw`, and use HCL's rules. If you get as far as modifying makefiles by hand, bear the possible HCL/SEP incompatibility in mind. You can rebuild the entire package simply by entering `make build` in the root directory. You will construct all of the executables and final results (`.ps` files in the `Fig`) directories, including this postscript version of this report. You can make individual results in the usual way. Note that all figures in this report except the first two will vary from build to build, as you choose a new pseudorandom seed each time you execute the commands. The command for deconvolution is `sfilter/decon.x`. You execute it by following it with a parameter file name: `decon.x par` - it reads all of its parameters from the file. Regrettably the "getpar" device used in current HCL programs is not flexible enough to permit specification of parameters on the command line. Probably I should just steal SEP's getpar! You can use the executable `sfilter/decon.x` with other data by altering its input parameters, and so explore the capabilities of the algorithm using data other than that supplied with this report. Parameter control requires you to edit the parameter files manually. HCL parameter files are `keyword=value` lists; the values can be integers, floating point numbers with any size of mantissa, and strings. Parameters to be read only by one part of the program (typically a class constructor) get an identifying string prepended, with a double colon. Thus the parameter `Tol` for the conjugate gradient algorithm becomes `CG::Tol`. That is, the parameter file can specify many variables named `Tol`, so long as they have been equipped with qualifiers which allow each program unit to choose a unique value. Files should be in either SU (stripped SEGY) or SEGY formats. `sVector` currently supports only native binary floating point representation, so if you port data from Linux to SGI etc. you will have to byte-swap it. Here is the parameter file structure for the deconvolution example:

```
Sigma=0.5
Lambda=0.0001
Wavelet="rick15.su"
DataTimeSeries="fnd.su"
CG::Tol=1.e-4
CG::MaxItn=50
CG::DispFlag=3
QCQM::Tol=0.01
QCQM::MaxItn=10
QCQM::DispFlag=1
```

The parameters are

- `Sigma`: target noise level
- `Lambda`: initial estimate of λ . It may be worth thinking about sensible defaults or crude estimates for this, and the scalar model actually suggests such an estimate. For the moment, set by hand.
- `wavelet`: name of wavelet data file, in either SU or SEGY format.
- `DataTimeSeries`: filename for data time series

- `CG::Tol`: convergence tolerance for HCL conjugate code - iteration terminates if *normal* residual, i.e. in this case $\lambda A^T(Ax - d) + R^T Rx$
- `CG::MaxItn`: maximum conjugate gradient iterations
- `CG::DispFlag`: controls verbosity of CG output, as explained in HCL reference manual (through TRIP web page: point your browser at

<http://www.trip.caam.rice.edu>

and follow the links to the HCL reference manual, page on the conjugate gradient algorithms). Level 3 is max output, including a summary of the progress at each iteration. Note that the residual reported is the *normal* residual for this application!

- `QCQM::Tol`: the Moré-Hebden algorithm considers itself converged when the relative error in the constraint is within this amount
- `QCQM::MaxItn`: maximum Moré-Hebden iterations
- `QCQM::DispFlag`: controls verbosity of Moré-Hebden code: 0 = silent, any other value prints diagnostic information about run

Amongst the diagnostics printed out at the end of a run when `QCQM::DispFlag` is set you will find “Lagrange cosine”. This is the cosine of the angle between the constraint gradient and the objective gradient. It diagnoses the success of the constrained optimization: if it is very close to 1, then the two gradients are parallel and the first order necessary condition has been satisfied. This occurs in the deconvolution examples in all cases except that depicted in Figure 8, in which nothing converges and you can’t fit the data. Apparently failure to get the Lagrange cosine close to 1 in a reasonable number of CG and Moré-Hebden iterations implies that the noise level has been set too small and you are trying to match data components associated with very small singular values.

Short Note

Backus revisited: Just in time

Francis Muir¹

INTRODUCTION

The essence of Backus theory (Backus, 1962) is that it allows a bunch of layers to be replaced by a single layer. The new homogeneous medium has elastic properties identical in the long-wavelength limit so that mass and travel-time are conserved, but wavelet shape is not. In this paper I show that for normal incidence plane waves the three elastic layer parameters of thickness, compliance, and density can be replaced by the two, travel-time and impedance, without losing reflection and transmission information.

DEVELOPMENT

In the normal incidence case of Backus averaging, the layer properties of thickness, ΔZ_j , compliance, S_j , and mass density, R_j , are replaced by the corresponding properties of the equivalent homogeneous medium, Z_{equiv} , S_{equiv} , and R_{equiv} :

$$Z_{\text{equiv}} = \sum \Delta Z_j \quad (1)$$

$$S_{\text{equiv}} = \sum S_j \Delta Z_j / \sum \Delta Z_j \quad (2)$$

$$R_{\text{equiv}} = \sum R_j \Delta Z_j / \sum \Delta Z_j \quad (3)$$

that is, the thickness of Z_{equiv} is the sum thickness of the layers, $Z = \sum \Delta Z_j$, and the equivalent medium mechanical properties are the thickness-weighted averages of those of the layered medium. However, these layers can also be described in terms of the layer properties of one-way travel-time, ΔT_j , and impedance, I_j , with slowness, L_j acting as the means for changing the independent variable between depth and time. It is well known that:

$$I_j = \sqrt{R_j/S_j} \quad (4)$$

$$L_j = \sqrt{R_j S_j} \quad (5)$$

and from these:

$$R_j = I_j L_j \quad (6)$$

$$S_j = I_j^{-1} L_j \quad (7)$$

¹email: francis@stanford.edu

and thus:

$$\begin{aligned}
 I_{\text{equiv}} &= \sqrt{R_{\text{equiv}}/S_{\text{equiv}}} \\
 &= \sqrt{\sum R_j \Delta Z_j / \sum S_j \Delta Z_j} \\
 &= \sqrt{\sum I_j L_j \Delta Z_j / \sum I_j^{-1} L_j \Delta Z_j}
 \end{aligned} \tag{8}$$

but $L_j \Delta Z_j = \Delta T_j$, so:

$$I_{\text{equiv}} = \sqrt{\sum I_j \Delta T_j / \sum I_j^{-1} \Delta T_j} \tag{9}$$

and by similar reasoning the slowness equivalent:

$$L_{\text{equiv}} = (\sum \Delta Z_j)^{-1} \sqrt{\sum I_j \Delta T_j \sum I_j^{-1} \Delta T_j} \tag{10}$$

and, since $T_{\text{equiv}} = Z_{\text{equiv}} L_{\text{equiv}}$:

$$T_{\text{equiv}} = \sqrt{\sum I_j \Delta T_j \sum I_j^{-1} \Delta T_j} \tag{11}$$

DISCUSSION

What have we done? Firstly we have shown that Backus exactly translates from depth and elastic properties to time and impedance. In other words we have transplanted Backus from the physical world to the data processor's world of traveltimes and reflectivity as characterized by Goupillaud's ladder and Chapter 8 of FGDP (Claerbout, 1976). Secondly we have swapped two elastic parameters, compliance and density, for one, impedance, without any loss of information. This may represent a key to extending Backus theory sufficiently away from zero frequency to provide useful approximations to changes in wavelet shape due to scattering.

PREVIOUS PUBLICATION

This result is new to me but it is possible — even likely — that it has been previously published in this explicit form. If this is so, then the author would much appreciate getting the citation. In fact, Bill Symes, in a personal communication, has let me know that Bamberger came up with an integral representation of my summation form.

ACKNOWLEDGMENTS

To the Department of Energy for funding this work through DE-FG03-99ER-14933, to Amos Nur for supplying the umbrella for my continuing studies, and to Jon Claerbout for never quite letting me leave his SEP.

REFERENCES

Backus, G. E., 1962, Long-wave anisotropy produced by horizontal layering: *J. Geophys. Res.*, **66**, 4427–4440.

Claerbout, J. F., 1976, *Fundamentals of geophysical data processing*: Blackwell.

Elastic wave propagation and attenuation in a double-porosity dual-permeability medium

James G. Berryman¹ and Herbert F. Wang²

ABSTRACT

To account for large-volume low-permeability storage porosity and low-volume high-permeability fracture/crack porosity in oil and gas reservoirs, phenomenological equations for the poroelastic behavior of a double porosity medium have been formulated and the coefficients in these linear equations identified. This generalization from a single porosity model increases the number of independent inertial coefficients from three to six, the number of independent drag coefficients from three to six, and the number of independent stress-strain coefficients from three to six for an isotropic applied stress and assumed isotropy of the medium. The analysis leading to physical interpretations of the inertial and drag coefficients is relatively straightforward, whereas that for the stress-strain coefficients is more tedious. In a quasistatic analysis, the physical interpretations are based upon considerations of extremes in both spatial and temporal scales. The limit of very short times is the one most pertinent for wave propagation, and in this case both matrix porosity and fractures are expected to behave in an undrained fashion, although our analysis makes no assumptions in this regard. For the very long times more relevant to reservoir drawdown, the double porosity medium behaves as an equivalent single porosity medium. At the macroscopic spatial level, the pertinent parameters (such as the total compressibility) may be determined by appropriate field tests. At the mesoscopic scale pertinent parameters of the rock matrix can be determined directly through laboratory measurements on core, and the compressibility can be measured for a single fracture. We show explicitly how to generalize the quasistatic results to incorporate wave propagation effects and how effects that are usually attributed to squirt flow under partially saturated conditions can be explained alternatively in terms of the double-porosity model. The result is therefore a theory that generalizes, but is completely consistent with, Biot's theory of poroelasticity and is valid for analysis of elastic wave data from highly fractured reservoirs.

INTRODUCTION

It is well-known in the phenomenology of earth materials that rocks are generally heterogeneous, porous, and often fractured or cracked. In situ, rock pores and cracks/fractures can

¹email: berryman@sep.stanford.edu

²Department of Geology and Geophysics, University of Wisconsin, Madison, WI 53706.
email: wang@geology.wisc.edu

contain oil, gas, or water. These fluids are all of great practical interest to us. Distinguishing these fluids by their seismic signatures is a key issue in seismic exploration and reservoir monitoring. Understanding their flow characteristics is typically the responsibility of the reservoir engineer.

Traditional approaches to seismic exploration have often made use of Biot's theory of poroelasticity (Biot, 1941; 1956a,b; 1962; Gassmann, 1951). Many of the predictions of this theory, including the observation of the slow compressional wave, have been confirmed by both laboratory and field experiment (Berryman, 1980; Plona, 1980; Johnson *et al.*, 1982; Chin *et al.*, 1985; Winkler, 1985; Pride and Morgan, 1991; Thompson and Gist, 1993; Pride, 1994). Nevertheless, this theory has always been limited by an explicit assumption that the porosity itself is homogeneous. Although this assumption is often applied to acoustic studies of many core samples in a laboratory setting, heterogeneity of porosity nevertheless exists in the form of pores and cracks. Also, single homogeneous porosity is often not a good assumption for application to realistic heterogeneous reservoirs in which porosity exists in the form of matrix and fracture porosity. One approach to dealing with the heterogeneity is to construct a model that is locally homogeneous, *i.e.*, a sort of finite element approach in which each block of the model satisfies Biot-Gassmann equations. This approach may be adequate in some applications, and is certainly amenable to study with large computers. However, such models necessarily avoid the question of how we are to deal with heterogeneity on the local scale, *i.e.*, much smaller than the size of blocks typically used in such codes.

Although it is clear that porosity in the earth can and does come in virtually all shapes and sizes, it is also clear that just two types of porosity are often most important at the reservoir scale: (1) Matrix porosity occupies a finite and substantial fraction of the volume of the reservoir. This porosity is often called the storage porosity, because this is the volume that stores the fluids of interest to us. (2) Fracture or crack porosity may occupy very little volume, but nevertheless has two very important effects on the reservoir properties. The first effect is that fractures/cracks drastically weaken the rock elastically, and at very low effective stress levels introduce nonlinear behavior since very small changes in stress can lead to large changes in the fracture/crack apertures (and at the same time change the fracture strength for future changes). The second effect is that the fractures/cracks often introduce a high permeability pathway for the fluid to escape from the reservoir. This effect is obviously key to reservoir analysis and the economics of fluid withdrawal.

It is therefore not surprising that many attempts have been made to incorporate fractures into rock models, and especially models that try to account for partial saturation effects and the possibility that fluid moves (or squirts) during the passage of seismic waves (Budiansky and O'Connell, 1975; O'Connell and Budiansky, 1977; Mavko and Nur, 1979; Mavko and Jizba, 1991; Dvorkin and Nur, 1993). Previous attempts to incorporate fractures have generally been rather ad hoc in their approach to the introduction of the fractures into Biot's theory, if Biot's theory was used at all. The present authors have recently started an effort to make a rigorous extension of Biot's poroelasticity to include fractures/cracks by making a generalization to double-porosity/dual-permeability modeling (Berryman and Wang, 1995). The previously published work concentrated on the fluid flow aspects of this problem in order to deal with the interactions between fluid withdrawal and the elastic behavior (closure) of fractures dur-

ing reservoir drawdown. The resulting equations have been applied recently to the reservoir consolidation problem by Lewallen and Wang (1998).

It is the purpose of the present work to point out that a similar analysis applies to the wave propagation problem. Just as Biot's early work on poroelasticity for consolidation (Biot, 1941) led to his later work on wave propagation (Biot, 1956; 1962), the present work follows our own work on consolidation (Berryman and Wang, 1995) with its extension to wave propagation. We expect it will be possible to incorporate all of the important physical effects in a very natural way into this double-porosity extension of poroelasticity for seismic wave propagation. The price we pay for this rigor is that we must solve a larger set of coupled equations of motion locally. Within traditional poroelasticity, there are two types of equations that are coupled. These are the equations for the elastic behavior of the solid rock and the equations for elastic and fluid flow behavior of the pore fluid. In the double-porosity extension of poroelasticity, we have not two types of equations but three. The equations for the elastic behavior of the solid rock will be unchanged except for the addition of a new coupling term, while there will be two types of pore-fluid equations (even if there is only one fluid present) depending on the environment of the fluid. Pore fluid in the matrix (storage) porosity will have one set of equations with coupling to fracture fluid and solid; while fluid in the fractures/cracks will have another set of equations with coupling to storage fluid and solid. Although solving these equations is surely more difficult than for simple acoustics/elasticity, finding solutions for the double-porosity equations is not significantly more difficult than for traditional single-porosity poroelasticity. We will solve these equations in the present paper. We will first derive them and then show that the various coefficients in these equations can be readily identified with measurable quantities. Then we develop and solve (numerically) the dispersion relation.

EQUATIONS OF MOTION

The seismic equations of motion for a double-porosity medium have been derived recently by Tuncay and Corapcioglu (1996) using a volume averaging approach. (These authors also provide a thorough review of the prior literature on this topic.) We will present instead a quick derivation based on ideas similar to those of Biot's original papers (Biot, 1956; 1962), wherein a Lagrangian formulation is presented and the phenomenological equations derived.

Physically what we need is quite simple — just equations embodying the concepts of *force = mass × acceleration*, together with dissipation due to viscous loss mechanisms. The forces are determined by taking a derivative of an energy storage functional. The appropriate energies are discussed at length later in this paper, so for our purposes in this section it will suffice to assume that the constitutive laws relating stress and strain are known, and so the pertinent forces are the divergence of the solid stress field $\tau_{ij,j}$ and the gradients of the two fluid pressures $p_{,i}^{(1)}$ and $p_{,i}^{(2)}$ for the matrix and fracture fluids, respectively. (In this notation, i, j index the three Cartesian coordinates x_1, x_2, x_3 and a comma preceding a subscript indicates a derivative with respect to the specified coordinate direction.) Then, the only new work we need to do to establish the equations of motion for dynamical double-porosity systems concerns the inertial terms arising from the kinetic energy of the system.

Generalizing Biot's approach (Biot, 1956) to the formulation of the kinetic energy terms, we find that, for a system with two fluids, the kinetic energy T is determined by

$$2T = \rho_{11}\dot{\mathbf{u}} \cdot \dot{\mathbf{u}} + \rho_{22}\dot{\mathbf{U}}^{(1)} \cdot \dot{\mathbf{U}}^{(1)} + \rho_{33}\dot{\mathbf{U}}^{(2)} \cdot \dot{\mathbf{U}}^{(2)} + 2\rho_{12}\dot{\mathbf{u}} \cdot \dot{\mathbf{U}}^{(1)} + 2\rho_{13}\dot{\mathbf{u}} \cdot \dot{\mathbf{U}}^{(2)} + 2\rho_{23}\dot{\mathbf{U}}^{(1)} \cdot \dot{\mathbf{U}}^{(2)}, \quad (1)$$

where \mathbf{u} is the displacement of the solid, $\mathbf{U}^{(k)}$ is the displacement of the k th fluid which occupies volume fraction $\phi^{(k)}$, and the various coefficients ρ_{11} , ρ_{12} , etc., are mass coefficients that take into account the fact that the relative flow of fluid through the pores is not uniform, and that oscillations of solid mass in the presence of fluid leads to induced mass effects. Clarifying the precise meaning of these displacements is beyond our current scope, but other recent publications help with these interpretations (Pride and Berryman, 1998).

Dissipation plays a crucial role in the motion of the fluids and so cannot be neglected in this context. The appropriate dissipation functional will take the form

$$2D = b_{12}(\dot{\mathbf{u}} - \dot{\mathbf{U}}^{(1)}) \cdot (\dot{\mathbf{u}} - \dot{\mathbf{U}}^{(1)}) + b_{13}(\dot{\mathbf{u}} - \dot{\mathbf{U}}^{(2)}) \cdot (\dot{\mathbf{u}} - \dot{\mathbf{U}}^{(2)}) + b_{23}(\dot{\mathbf{U}}^{(1)} - \dot{\mathbf{U}}^{(2)}) \cdot (\dot{\mathbf{U}}^{(1)} - \dot{\mathbf{U}}^{(2)}). \quad (2)$$

This formula assumes that all dissipation is caused by motion of the fluids either relative to the solid, or relative to each other. (Other potential sources of attenuation, especially for partially saturated porous media (Stoll, 1985; Miksis, 1988), should also be treated, but will not be considered here.) We expect the fluid-fluid coupling coefficient b_{23} will generally be small and probably negligible, whenever the double-porosity model is appropriate for the system under study.

Lagrange's equations then show easily that

$$\frac{\partial}{\partial t} \left(\frac{\partial T}{\partial \dot{u}_i} \right) + \frac{\partial D}{\partial \dot{u}_i} = \tau_{ij,j}, \quad \text{for } i = 1, 2, 3, \quad (3)$$

and that

$$\frac{\partial}{\partial t} \left(\frac{\partial T}{\partial \dot{U}_i^{(k)}} \right) + \frac{\partial D}{\partial \dot{U}_i^{(k)}} = -\bar{p}_i^{(k)}, \quad \text{for } i = 1, 2, 3; k = 1, 2, \quad (4)$$

where the pressures $\bar{p}^{(k)}$ are the macroscopic fluid pressures across interfaces and are related to the internal pore pressures $p^{(k)}$ by factors of the porosity so that $\bar{p}^{(1)} = (1 - v^{(2)})\phi^{(1)}p^{(1)}$ and $\bar{p}^{(2)} = v^{(2)}\phi^{(2)}p^{(2)}$, with $v^{(2)}$ being the total volume fraction of the fracture porosity and $\phi^{(1)}$ and $\phi^{(2)}$ being the matrix and fracture porosities, respectively. (Note that in this method of accounting for the void space, $\phi^{(2)} \equiv 1$.) These equations now account properly for inertia and elastic energy, strain, and stress, as well as for the specified types of dissipation mechanisms, and are in complete agreement with those developed by Tuncay and Corapcioglu (1996) using a different approach. In (4), the parts of the equation not involving the kinetic energy can be shown to be equivalent to a two-fluid Darcy's law in this context, so b_{12} and b_{13} are related to Darcy's constants for two single phase flow and b_{23} is the small coupling coefficient. Explicit relations between the b 's and the appropriate permeabilities [see Eqs. (53) and (54) of Berryman and Wang (1995)] are not difficult to establish. The harder part of the analysis concerns

the constitutive equations required for the right hand side of (3). After the following section on inertia and drag, the remainder of the paper will necessarily be devoted to addressing some of these issues concerning stress-strain relations.

In summary, equations (3) and (4) can be combined into

$$\begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{12} & \rho_{22} & \rho_{23} \\ \rho_{13} & \rho_{23} & \rho_{33} \end{pmatrix} \begin{pmatrix} \ddot{u}_i \\ \dot{U}_i^{(1)} \\ \dot{U}_i^{(2)} \end{pmatrix} + \begin{pmatrix} b_{12} + b_{13} & -b_{12} & -b_{13} \\ -b_{12} & b_{12} + b_{23} & -b_{23} \\ -b_{13} & -b_{23} & b_{13} + b_{23} \end{pmatrix} \begin{pmatrix} \dot{u}_i \\ \dot{U}_i^{(1)} \\ \dot{U}_i^{(2)} \end{pmatrix} = \begin{pmatrix} \tau_{ij,j} \\ -\bar{p}_{,i}^{(1)} \\ -\bar{p}_{,i}^{(2)} \end{pmatrix}, \quad (5)$$

showing the coupling between the solid and both types of fluid components.

In the next section we show how to relate the inertial and drag coefficients to physically measurable quantities.

INERTIAL AND DRAG COEFFICIENTS

Inertial coefficients

It is easy to understand that the inertial coefficients appearing in the kinetic energy T must depend on the densities of solid and fluid constituents ρ_s and ρ_f , and also on the volume fractions $v^{(1)}$, $v^{(2)}$ and porosities $\phi^{(1)}$, $\phi^{(2)}$ of the matrix material and fractures, respectively. The total porosity is given by $\phi = v^{(1)}\phi^{(1)} + v^{(2)}\phi^{(2)}$ and the volume fraction occupied by the solid material is therefore $1 - \phi$.

For a single porosity material, there are only three inertial coefficients and the kinetic energy can be written as

$$2T = (\dot{\mathbf{u}} \quad \dot{\mathbf{U}}) \begin{pmatrix} \bar{\rho}_{11} & \bar{\rho}_{12} \\ \bar{\rho}_{12} & \bar{\rho}_{22} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{U}} \end{pmatrix}, \quad (6)$$

where $\dot{\mathbf{U}}$ is the velocity of the only fluid present. Then, it is easy to see that, if $\dot{\mathbf{u}} = \dot{\mathbf{U}}$, the total inertia $\bar{\rho}_{11} + 2\bar{\rho}_{12} + \bar{\rho}_{22}$ must equal the total inertia present in the system $(1 - \phi)\rho_s + \phi\rho_f$. Furthermore, Biot (1956) has shown that $\bar{\rho}_{11} + \bar{\rho}_{12} = (1 - \phi)\rho_s$, and that $\bar{\rho}_{22} + \bar{\rho}_{12} = \phi\rho_f$. These three equations are not linearly independent and therefore do not determine the three coefficients. So we make the additional assumption that $\bar{\rho}_{22} = \tau\phi\rho_f$, where τ (Note: This τ without subscripts should not be confused with the stress tensor introduced earlier in the paper.) was termed the structure factor by Biot (1956), but has more recently been termed the electrical tortuosity (Brown, 1980; Johnson *et al.*, 1982), since $\tau = \phi F$, where F is the electrical formation factor. Berryman (1980) has shown that

$$\tau = 1 + r \left(\frac{1}{\phi} - 1 \right), \quad (7)$$

follows from interpreting the coefficient $\bar{\rho}_{11}$ as resulting from the solid density plus the induced mass due to the oscillation of the solid in the surrounding fluid. Then, $\bar{\rho}_{11} = (1 - \phi)(\rho_s + r\rho_f)$, where r is a factor dependent on microgeometry that is expected to lie in the range $0 \leq r \leq 1$, with $r = \frac{1}{2}$ for spherical grains. For example, if $\phi = 0.2$ and $r = 0.5$, equation (7) implies $\tau = 3.0$, which is a typical value for tortuosity of sandstones.

For double porosity, the kinetic energy may be written as

$$2T = (\dot{\mathbf{u}} \quad \dot{\mathbf{U}}^{(1)} \quad \dot{\mathbf{U}}^{(2)}) \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{12} & \rho_{22} & \rho_{23} \\ \rho_{13} & \rho_{23} & \rho_{33} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{U}}^{(1)} \\ \dot{\mathbf{U}}^{(2)} \end{pmatrix}. \quad (8)$$

We now consider some limiting cases: First, suppose that all the solid and fluid material moves in unison. Then, in complete analogy to the single porosity case, we have the result that $\rho_{11} + \rho_{22} + \rho_{33} + 2\rho_{12} + 2\rho_{13} + 2\rho_{23}$ must equal the total inertia of the system $(1 - \phi)\rho_s + \phi\rho_f$. Next, if we suppose that the two fluids can be made to move in unison, but independently of the solid, then we can take $\dot{\mathbf{U}} = \dot{\mathbf{U}}^{(1)} = \dot{\mathbf{U}}^{(2)}$, and telescope the expression for the kinetic energy to

$$2T = (\dot{\mathbf{u}} \quad \dot{\mathbf{U}}) \begin{pmatrix} \rho_{11} & (\rho_{12} + \rho_{13}) \\ (\rho_{12} + \rho_{13}) & (\rho_{22} + 2\rho_{23} + \rho_{33}) \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{U}} \end{pmatrix}. \quad (9)$$

We can now relate the matrix elements in (9) directly to the barred matrix elements appearing in (6), which then gives us three equations for our six unknowns. Again these three equations are not linearly independent, so we still need four more equations.

Next we consider the possibility that the fracture fluid can oscillate independently of the solid and the matrix fluid, and furthermore that the matrix fluid velocity is locked to that of the solid so that $\dot{\mathbf{u}} = \dot{\mathbf{U}}^{(1)}$. For this case, the kinetic energy telescopes in a different way to

$$2T = (\dot{\mathbf{u}} \quad \dot{\mathbf{U}}^{(2)}) \begin{pmatrix} (\rho_{11} + 2\rho_{12} + \rho_{22}) & (\rho_{13} + \rho_{23}) \\ (\rho_{13} + \rho_{23}) & \rho_{33} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{U}}^{(2)} \end{pmatrix}. \quad (10)$$

This equation is also of the form (6), but we must be careful to account properly for the parts of the system included in the matrix elements. Now we treat the solid and matrix fluid as a single unit, so

$$\rho_{11} + 2\rho_{12} + \rho_{22} = (1 - \phi)\rho_s + (1 - v^{(2)})\phi^{(1)}\rho_f + (\tau^{(2)} - 1)v^{(2)}\phi^{(2)}\rho_f, \quad (11)$$

$$\rho_{13} + \rho_{23} = -(\tau^{(2)} - 1)v^{(2)}\phi^{(2)}\rho_f, \quad (12)$$

and

$$\rho_{33} = \tau^{(2)}v^{(2)}\phi^{(2)}\rho_f, \quad (13)$$

where $\tau^{(2)}$ is the tortuosity of fracture porosity alone and $v^{(2)}$ is the volume fraction of the fractures in the system.

Finally, we consider the possibility that the matrix fluid can oscillate independently of the solid and the fracture fluid, and furthermore that the fracture fluid velocity is locked to that of

the solid so that $\dot{\mathbf{u}} = \dot{\mathbf{U}}^{(2)}$. The kinetic energy telescopes in a very similar way to the previous case with the result

$$2T = (\dot{\mathbf{u}} \quad \dot{\mathbf{U}}^{(1)}) \begin{pmatrix} (\rho_{11} + 2\rho_{13} + \rho_{33}) & (\rho_{12} + \rho_{23}) \\ (\rho_{12} + \rho_{23}) & \rho_{22} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{U}}^{(1)} \end{pmatrix}. \quad (14)$$

We imagine that this thought experiment amounts to analyzing the matrix material alone without fractures being present. The equations resulting from this identification are completely analogous to those in (11)-(13), so we will not show them explicitly here.

We now have nine equations in the six unknowns and six of these are linearly independent, so the system can be solved. The result of this analysis is that the off-diagonal terms are given by

$$2\rho_{12}/\rho_f = (\tau^{(2)} - 1)v^{(2)}\phi^{(2)} - (\tau^{(1)} - 1)(1 - v^{(2)})\phi^{(1)} - (\tau - 1)\phi, \quad (15)$$

$$2\rho_{13}/\rho_f = (\tau^{(1)} - 1)(1 - v^{(2)})\phi^{(1)} - (\tau^{(2)} - 1)v^{(2)}\phi^{(2)} - (\tau - 1)\phi, \quad (16)$$

and

$$2\rho_{23}/\rho_f = (\tau - 1)\phi - (\tau^{(1)} - 1)(1 - v^{(2)})\phi^{(1)} - (\tau^{(2)} - 1)v^{(2)}\phi^{(2)}. \quad (17)$$

The diagonal terms are given by

$$\rho_{11} = (1 - \phi)\rho_s + (\tau - 1)\phi\rho_f, \quad (18)$$

$$\rho_{22} = \tau^{(1)}(1 - v^{(2)})\phi^{(1)}\rho_f, \quad (19)$$

and ρ_{33} is given by (13).

Estimates of the three tortuosities τ , $\tau^{(1)}$, and $\tau^{(2)}$ may be obtained using (7), or direct measurements may be made using electrical methods as advocated by Brown (1980) and Johnson *et al.* (1982). Appendix A explains one method of estimating τ for the whole medium when the constituent tortuosities and volume fractions are known.

Drag coefficients

The drag coefficients may be determined by first noting that the equations presented here reduce to those of Berryman and Wang (1995) in the low frequency limit by merely neglecting the inertial terms. What is required to make the direct identification of the coefficients is a pair of coupled equations for the two increments of fluid content $\zeta^{(1)}$ and $\zeta^{(2)}$. These quantities are related to the displacements by $\zeta^{(1)} = -(1 - v^{(2)})\phi^{(1)}\nabla \cdot (\mathbf{U}^{(1)} - \mathbf{u})$ and $\zeta^{(2)} = -v^{(2)}\phi^{(2)}\nabla \cdot (\mathbf{U}^{(2)} - \mathbf{u})$.

The pertinent equations from Berryman and Wang (1995) are

$$\eta \begin{pmatrix} \dot{\zeta}^{(1)} \\ \dot{\zeta}^{(2)} \end{pmatrix} = \begin{pmatrix} k^{(11)} & k^{(12)} \\ k^{(21)} & k^{(22)} \end{pmatrix} \begin{pmatrix} P_{,ii}^{(1)} \\ P_{,ii}^{(2)} \end{pmatrix}, \quad (20)$$

where η is the shear viscosity of the fluid, the $k^{(ij)}$ are permeabilities including possible cross-coupling terms. The pressures appearing here are the actual pore pressures in the storage and fracture porosity. We can extract the terms we need from (5), and then take the divergence to obtain

$$\begin{pmatrix} 1/(1-v^{(2)})\phi^{(1)} & 0 \\ 0 & 1/v^{(2)}\phi^{(2)} \end{pmatrix} \begin{pmatrix} b_{12}+b_{23} & -b_{23} \\ -b_{23} & b_{13}+b_{23} \end{pmatrix} \begin{pmatrix} \nabla \cdot (\dot{\mathbf{U}}^{(1)} - \dot{\mathbf{u}}) \\ \nabla \cdot (\dot{\mathbf{U}}^{(2)} - \dot{\mathbf{u}}) \end{pmatrix} = - \begin{pmatrix} p_{,ii}^{(1)} \\ p_{,ii}^{(2)} \end{pmatrix}. \quad (21)$$

Comparing these two sets of equations and solving for the b coefficients, we find

$$b_{12} = \frac{\eta(1-v^{(2)})\phi^{(1)} [(1-v^{(2)})\phi^{(1)}k^{(22)} - v^{(2)}\phi^{(2)}k^{(21)}]}{k^{(11)}k^{(22)} - k^{(12)}k^{(21)}}, \quad (22)$$

$$b_{13} = \frac{\eta v^{(2)}\phi^{(2)} [v^{(2)}\phi^{(2)}k^{(11)} - (1-v^{(2)})\phi^{(1)}k^{(12)}]}{k^{(11)}k^{(22)} - k^{(12)}k^{(21)}}, \quad (23)$$

and

$$b_{23} = \frac{\eta v^{(2)}(1-v^{(2)})\phi^{(1)}\phi^{(2)}k^{(21)}}{k^{(11)}k^{(22)} - k^{(12)}k^{(21)}} = \frac{\eta v^{(2)}(1-v^{(2)})\phi^{(1)}\phi^{(2)}k^{(12)}}{k^{(11)}k^{(22)} - k^{(12)}k^{(21)}}. \quad (24)$$

For wave propagation, it will often be adequate to assume that the cross-coupling vanishes, as this effect is presumably more important for long term drainage of fluids than it is for short term propagation of waves. When this approximation is satisfactory, we have $b_{23} = 0$, and

$$b_{12} = \frac{\eta(1-v^{(2)})^2(\phi^{(1)})^2}{k^{(11)}} \quad (25)$$

and

$$b_{13} = \frac{\eta(v^{(2)}\phi^{(2)})^2}{k^{(22)}}, \quad (26)$$

which also provides a simple interpretation of these coefficients in terms of the porosities and diagonal permeabilities.

This completes the identification of the inertial and drag coefficients introduced in the previous section.

CONSTITUTIVE EQUATIONS

In previous work (Berryman and Wang, 1995) chose the external confining pressure, p_c , and the fluid pressures in the matrix, $p^{(1)}$, and in the fracture, $p^{(2)}$, to be the independent variables. The dependent variables were chosen to be the volumetric strain, e , and Biot's increment of fluid content (fluid volume accumulation per unit bulk volume) in the matrix, $\zeta^{(1)}$, and fracture, $\zeta^{(2)}$, separately. The phenomenological approach then relates each dependent variable linearly

to the independent variables. This choice of variables leads to a symmetric coefficient matrix because the scalar product of the dependent and independent variables is an energy density. The double-porosity theory with six independent coefficients, a_{ij} , for hydrostatic loading is a straightforward generalization of Biot's original equations.

$$\begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} -p_c \\ -p^{(1)} \\ -p^{(2)} \end{pmatrix} \quad (27)$$

The six coefficients occur in three classes, which correspond to the three original Biot coefficients. The coefficient $a_{11} = 1/K$ is an effective *compressibility* of the combined fracture-matrix system. The coefficients a_{12} and a_{13} are generalized *poroelastic expansion coefficients*, i.e., the ratio of bulk strain to matrix pressure and fracture pressure, respectively. The terms a_{22} , a_{23} , $a_{32} = a_{23}$, and a_{33} are generalized *storage coefficients*, i.e., a_{ij} is the volume of fluid that flows into a control volume (normalized by the control volume) of phase $i - 1$ due to a unit increase in fluid pressure in phase $j - 1$.

Formulas relating these parameters to properties of the constituents are summarized in TABLE 1, in which values are used for Berea sandstone from the results tabulated in TABLE 2. The definitions of the input parameters are: K and $K^{(1)}$ are the (jacketed) frame bulk moduli of the whole and the matrix respectively, K_s and $K_s^{(1)}$ are the unjacketed bulk moduli for the whole and the matrix, $\alpha = 1 - K/K_s$ and $\alpha^{(1)} = 1 - K^{(1)}/K_s^{(1)}$ are the corresponding Biot-Willis parameters, K_f is the pore fluid bulk modulus, $v^{(2)} = 1 - v^{(1)}$ is the total volume fraction of the fractures in the whole, and $B^{(1)}$ is Skempton's pore-pressure buildup coefficient for the matrix. Poisson's ratio and the porosity of the matrix are $\nu^{(1)}$ and $\phi^{(1)}$, respectively. It was observed by Berryman and Wang (1995) that the fluid-fluid coupling term a_{23} was small or negligible for the examples considered, and that it is expected to be small or negligible in most situations in which it makes sense to use the double-porosity model at all. Since our main goal for this paper is to extract and evaluate a somewhat simplified version of these formulas, from the more general analysis presented so far, we will therefore make the approximation in the remainder of this paper that $a_{23} \equiv 0$. This physically reasonable choice will also make the subsequent analysis somewhat less tedious.

We need to express the vector on the right hand side of (35) in terms of the macroscopic variables, using the constitutive relations in (27), plus the usual relations of linear elasticity. The basic set of equations for assumed isotropic media [analogous expressions for single-porosity with and without elastic anisotropy are given in Berryman (1998)] has the form

$$\begin{pmatrix} S_{11} & S_{12} & S_{12} & -\beta^{(1)} & -\beta^{(2)} \\ S_{12} & S_{11} & S_{12} & -\beta^{(1)} & -\beta^{(2)} \\ S_{12} & S_{12} & S_{11} & -\beta^{(1)} & -\beta^{(2)} \\ -\beta^{(1)} & -\beta^{(1)} & -\beta^{(1)} & a_{22} & a_{23} \\ -\beta^{(2)} & -\beta^{(2)} & -\beta^{(2)} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} \tau_{11} \\ \tau_{22} \\ \tau_{33} \\ -p^{(1)} \\ -p^{(2)} \\ \tau_{23} \\ \tau_{31} \\ \tau_{12} \end{pmatrix} = \begin{pmatrix} e_{11} \\ e_{22} \\ e_{33} \\ -\zeta^{(1)} \\ -\zeta^{(2)} \\ e_{23} \\ e_{31} \\ e_{12} \end{pmatrix}, \quad (28)$$

where the S_{ij} 's are the usual drained elastic compliances, and the β 's are poroelastic expansion coefficients approximately of the form $\beta = \alpha/3K$, where α is the Biot-Willis parameter (Biot-Willis, 1957) for single-porosity and K is the drained bulk modulus. We will not show our work here, but it is not hard to derive the following three relations:

$$\tau_{ij,j} = (\lambda + \mu)e_{,i} + \mu u_{i,jj} - 3K \left[\beta^{(1)} p_{,i}^{(1)} + \beta^{(2)} p_{,i}^{(2)} \right], \quad (29)$$

$$-3K \left[\beta^{(1)} p_{,i}^{(1)} + \beta^{(2)} p_{,i}^{(2)} \right] = -p_{c,i} - K e_{,i}, \quad (30)$$

and

$$-3 \left[\beta^{(1)} p_{,i}^{(1)} + \beta^{(2)} p_{,i}^{(2)} \right] = B^{(1)} \left[-\zeta_i^{(1)} - 3\beta^{(1)} p_{c,i} \right] + B^{(2)} \left[-\zeta_i^{(2)} - 3\beta^{(2)} p_{c,i} \right]. \quad (31)$$

Appearing in (29) are λ and μ , which are the Lamé parameters for the drained medium. A linear combination of the last two equations can be found to eliminate the appearance of $p_{c,i}$, and then this result can be substituted into (7) to show that

$$\tau_{ij,j} = \left(K_u + \frac{1}{3}\mu \right) e_{,i} + \mu u_{i,jj} + K_u \left[B^{(1)} \zeta_i^{(1)} + B^{(2)} \zeta_i^{(2)} \right], \quad (32)$$

where

$$K_u = \frac{K}{1 - 3K \left[\beta^{(1)} B^{(1)} + \beta^{(2)} B^{(2)} \right]} \quad (33)$$

is the undrained bulk modulus for the double porosity medium — specifically, it is the undrained bulk modulus for intermediate time scales, *i.e.*, undrained at the representative elementary volume (REV) scale but equilibrated between pore and fracture porosity locally. (This statement is consistent with our assumption that $a_{23} \equiv 0$, but needs some qualification if $a_{23} \neq 0$.)

Combining (32) with (27) and taking the divergence, we finally obtain the expression we need:

$$\begin{pmatrix} \tau_{ij,j} \\ -p_{,ii}^{(1)} \\ -p_{,ii}^{(2)} \end{pmatrix} = \begin{pmatrix} K_u + \frac{4}{3}\mu & B^{(1)} K_u & B^{(2)} K_u \\ -a_{12} a_{33} / D & (a_{11} a_{33} - a_{13}^2) / D & a_{12} a_{13} / D \\ -a_{13} a_{22} / D & a_{12} a_{13} / D & (a_{11} a_{22} - a_{12}^2) / D \end{pmatrix} \begin{pmatrix} e_{,ii} \\ -\zeta_{,ii}^{(1)} \\ -\zeta_{,ii}^{(2)} \end{pmatrix}, \quad (34)$$

where $D = a_{11} a_{22} a_{33} - a_{12}^2 a_{33} - a_{13}^2 a_{22}$.

THE DISPERSION RELATION AND ITS SOLUTION

It is now possible to write down and solve the dispersion relation for waves propagating through the double-porosity medium that we have been developing in the previous sections.

Derivation of the dispersion relation

We will first take a Fourier transform of (5) in the time domain, equivalent to assuming a time dependence of the form $\exp(-i\omega t)$. (Strictly speaking we should now introduce new notation for the variables that follow to account for the differences between the time-dependent coefficients and the Fourier coefficients. But we will not refer further to the time-dependent coefficients in this paper, so no confusion should arise if we use the same notation from now on for the Fourier coefficients.) Then, (5) becomes

$$-\omega^2 \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{pmatrix} \begin{pmatrix} u_i \\ U_i^{(1)} \\ U_i^{(2)} \end{pmatrix} = \begin{pmatrix} \tau_{ij,j} \\ -\bar{p}_{,i}^{(1)} \\ -\bar{p}_{,i}^{(2)} \end{pmatrix}, \quad (35)$$

where

$$\begin{aligned} q_{11} &= \rho_{11} + \frac{i}{\omega}(b_{12} + b_{13}), \\ q_{12} &= \rho_{12} - \frac{i}{\omega}b_{12}, \quad \text{etc.} \end{aligned} \quad (36)$$

It is also convenient to notice that

$$\frac{\partial}{\partial x_i} \begin{pmatrix} u_i \\ U_i^{(1)} \\ U_i^{(2)} \end{pmatrix} = \begin{pmatrix} e \\ U_{i,i}^{(1)} \\ U_{i,i}^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & & \\ 1 & \frac{1}{(1-v^{(2)})\phi^{(1)}} & \\ 1 & & \frac{1}{v^{(2)}\phi^{(2)}} \end{pmatrix} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix} \equiv \mathbf{R} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix}, \quad (37)$$

which will permit us to write the final equation in terms of the macroscopic strain and fluid contents e , $\zeta^{(1)}$, and $\zeta^{(2)}$. The final equality in (37) defines the matrix \mathbf{R} , which we need again later in the analysis.

Taking the divergence of (35), then substituting (37) and (34), and finally taking the spatial Fourier transform (having wavenumber k) gives the complex eigenvalue problem associated with wave propagation:

$$\begin{pmatrix} K_u + \frac{4}{3}\mu & B^{(1)}K_u & B^{(2)}K_u \\ -a_{12}a_{33}/D & (a_{11}a_{33} - a_{13}^2)/D & a_{12}a_{13}/D \\ -a_{13}a_{22}/D & a_{12}a_{13}/D & (a_{11}a_{22} - a_{12}^2)/D \end{pmatrix} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix} = v^2(\omega) \begin{pmatrix} 1 & & \\ \frac{1}{(1-v^{(2)})\phi^{(1)}} & & \\ & \frac{1}{v^{(2)}\phi^{(2)}} & \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{pmatrix} \begin{pmatrix} 1 & & \\ 1 & \frac{1}{(1-v^{(2)})\phi^{(1)}} & \\ & & \frac{1}{v^{(2)}\phi^{(2)}} \end{pmatrix} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix}, \quad (38)$$

where the eigenvalue $v^2(\omega) = \omega^2/k^2$ has the physical significance of being the square of the complex wave velocity. With obvious definitions for the matrices \mathbf{A} , \mathbf{P} , and \mathbf{Q} , while \mathbf{R} was previously defined in (37), we rewrite (38) as

$$\mathbf{A} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix} = v^2(\omega)\mathbf{PQR} \begin{pmatrix} e \\ -\zeta^{(1)} \\ -\zeta^{(2)} \end{pmatrix}, \quad (39)$$

and then, in terms of these matrices, the dispersion relation determining $v^2(\omega)$ at all angular frequencies ω is

$$\det(\mathbf{A} - v^2(\omega)\mathbf{PQR}) = 0. \quad (40)$$

This is a 3×3 determinant of complex numbers that must be solved for v^2 . A method for finding the three solutions is discussed in the next subsection.

Solution of the dispersion relation

A variety of numerical methods may be used to solve (40), including for example Crout's reduction method (Crout, 1941). However, since the system is relatively small (3×3) and since our purposes include gaining some physical insight into the processes involved, it will prove instructive to do some more analysis on the problem prior to the ultimate numerical calculations.

First, note that our analysis will be considerably simplified by an obvious rearrangement of the determinant (40) so that

$$\det(\mathbf{G} - v^2\mathbf{I}) = 0, \quad (41)$$

where \mathbf{I} is the identity matrix and

$$\mathbf{G} = \mathbf{AR}^{-1}\mathbf{Q}^{-1}\mathbf{P}^{-1}, \quad (42)$$

having complex matrix elements g_{ij} , for $i, j = 1, 2, 3$. The matrix \mathbf{R} was given previously in (37) and is clearly invertible for finite values of the two porosities. Matrix \mathbf{Q} is also always invertible for realistic choices of material parameters.

Using the properties of determinants, it is not hard to show that the 3×3 determinant (41) can be reduced to a 2×2 determinant in either of two convenient forms:

$$\begin{aligned} & \det \begin{pmatrix} (g_{11} - v^2)(g_{22} - v^2) - g_{12}g_{21} & -1 \\ g_{12}g_{23}g_{31} + g_{13}g_{32}g_{21} - (g_{22} - v^2)g_{13}g_{31} - (g_{11} - v^2)g_{23}g_{32} & g_{33} - v^2 \end{pmatrix} \\ &= \det \begin{pmatrix} (g_{11} - v^2)(g_{33} - v^2) - g_{13}g_{31} & -1 \\ g_{12}g_{23}g_{31} + g_{13}g_{32}g_{21} - (g_{33} - v^2)g_{12}g_{21} - (g_{11} - v^2)g_{23}g_{32} & g_{22} - v^2 \end{pmatrix} \\ &= 0. \end{aligned} \quad (43)$$

It is useful to write the determinant in each of these two ways in order to make connections with single-porosity models. The first version in (43) has as its upper left-hand element an expression that corresponds precisely to the determinant for the 2×2 system when the only porosity present is the matrix (or storage) porosity. Similarly, the second version in (43) has as its element in the upper left-hand position a term that corresponds to the determinant for the 2×2 system when the only porosity present is the fracture/crack porosity. In each case the remaining terms determine the effect of coupling to the second type of porosity. Depending on other system parameters such as permeabilities and porosities, either of these limits may

be useful to consider, and may also provide good starting points for the iterative (Newton-Raphson) procedure that we will use to solve the determinant equation for the complex velocity v .

If we express the determinant (41) as a polynomial $D(x)$ where $x = v^2$, then

$$\begin{aligned} D(x) = & -x^3 + (g_{11} + g_{22} + g_{33})x^2 \\ & - (g_{22}g_{33} + g_{33}g_{11} + g_{11}g_{22} \\ & - g_{23}g_{32} - g_{13}g_{31} - g_{12}g_{21})x \\ & + (g_{11}g_{22}g_{33} - g_{23}g_{32}g_{11} - g_{13}g_{31}g_{22} \\ & - g_{12}g_{21}g_{33} + g_{12}g_{23}g_{31} + g_{13}g_{32}g_{21}) = 0. \end{aligned} \quad (44)$$

The Newton-Raphson method (Hildebrand, 1956) for solving this equation for x is an iteration process starting from some initial choice $x^{(0)}$ and computing

$$x^{(i)} = x^{(i-1)} - D(x^{(i-1)})/D'(x^{(i-1)}), \quad \text{for } i = 1, 2, \dots, N_c \quad (45)$$

(where D' is the first derivative of D with respect to x) until some convergence criterion has been met at $i = N_c$. The fact that the coefficients and the solution of the problem are complex adds no special complication to this procedure. However, since the polynomial is complex, it is important that good starting values $x^{(0)}$ be obtained for at least two of the three roots of (44), as a search procedure in the complex plane would be considerably more difficult to implement than is pure Newton-Raphson iteration.

From (43), we see that two choices of starting values for the complex parameter x are given by (for example)

$$x_{1,2}^{(0)} = \frac{1}{2} \left(g_{11} + g_{22} \pm \sqrt{(g_{11} - g_{22})^2 + 4g_{12}g_{21}} \right). \quad (46)$$

Once the Newton-Raphson iteration has converged from both of these starting values to their final values of x_1 and x_2 , then the third solution of the dispersion relation is obtained directly by recalling that

$$D(x) \equiv \sum_{n=0}^3 D_n (-x)^n = -(x - x_1)(x - x_2)(x - x_3). \quad (47)$$

Using the linear independence of the terms in powers of x in (47), we have three equations showing that

$$x_3 = D_2 - x_1 - x_2 = \frac{D_1 - x_1 x_2}{x_1 + x_2} = \frac{D_0}{x_1 x_2}. \quad (48)$$

Any one of these three formulas may be used to calculate x_3 directly, or combinations of them (such as the geometric mean of any two) may be used to reduce the errors that might be introduced by premature termination of the Newton-Raphson iteration process for x_1 and x_2 . An alternative method (but not the one we used in the examples of the present paper) is to use

the formulas of (48) in a different way to arrive at a quadratic formula for x_2 and x_3 , once that x_1 is known from the iteration procedure. The resulting formulas are given (for example) by

$$x_{2,3} = \frac{1}{2} \left[(D_2 - x_1) \pm \sqrt{(D_2 - x_1)^2 - 4D_0/x_1} \right]. \quad (49)$$

In this approach it is clearly preferable to solve for the fast wave solution as x_1 , and then use the formulas in (49) to obtain the two slow wave solutions since the fast wave solution will virtually always be well-behaved.

For each value of x that solves the dispersion equation, we can then compute the wave velocity and inverse of the quality factor Q using the definition

$$\frac{1}{\sqrt{x}} \equiv \frac{1}{v(\omega)} \equiv \frac{1}{v_r} \left(1 + \frac{i}{2Q} \right), \quad (50)$$

where the choice of square root is determined by the requirement of physical realizability as discussed in the following section. This definition of $1/Q$ is accurate when attenuation losses are low, but should be carefully interpreted for high loss situations (*e.g.*, low frequency diffusive modes). In particular, it is often stated that, when losses are high, (50) needs to be modified so that if

$$\frac{1}{v(\omega)} \equiv \frac{1}{v_r} \left(1 + \frac{i\alpha v_r}{\omega} \right), \quad (51)$$

where α is the attenuation coefficient (having units of inverse length), then [as Hamilton (1972) and Bourbié *et al.* (1987) show]

$$\frac{1}{Q} \equiv \frac{2\alpha v_r}{\omega - \alpha^2 v_r^2 / \omega}. \quad (52)$$

However, for diffusive modes the imaginary and real parts of the velocity are of comparable size, and therefore it is possible that the formula (52) will become singular at low frequencies for such modes. This complication can and does happen in practice. To avoid this complication, we use the definition (50) in all cases, and then interpret those situations in which $1/Q \rightarrow 2$ as an indication that the mode under consideration is actually diffusive rather than propagatory.

Physical realizability at higher frequencies

For the purposes of this paper, we will take (25) and (26) to be the low frequency limits of the drag coefficients and also assume that $b_{23} \equiv 0$. Then, the coefficients b_{12} and b_{13} must be modified at higher frequencies in order to assure that the theory as a whole preserves obvious physical requirements such as nonnegative dissipation for all modes at all times. (If the theory always predicts nonnegative dissipation, then we will say it is “realizable.” If the theory predicts negative dissipation for any of the modes of propagation, then the theory is not realizable, and further effort will be required to make the theory fully realizable.) This issue arises naturally when we have obtained the solution $x = v^2$ to (45) for any one of the three

compressional modes. Then, taking the complex square root, we get two roots that differ only by + and - signs. We want the solution that has both positive real velocity and a negative imaginary part. This is so because $k = \omega/v = k_0 + i\alpha$, where α should be a physical attenuation coefficient such that

$$\exp i(kz - \omega t) = \exp(-\alpha z) \exp i k_0(z - v_0 t) \quad (53)$$

leads to a decrease in the overall amplitude of the compressional mode. If, for any of the three compressional modes, no root exists with both positive real velocity and negative imaginary part, then the dispersion relation is unphysical and the approximations we have made in deriving it are suspect. In our examples to follow, we will tentatively take the results from Appendix B as the proper way to modify the drag coefficients at higher frequencies, but must always be careful to check that this choice does not lead to unphysical behavior.

EXAMPLE

The example we present here is for Berea sandstone saturated with water. The parameters used in the calculations are taken from TABLES 1 and 2. Most of the mechanical properties were obtained from measurements made by Coyner (1984). The permeability values for the matrix $k^{(1)}$ and the fractures $k^{(2)}$ are the same as those used by Lewallen and Wang (1998). The approach described in the preceding text, together with the results obtained in Appendices A and B, has been implemented by writing a Fortran code and computing the eigenvalues for the three compressional modes and their corresponding eigenvectors in the frequency range 10.0 Hz to 1 MHz. The results for the computed velocities and inverse quality factors are then displayed in Figures 1-6. The results for the eigenvectors will be described but not displayed here. We will not present results for the shear modes, but expect them to differ little from results for single-porosity calculations for the same material, since the pore fluid does not significantly affect the shear moduli in these models. (However, shear dispersion effects due to pore-fluid will still be of importance.) Figures 1 and 2 show that the first compressional wave is dispersive and has its main contributions to attenuation ($1/Q \simeq .001$) centered at about 3 kHz, with significant decrease in the attenuation envelope (by about an order of magnitude) at 100 Hz and 100 kHz. Wave velocity dispersion is localized approximately to the frequency range 1 kHz to 10 kHz, and the total dispersion is less than 1 %. The eigenvector for this mode shows that the storage pore fluid is essentially moving in concert with the solid frame throughout the frequency range considered, with some small but largely negligible deviations above 1 kHz. On the other hand, the fracture pore fluid oscillates out of phase with respect to the solid frame with an amplitude as much as about one half that of the frame amplitude above about 10 kHz. The observed dissipation for this mode is clearly tied to the out of phase motion of the fracture fluid.

Figures 3 and 4 show that the second compressional mode is diffusive at low frequencies, but becomes propagatory with a $Q \simeq 5$ or greater at about 10 kHz. The wave speed is quite small at these higher frequencies (about 550 m/s), indicating that the wave is probably propagating mostly through some pore fluid along a tortuous path. The eigenvector analysis shows

that the storage fluid excitation is again quite small compared to that of the fracture fluid, although it is about two orders of magnitude larger than that observed for the first compressional wave. The main effect observed for the second compressional wave is a large oscillation of the fracture fluid relative to the solid frame, so that this mode can be properly characterized in this example as a slow wave (in the sense of single-porosity poroelasticity) through the fracture fluid.

Figures 5 and 6 show that the third compressional mode is diffusive at all frequencies in the range considered. The apparent velocity is much lower (less than 100 m/s) — even than that of the second compressional mode. The eigenvector analysis for this mode shows that both the storage fluid and the fracture fluid are oscillating with significant and comparable amplitudes, larger than that of the solid frame. The two fluids are also oscillating out of phase with each other. The amplitude of the storage fluid oscillation slightly dominates that of the fracture fluid, which partly explains the increased attenuation for this mode. For this example, we might characterize this mode as a slow wave through the storage fluid, but note that this interpretation is slightly over simplified.

Finally, we note that other examples have been computed in an attempt to verify the nature of the dependence of the mode parameters on the input parameters. It has been observed for example that the second and third compressional wave “velocities” at low frequencies are proportional to the square root of the fluid bulk modulus as would be expected for a single-porosity slow wave (Berryman, 1981; Chandler and Johnson, 1981). In the single porosity case, the slow wave velocity is inversely proportional to the square root of the specific storage coefficient (Green and Wang, 1990). If the solid compressibility is small compared with that of the fluid, then the slow wave velocity would be proportional to the square root of the fluid bulk modulus.

DISCUSSION AND CONCLUSIONS

The physical realizability issue is an important one, but not one we have dwelt much on in this paper. Because of the complex nature of the dispersion equations and the existence of three compressional modes for this theory, it will require some substantial amount of effort to clarify the proper analytic structure for the theory so that unphysical results are not generated. The main issues here are whether or not it is always appropriate to use the results presented in Appendix B. It is known that these types of results are valid for single-porosity analysis, but when used together with some of the approximations we have made here — such as (for example) the neglect of cross-terms in the quasistatic permeability equations — may introduce some unphysical behavior at high frequencies for some combinations of parameters. We leave this analysis to future work.

Another important practical issue involves the careful comparison of these results with the extensive literature on wave propagation and attenuation in fractured earth materials. These comparisons will also be left to the future.

We conclude that the double-porosity dual-permeability analysis that has been presented

here has the capability to explain both wave propagation and attenuation in earth materials when the attenuation is due to out-of-phase motion of pore fluids in storage and fracture porosity. However, there remains quite a lot of work to do yet both on the theory and on its applications to real data before we can consider the story to be complete.

REFERENCES

- Berryman, J. G., 1980, Confirmation of Biot's theory: *Appl. Phys. Lett.*, **37**, 382–384.
- Berryman, J. G., 1981, Elastic wave propagation in fluid-saturated porous media: *J. Acoust. Soc. Am.*, **69**, 416–424.
- Berryman, J. G., 1995, Mixture theories for rock properties: in *Rock Physics and Phase Relations: A Handbook of Physical Constants*: edited by T. J. Ahrens AGU, Washington, D. C., pp. 205–228.
- Berryman, J. G., 1998, Transversely isotropic poroelasticity arising from thin isotropic layers: in *Mathematics of Multiscale Materials*: K. M. Golden, G. R. Grimmett, R. D. James, G. W. Milton, and P. N. Sen (editors), Springer-Verlag, New York, pp. 37–50.
- Berryman, J. G., and H. F. Wang, 1995, The elastic coefficients of double-porosity models for fluid transport in jointed rock: *J. Geophys. Res.*, **100**, 24611–24627.
- Biot, M. A., 1941, General theory of three dimensional consolidation: *J. Appl. Phys.*, **12**, 155–164.
- Biot, M. A., 1956a, Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range: *J. Acoust. Soc. Am.*, **28**, 168–178.
- Biot, M. A., 1956b, Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range: *J. Acoust. Soc. Am.*, **28**, 179–191.
- Biot, M. A., 1962, Mechanics of deformation and acoustic propagation in porous media: *J. Appl. Phys.*, **33**, 1482–1498.
- Biot, M. A., and D. G. Willis, 1957, The elastic coefficients of the theory of consolidation: *J. App. Mech.*, **24**, 594–601.
- Bourbié, T., O. Coussy, and B. Zinszner, 1987, *Acoustics of Porous Media*: Gulf, Houston, p. 113.
- Brown, R. J. S., 1980, Connection between formation factor for electrical-resistivity and fluid-solid coupling factor in Biot equations for acoustic-waves in fluid-filled porous media: *Geophysics*, **45**, 1269–1275.
- Brown, R. J. S., and J. Korringa, 1975, On the dependence of the elastic properties of a porous rock on the compressibility of a pore fluid: *Geophysics*, **40**, 608–616.

- Budiansky, B., and R. J. O'Connell, 1976, Elastic moduli of a cracked solid: *Int. J. Solids Struct.*, **12**, 81–97.
- Carroll, M. M., 1980, Mechanical response of fluid-saturated porous materials: in *Theoretical and Applied Mechanics*: F. P. J. Rimrott and B. Tabarrok (eds.), Proceedings of the 15th International Congress of Theoretical and Applied Mechanics, Toronto, August 17–23, 1980, North-Holland, Amsterdam, pp. 251–262.
- Chandler, R. N., and D. L. Johnson, 1981, The equivalence of quasistatic flow in fluid-saturated porous media and Biot's slow wave in the limit of zero frequency: *J. Appl. Phys.*, **52**, 3391–3395.
- Chin, R. C. Y., J. G. Berryman, and G. W. Hedstrom, 1985, Generalized ray expansion for pulse-propagation and attenuation in fluid-saturated porous media: *Wave Motion*, **7**, 43–65.
- Corapcioglu, M. Y., and K. Tuncay, 1996, Propagation of waves in porous media: in *Advances in Porous Media 3*: edited by M. Y. Corapcioglu, Elsevier, Amsterdam, pp. 361–440.
- Corapcioglu, M. Y., and K. Tuncay, 1998, Wave propagation in fractured porous media saturated by two immiscible fluids: in *Poromechanics: A Tribute to Maurice A. Biot*: edited by J.-F. Thimus, Y. Abousleiman, A. H.-D. Cheng, O. Coussy, and E. Detournay, Balkema, Rotterdam, pp. 197–203.
- Coyner, K. B., 1984, *Effects of Stress, Pore Pressure, and Pore Fluids on Bulk Strain, Velocity, and Permeability of Rocks*: Ph. D. Thesis, Massachusetts Institute of Technology.
- Crout, P. D., 1941, A short method for evaluating determinants and solving systems of linear equations with real or complex coefficients: *Trans. AIEE*, **60**, 1235–1240.
- Dvorkin, J., and A. Nur, 1993, Dynamic poroelasticity: A unified model with the squirt and the Biot mechanisms: *Geophysics*, **58**, 524–533.
- Gassmann, F., 1951, Über die elastizität poröser medien: *Veierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **96**, 1–23.
- Geertsma, J., 1957, The effect of fluid pressure decline on volumetric changes of porous rocks: *Trans. AIME*, **210**, 331–340.
- Green, D. H., and H. F. Wang, 1990, Specific storage as a poroelastic coefficient: *Water Resources Res.*, **26**, 1631–1637.
- Hamilton, E. L., 1972, Compressional-wave attenuation in marine sediments: *Geophysics*, **37**, 620–646.
- Hashin, Z., and S. Shtrikman, 1962, A variational approach to the theory of the effective magnetic permeability of multiphase materials: *J. Appl. Phys.*, **33**, 3125–3131.
- Hildebrand, F. B., 1956, *Introduction to Numerical Analysis*: Dover, New York, Section 10.11.

- Johnson, D. L., J. Koplik, and R. Dashen, 1987, Theory of dynamic permeability and tortuosity in fluid-saturated porous-media: *J. Fluid Mech.*, **176**, 379–402.
- Johnson, D. L., J. Koplik, and L. M. Schwartz, 1986, New pore-size parameter characterizing transport in porous media: *Phys. Rev. Lett.*, **57**, 2564–2567.
- Johnson, D. L., T. J. Plona, C. Scala, F. Pasierb, and H. Kojima, 1982, Tortuosity and acoustic slow waves: *Phys. Rev. Lett.*, **49**, 1840–1844.
- Kümpel, H.-J., 1991, Poroelasticity: parameters reviewed: *Geophys. J. Int.*, **105**, 783–799.
- Lewallen, K. T., and H. F. Wang, 1998, Consolidation of a double-porosity medium: *Int. J. Solids Structures*, **35**, 4845–4867.
- Love, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity*: Dover, New York, pp. 173–174.
- Mavko, G., and D. Jizba, 1991, Estimating grain-scale fluid effects on velocity dispersion in rocks: *Geophysics*, **56**, 1940–1949.
- Mavko, G., and A. Nur, 1979, Wave attenuation in partially saturated rocks: *Geophysics*, **44**, 161–178.
- Miksis, M. J., 1988, Effects of contact line movement on the dissipation of waves in partially saturated rocks: *J. Geophys. Res.*, **93**, 6624–6634.
- O’Connell, R. J., and B. Budiansky, 1974, Seismic velocities in dry and saturated cracked solids: *J. Geophys. Res.*, **79**, 5412–5426.
- O’Connell, R. J., and B. Budiansky, 1977, Viscoelastic properties of fluid-saturated cracked solids: *J. Geophys. Res.*, **82**, 5719–5736.
- Plona, T. J., 1980, Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies: *Appl. Phys. Lett.*, **36**, 259–261.
- Pride, S., 1994, Governing equations for the coupled electromagnetics and acoustics of porous media: *Phys. Rev. B*, **50**, 15678–15696.
- Pride, S. R., and J. G. Berryman, 1998, Connecting theory to experiment in poroelasticity: *J. Mech. Phys. Solids*, **46**, 719–747.
- Pride, S. R., and F. D. Morgan, 1991, Electrokinetic dissipation induced by seismic-waves: *Geophysics*, **56**, 914–925.
- Rice, J. R., 1975, On the stability of dilatant hardening for saturated rock masses: *J. Geophys. Res.*, **80**, 1531–1536.
- Rice, J. R., and M. P. Cleary, 1976, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents: *Rev. Geophys. Space Phys.*, **14**, 227–241.

Skempton, A. W., 1954, The pore-pressure coefficients A and B: *Geotechnique*, **4**, 143–147.

Stoll, R. D., 1985, Marine sediment acoustics: *J. Acoust. Soc. Am.*, **77**, 1789–1799.

Thompson, A. H., and G. A. Gist, 1993, Geophysical applications of electrokinetic conversion: *The Leading Edge* **12**, 1169–1173.

Tuncay, K., and M. Y. Corapcioglu, 1996, Wave propagation in fractured porous media: *Transp. Porous Media*, **23**, 237–258.

Winkler, K. W., 1985, Dispersion analysis of velocity and attenuation in Berea sandstone: *J. Geophys. Res.*, **90**, 6793–6800.

APPENDIX A

Tortuosity for Double-Porosity Media

Theoretical estimates of tortuosity for the matrix and fracture components of the double-porosity medium may be obtained by noting that equation (7) implies

$$\tau^{(1)} = \frac{1}{2} \left(1 + \frac{1}{\phi^{(1)}} \right) \quad (54)$$

for storage porosity that is spherical in shape, while

$$\tau^{(2)} = 1, \quad (55)$$

for the fracture porosity, because $\phi^{(2)} = 1$ by assumption.

It is more difficult to estimate the overall tortuosity τ , but a physically reasonable value can be obtained by considering the Hashin-Shtrikman bounds on electrical conductivity of a composite medium (Hashin and Shtrikman, 1962). These bounds show that for a two-component medium the effective conductivity will lie between the values σ_{HS}^{\pm} given by the formula (Berryman, 1995)

$$\frac{1}{\sigma_{HS}^{\pm} + 2\sigma_m} = \frac{1 - v^{(2)}}{\sigma_1 + 2\sigma_m} + \frac{v^{(2)}}{\sigma_2 + 2\sigma_m}, \quad (56)$$

where

$$\sigma_m = \max(\sigma_1, \sigma_2) \quad \text{for} \quad \sigma_{HS}^+ \quad (57)$$

and

$$\sigma_m = \min(\sigma_1, \sigma_2) \quad \text{for} \quad \sigma_{HS}^-. \quad (58)$$

This notation means that σ_{HS}^+ is the upper bound, while σ_{HS}^- is the lower bound.

Recalling that electrical tortuosity is related to formation factor F by $\tau = \phi F$, where $F = \sigma_f / \sigma$, we find that the tortuosity bounds for the double-porosity medium are:

$$\frac{1}{\phi/\tau^\pm + 2/F_m} = \frac{1 - v^{(2)}}{\phi^{(1)}/\tau^{(1)} + 2/F_m} + \frac{v^{(2)}}{1 + 2/F_m}. \quad (59)$$

We will assume that the overall tortuosity of the fractured double-porosity medium is in fact dominated by the fractures, in which case it is appropriate to assume that the actual electrical conductivity will be close to the upper bound σ_{HS}^+ . In this case we choose $F_m = 1$ and, after rearranging the formula, we find

$$\tau \simeq \phi \frac{v^{(2)}\phi^{(1)} + (3 - v^{(2)})\tau^{(1)}}{(3 - 2v^{(2)})\phi^{(1)} + 2v^{(2)}\tau^{(1)}}. \quad (60)$$

Also, recall that the overall porosity is given by $\phi = (1 - v^{(2)})\phi^{(1)} + v^{(2)}$. The formula (60) is expected to be valid for situations in which $v^{(2)} \ll 1$, and then (60) reduces approximately to $\tau \simeq \tau^{(1)}$. For applications to media in which such an assumption is not valid, the bounds in (56) should generally be used instead of (60).

Another physical constraint imposed by our model is that, if the drag/permeability coupling terms b_{23} are neglected, then internal consistency of the theory may also require that $\rho_{23} \simeq 0$. Then, (17) can be used to solve for the effective τ that gives $\rho_{23} = 0$. Interestingly, the result in the limit $v^{(2)} \ll 1$ is again that $\tau \simeq \tau^{(1)}$. So these two approaches give very consistent results, and suggest that $\rho_{23} \simeq 0$ may also be a physically reasonable approximation in many situations.

APPENDIX B

Frequency dependent permeability, tortuosity, and viscosity

Johnson *et al.* (1987) have shown that, for a single-porosity medium, the frequency dependence of the dynamic permeability and tortuosity can be well-approximated by

$$k(\omega) = \frac{k_0}{\left[1 - 4ik_0^2\tau_\infty^2\rho_f\omega/\eta\Lambda^2\phi^2\right]^{\frac{1}{2}} - ik_0\tau_\infty\rho_f\omega/\eta\phi} \quad (61)$$

and

$$\tau(\omega) = \tau_\infty + \frac{i\eta\phi}{k_0\rho_f\omega} \left[1 - 4ik_0^2\tau_\infty^2\rho_f\omega/\eta\Lambda^2\phi^2\right]^{\frac{1}{2}}. \quad (62)$$

The symbol η stands for the fluid viscosity (in units of $kgm/m \cdot s$), while η/ρ_f is the kinematic viscosity (in units of m^2/s). The new symbols appearing in these formulas are the d.c. permeability k_0 , the high frequency tortuosity τ_∞ , and the lambda parameter Λ introduced by Johnson *et al.* (1986). The values of the high frequency tortuosity τ_∞ for double-porosity media were considered here in Appendix A. The d.c. permeabilities for double-porosity media have been considered in Berryman and Wang (1995) and in Lewallen and Wang (1998).

For single porosity media, Johnson *et al.* (1987) show that the lambda parameter approximately satisfies

$$\Lambda^2 = 8k_0 F = 8k_0 \tau_\infty / \phi. \quad (63)$$

For the present purposes, we will assume that this relation holds independently for the storage porosity and the fracture porosity. Then we have

$$\Lambda^{(1)} = [8k^{(11)} \tau^{(1)} / \phi^{(1)}]^{1/2} \quad (64)$$

and

$$\Lambda^{(2)} = [8k^{(22)} \tau^{(2)} / \phi^{(2)}]^{1/2} = [8k^{(22)}]^{1/2}. \quad (65)$$

Finally, we see that for the double-porosity medium, the corrections due to frequency dependence can be viewed alternatively as a frequency dependent viscosity, since equations (61) and (62) follow by assuming that

$$\eta(\omega) \equiv \eta (1 - i \Lambda^2 \rho_f \omega / 16\eta)^{1/2} \quad (66)$$

These corrections need to be made separately for the two types of pores. This interpretation of the frequency dependence as being associated specifically with the viscosity is the one advocated by Biot (1956b), and has some advantages in the present context as it makes it quite straightforward to determine what the corrections should be for the multiple porosity problem. Note that we used (63) to simplify the factors inside the square root in (66).

It is not yet clear how to generalize these expressions for the permeability coupling terms $k^{(12)}$, but our assumption following (24) that $b_{23} = 0$ has eliminated this potential problem from the present paper. Achieving an understanding of this issue will be one goal of our future work.

TABLE 1. Stress-strain parameters in double-porosity modeling as derived by Berryman and

Parameter	Formula	Berea Sandstone
a_{11} (GPa ⁻¹)	$1/K$	0.167
a_{12} (GPa ⁻¹)	$-\alpha^{(1)}K_s^{(1)}/K^{(1)}K_s$	-0.074
Wang (1995). a_{13} (GPa ⁻¹)	$-\alpha/K - a_{12}$	-0.068
a_{22} (GPa ⁻¹)	$v^{(1)}\alpha^{(1)}/B^{(1)}K^{(1)}$	0.144
a_{23} (GPa ⁻¹)	$-v^{(1)}\alpha^{(1)}/K^{(1)} - a_{12}$	0.001
a_{33} (GPa ⁻¹)	$v^{(2)}/K_f + v^{(1)}/K^{(1)} - (1 - 2\alpha)/K + 2a_{12}$	0.075
\bar{a}_{33} (GPa ⁻¹)	$a_{33} - v^{(2)}/K_f$	0.067

TABLE 2. Material Properties for Berea Sandstone and Water

Parameter	Berea Sandstone
K (GPa)	6.0 ^a
K_s (GPa)	39.0 ^a
α	0.85 ^a
$K^{(1)}$ (GPa)	10.0 ^a
$v^{(1)}$	0.15
$K_s^{(1)}$ (GPa)	39.0 ^a
$\alpha^{(1)}$	0.74 ^a
$\phi^{(1)}$	0.178 ^a
$B^{(1)}$	0.600
$v^{(2)}$	0.0178
$k^{(11)}$ (m ²)	1.0×10^{-16}
$k^{(22)}$ (m ²)	1.0×10^{-12}
	Water
K_f (GPa)	2.3
ρ_f (Kgm/m ³)	1000.0
η (Pa·s)	0.001

^aFrom Coyner (1984)

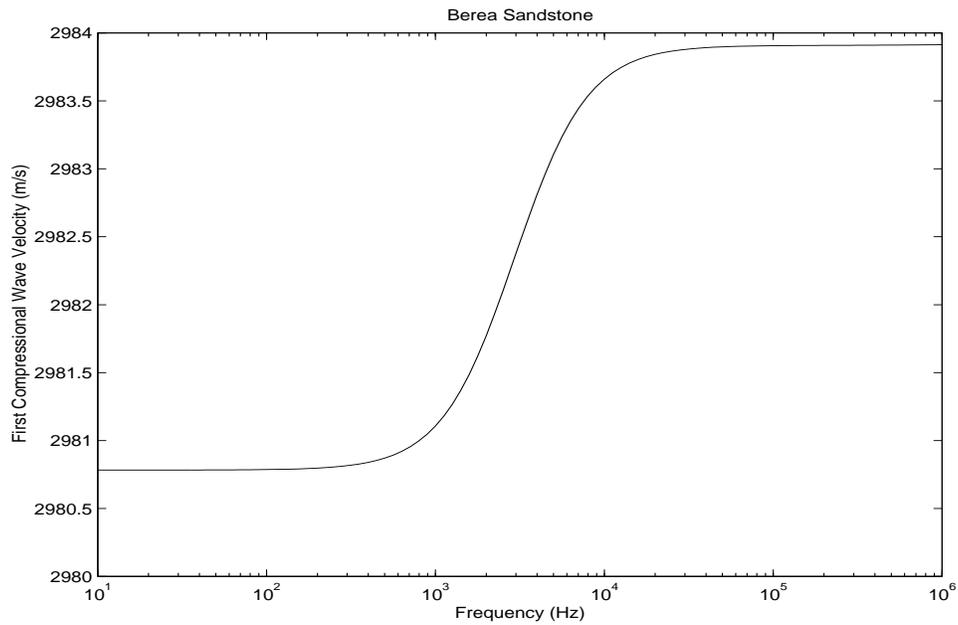


Figure 1: First compressional wave velocity as a function of frequency for fractured Berea sandstone. `jim1-vp` [NR]

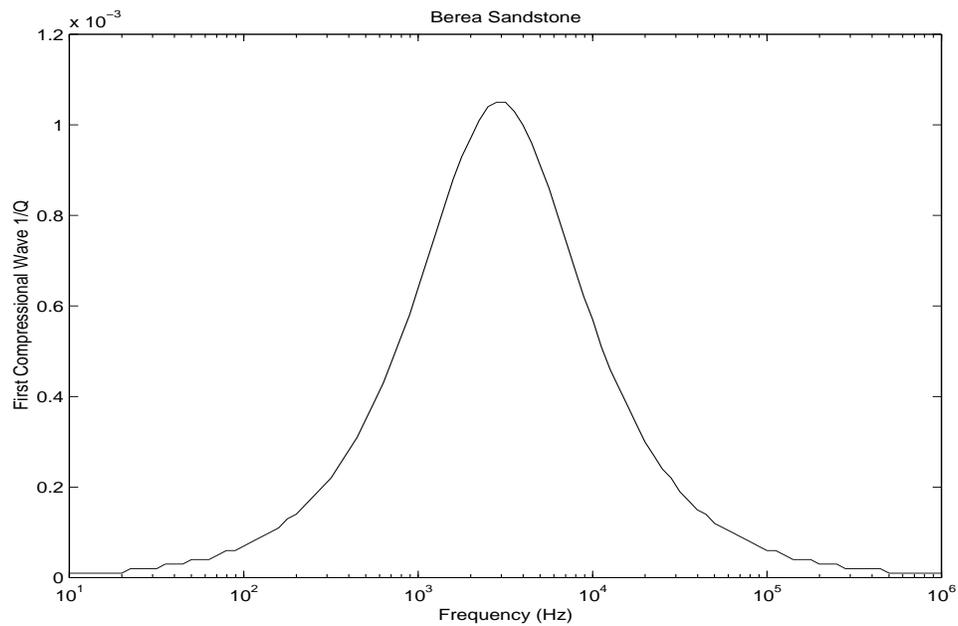


Figure 2: First compressional wave inverse quality factor $1/Q$ as a function of frequency for fractured Berea sandstone. `jim1-qp` [NR]

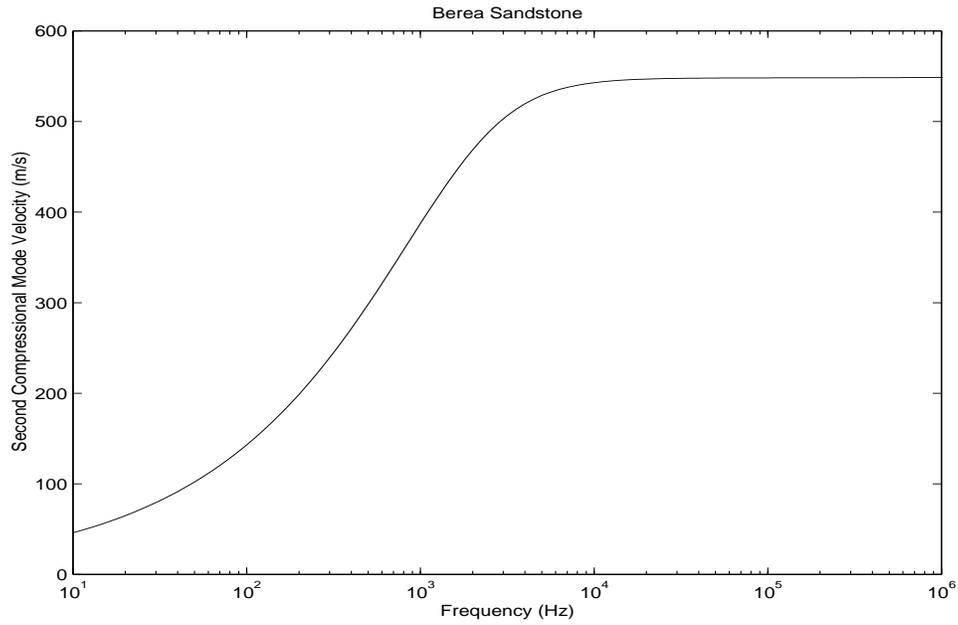


Figure 3: Second compressional wave velocity as a function of frequency for fractured Berea sandstone. jim1-vm [NR]

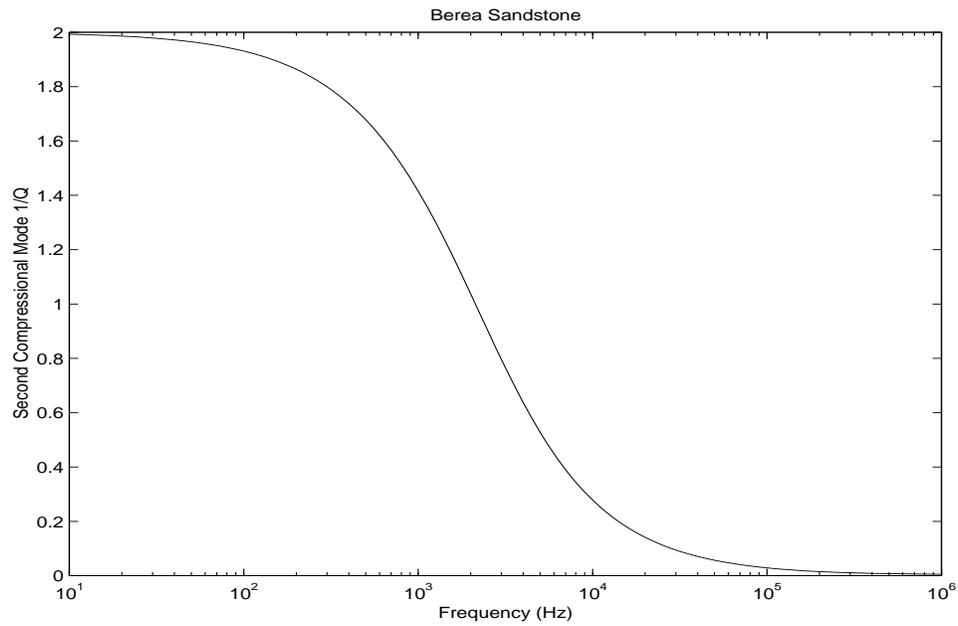


Figure 4: Second compressional wave inverse quality factor $1/Q$ as a function of frequency for fractured Berea sandstone. jim1-qm [NR]

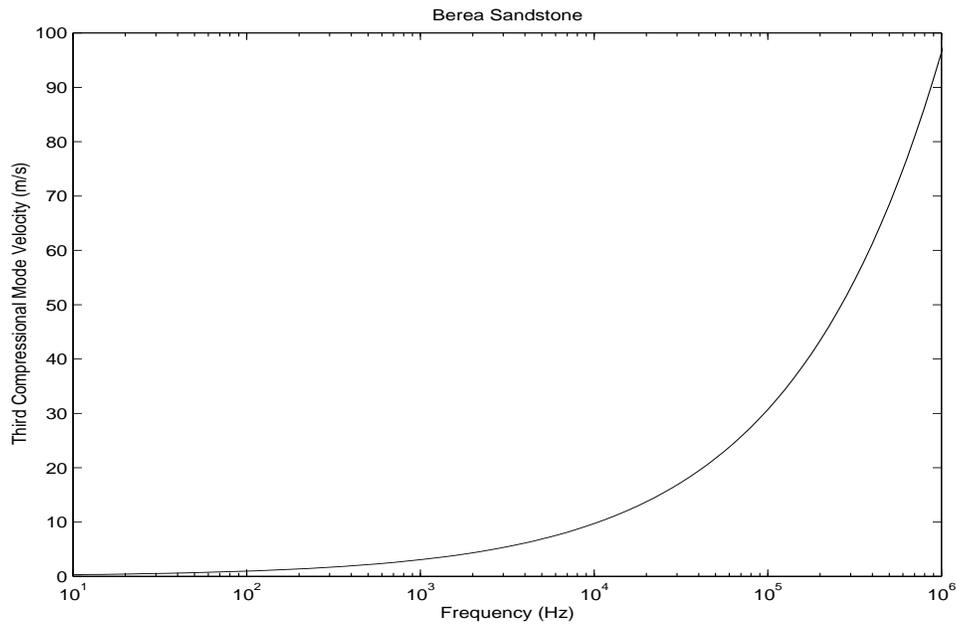


Figure 5: Third compressional wave velocity as a function of frequency for fractured Berea sandstone. jim1-vf [NR]

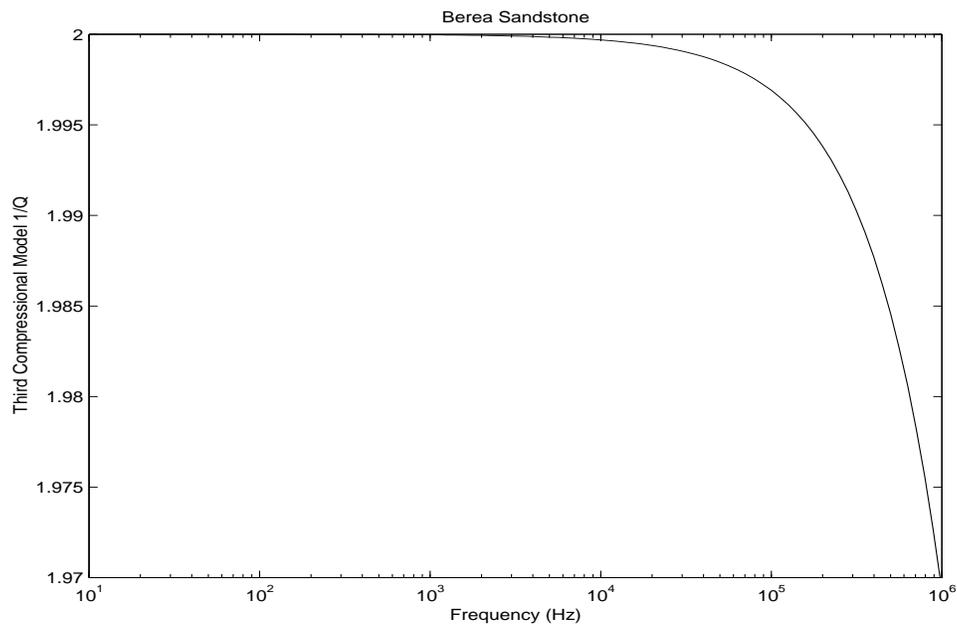


Figure 6: Third compressional wave inverse quality factor $1/Q$ as a function of frequency for fractured Berea sandstone. jim1-ql [NR]

Short Note

A short tour through the Stanford Exploration Project contributions to anisotropy

Tariq Alkhalifah¹

Browsing through the countless Stanford exploration Project (SEP) reports, I was amazed to find a large number of quality papers that tackle the sensitive issue of anisotropy. The amazement stems from the fact that a hidden principle, which states “anisotropy is not an important”, existed in SEP. This principle evolved from statements like “If you believe in anisotropy!” and “avoid the two A’s (anisotropy and AVO)” uttered by SEP’s leader and chief executive, Jon Claerbout, through the years. Despite such statements Jon has maintained an environment in SEP that allows its students to work on any subject they desire. The fact that some students choose anisotropy is a testament to the importance of this subject, and a testament to Francies Muir (past SEP professor and a devote anisotropist) positive influence on the group.

SEP exploits into this subject ranged for modeling seismic waves in anisotropic media to imaging data from such media. Jacobs (1979) was the first to write an anisotropic paper, titled “Velocity anisotropy”, in an SEP report. This paper, among other things, suggested stretching the z-axis to accommodate the ambiguity of resolving depth from surface seismic data in anisotropic media. This was one of the earliest papers to acknowledge the depth problem faced in practice due to anisotropy– the difference between the vertical and imaging velocities. Jacobs (1982) later included anisotropy in his thesis in a chapter titled “Pseudo-*P* and Pseudo-*S* Waves in a Hexagonally Anisotropic Earth”, where he discussed the downward continuation of the Pseudo-*S* wave despite that shear waves was another *taboo* at SEP.

Consistent with the low popularity of anisotropy in the early 80’s, the next paper to treat anisotropy at SEP appeared years later (Dellinger and Muir, 1985) and was titled “Axisymmetric Anisotropy I : Kinematics”. This paper seemingly had opened a flood gate on this subject because the very next report included four papers (Muir and Dellinger, 1985; Dellinger, 1985a; Woodward and Muir, 1985; Dellinger, 1985b) tackling issues ranging from modeling to moveout. The following year, Muir (1987) looked at Dix and Backus averaging for obtaining equivalent parameters and Etgen (1987) applied finite difference to the anisotropic elastic wave equation. One report later, Muir and Dellinger (1988) examined the relation between layering and anisotropy; they sought to obtain an equivalent homogeneous anisotropic media that approximated the layered model. In the same report, Nichols (1988) generated synthetic seismograms for anisotropic models and Guiziou (1988) did some 3-D raytracing in anisotropic media.

¹email: tariq@sep.stanford.edu

Dellinger (1988) built a filter capable of separating the P and S -wave components of the elastic anisotropic wavefield. Inversion of anisotropic parameters was discussed first in SEP by Dellinger (1989), where he examined the feasibility of inverting for the elastic coefficients in a Vertical-seismic-profile (VSP) type of setting. Nichols (1989) showed the equivalence between the Hudson model of the elastic properties of fractured rocks and the Schoenberg model. Karrenbach (1989) uses a paraxial approximation to separate the effect of anisotropy versus $v(z)$ variation of velocity. Popovici and Muir (1989) use an algorithm based on a Markov random field to model anisotropic porous rocks. Muir and Nichols (1989) use the compliance rather than the stiffness tensor to show that rocks can be fractured in any order.

A year later, Dellinger and van Trier (1990) devised one of the earliest implementations of finite difference scheme to the eikonal equation in anisotropic media. Though they had some stability problems at the time, that was treated later (Dellinger, 1991), they managed to produce interesting traveltimes curves in inhomogeneous media. Ji (1990) compares the nonhyperbolic moveout equation of Muir with that of Byun in inverting for an anellipticity parameter. Dellinger et al. (1990) show the success of the Schoenberg-Muir technique of equivalent media on synthetic data. More study on equivalent media followed (Nichols and Karrenbach, 1990; Karrenbach, 1990; Muir, 1993). Cunha (1990) inverted for an elliptical anisotropic parameters using data from a cross well configuration.

In 1991, the double elliptic approximation of Muir (1990) was used in moveout and imaging applications (Dellinger and Muir, 1991; Karrenbach, 1991a); an approximation built to utilize the simplicity of the elliptical anisotropic model to represent a transversely isotropic model. Also, Michelena and Muir (1991) looked at anisotropic tomography. Many other papers that range from azimuthal anisotropy to fine layering to imaging exists in the newer SEP reports (Karrenbach et al., 1992; Berryman, 1996; Alkhalifah and Rampton, 1997b; Alkhalifah and Fomel, 1997; Karrenbach, 1991b; Michelena, 1992; Alkhalifah, 1997a; Alkhalifah et al., 1997; Karrenbach and Muir, 1994; Alkhalifah, 1997c,b; Fomel and Grechka, 1996; Alkhalifah and Rampton, 1997a; Alkhalifah, 1998).

The above is just a highlight of the many papers at SEP that promote and discuss the important subject of anisotropy. The above is in no way inclusive of all the SEP papers published on this subject. The only way to benefit from and appreciate the level of anisotropy contribution in SEP is to directly read and study the papers involved.

REFERENCES

- Alkhalifah, T., and Fomel, S., 1997, Residual migration in VTI media using anisotropy continuation: SEP-94, 327-337.
- Alkhalifah, T., and Rampton, D., 1997a, Seismic anisotropy in trinidad: More parameter estimation: SEP-95, 205-218.
- Alkhalifah, T., and Rampton, D., 1997b, Seismic anisotropy in Trinidad: Processing and interpretation: SEP-94, 299-325.
- Alkhalifah, T., Fomel, S., and Biondi, B., 1997, Time-domain anisotropic processing in arbitrarily inhomogeneous media: SEP-95, 77-99.
- Alkhalifah, T., 1997a, An acoustic wave equation for anisotropic media: SEP-95, 283-307.
- Alkhalifah, T., 1997b, An anisotropic Marmousi model: SEP-95, 265-282.
- Alkhalifah, T., 1997c, Prestack time migration for anisotropic media: SEP-94, 263-298.
- Alkhalifah, T., 1998, The fast marching method in spherical coordinates: SEG/EAGE salt-dome model: SEP-97, 251-264.
- Berryman, J. G., 1996, Range of the P-wave anisotropy parameter for finely layered VTI media: SEP-93, 177-190.
- Cunha, C., 1990, Traveltime inversion of a cross well dataset for elliptically anisotropic media: SEP-67, 31-40.
- Dellinger, J., and Muir, F., 1985, Axisymmetric anisotropy I : Kinematics: SEP-42, 1-24.
- Dellinger, J., and Muir, F., 1991, The double elliptic approximation in the group and phase domains: SEP-70, 361-366.
- Dellinger, J., and van Trier, J., 1990, Anisotropic finite difference traveltimes: SEP-65, 59-66.
- Dellinger, J., Muir, F., and Etgen, J., 1990, Putting Schoenberg Muir to the test: SEP-65, 157-166.
- Dellinger, J., 1985a, Some anisotropic modeling examples: SEP-44, 63-66.
- Dellinger, J., 1985b, Two domains of anisotropy: SEP-44, 59-62.
- Dellinger, J., 1988, Wavefield separation in anisotropic media: SEP-59, 55-72.
- Dellinger, J., 1989, Anisotropic travel time inversion with error bars : Theory: SEP-60, 253-260.
- Dellinger, J., 1991, Finite-difference traveltimes: SEP-69, 114-136.
- Etgen, J., 1987, Finite difference elastic anisotropic wave propagation: SEP-56, 23-58.

- Fomel, S., and Grechka, V., 1996, On nonhyperbolic reflection moveout in anisotropic media: SEP-92, 135–158.
- Guiziou, J. L., 1988, Three dimensional ray tracing in anisotropic media: SEP-57, 207–226.
- Jacobs, B., 1979, Velocity anisotropy: SEP-16, 227–231.
- Jacobs, B., 1982, Pseudo- p and pseudo- s waves in a hexagonally anisotropic earth: SEP-34, 89.
- Ji, J., 1990, Anisotropic velocity analysis: SEP-65, 261–270.
- Karrenbach, M., and Muir, F., 1994, The quest for anisotropic moveout: SEP-82, 91–96.
- Karrenbach, M., Nichols, D., and Muir, F., 1992, Modeling reflections from the Austin chalk – A practical application of azimuthal anisotropy: SEP-75, 251–272.
- Karrenbach, M., 1989, Velocity and Q in transverse isotropic media: SEP-60, 271–284.
- Karrenbach, M., 1990, A general theory for equivalent media: SEP-67, 5–10.
- Karrenbach, M., 1991a, Anisotropic scalar imaging using the double elliptic approximation: SEP-70, 123–130.
- Karrenbach, M., 1991b, Prestack reverse time migration in anisotropic media: SEP-70, 113–122.
- Michelena, R., and Muir, F., 1991, Anisotropic tomography: SEP-70, 201–226.
- Michelena, R. J., 1992, Kinematic ray tracing in anisotropic layered media: Practical details: SEP-73, 345–360.
- Muir, F., and Dellinger, J., 1985, A practical anisotropic system: SEP-44, 55–58.
- Muir, F., and Dellinger, J., 1988, A calculus for layered elastic media: SEP-57, 179–190.
- Muir, F., and Nichols, D., 1989, Elastic properties of fractured rocks: SEP-61, 235–242.
- Muir, F., 1987, Layered models, equivalence, and Abelian groups: SEP-56, 59–70.
- Muir, F., 1990, A modified anisotropic system: SEP-67, 41–42.
- Muir, F., 1993, A result in equivalent medium theory: SEP-77, 319–324.
- Nichols, D., and Karrenbach, M., 1990, What is equivalent in an equivalent medium?: SEP-67, 1–4.
- Nichols, D., 1988, Building anisotropic models: SEP-57, 191–206.
- Nichols, D., 1989, A comparison of two models for the elastic properties of fractured rock: SEP-60, 261–270.
- Popovici, M., and Muir, F., 1989, Modeling anisotropic porous rocks: SEP-60, 295–302.
- Woodward, M., and Muir, F., 1985, Anisotropic NMO removal: SEP-44, 67–74.

25 for 100: Remembrances of the SEP 25th Anniversary Reunion

Stewart A. Levin, Diane P. Levin, and Joe Dellinger¹

¹**email:** not available

SEP ARTICLES PUBLISHED OR IN PRESS

- Alkhalifah, T., Biondi, B., and Fomel, S., 1998, Time-domain processing in arbitrarily inhomogeneous media: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1756–1759.
- Alkhalifah, T., 1998, An acoustic wave equation for anisotropic media: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1913–1916.
- Bachrach, R., and Rickett, J., 1998, Seismic detection of viscous contaminant using shallow seismic reflection: Seismic detection of viscous contaminant using shallow seismic reflection: Env. Eng. Geophys. Soc., Symposium on the Application of Geophysics to Environmental and Engineering Problems.
- Bachrach, R., and Rickett, J., 1999, Ultra shallow seismic reflection in depth: examples from 3D ultra shallow survey with application to joint seismic and GPR imaging: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 488–491.
- Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1723–1726.
- Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: *Geophysics*, **63**, no. 2, 574–588.
- Biondi, B., 1998a, Azimuth moveout vs. dip moveout in inhomogeneous media: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1740–1743.
- Biondi, B., 1998b, Cost-effective prestack depth imaging of marine data: Ann. Internat. Meeting, Offshore Technology Conference.
- Biondi, B., 1998c, Robust reflection tomography in the time domain: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1847–1850.
- Chemingui, N., and Biondi, B., 1999, Data regularization by inversion to common offset (ICO): 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1398–1401.
- Claerbout, J., 1998a, Multidimensional recursive filters via a helix: *Geophysics*, **63**, 1532–1541.
- Claerbout, J., 1998b, Multidimensional recursive filters via a helix with application to velocity estimation and 3-D migration: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1995–1998.
- Clapp, R., Biondi, B., Fomel, S., and Claerbout, J., 1998, Regularizing velocity estimation using geologic dip information: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1851–1854.
- Crawley, S., Clapp, R., and Claerbout, J., 1999, Interpolation with smoothly nonstationary prediction-error filters: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1154–1157.

- Crawley, S., 1998, Shot interpolation for Radon multiple suppression: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1238–1241.
- Guitton, A., and Symes, W. W., 1999, Robust and stable velocity analysis using the Huber function: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1166–1169.
- Prucha, M., Biondi, B., and Symes, W., 1999, Angle-domain common-image gathers by wave-equation migration: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 824–827.
- Qian, J., and Symes, W. W., 1999, Upwind finite difference traveltimes for anisotropic media: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1410–1413.
- Qian, J., Belfi, C. D., and Symes, W. W., 1999, Adaptive finite difference method for traveltimes and amplitude: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1763–1766.
- Rickett, J., and Claerbout, J., 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1675–1678.
- Rickett, J., and Lumley, D., 1998, A cross-equalization processing flow for off-the-shelf 4-D seismic data: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 16–19.
- Rickett, J., Claerbout, J., and Fomel, S., 1998, Implicit 3-D depth migration by wavefield extrapolation with helical boundary conditions: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1124–1127.
- Sava, P., and Fomel, S., 1998, Huygens wavefront tracing: A robust alternative to ray tracing: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1961–1964.
- Sun, Y., and Alkhalifah, T., 1998, Building the 3-D integral DMO operator in the slant-stack domain: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1752–1755.
- Sun, Y., and Fomel, S., 1998, Fast-marching eikonal solver in the tetragonal coordinates: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1949–1952.
- Sun, Y., 1999, Anti-aliasing multiple prediction beyond 2-D: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1338–1341.
- Symes, W. W., and Gockenbach, M. S., 1999, Coherent noise suppression in velocity inversion: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1719–1722.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Berryman, James	723-3187	berryman
Biondi, Biondo	723-1319	biondo
Brown, Morgan	723-6006	morgan
Chemingui, Nizar	723-1250	nizar
Claerbout, Jon	723-3717	jon
Clapp, Robert	723-0253	bob
Crawley, Sean	723-0253	sean
Fomel, Sergey	723-0463	sergey
Guitton, Antoine	723-6007	antoine
Lau, Diane	723-1703	diane
Mora, Carmen	723-5911	cmora
Prucha, Marie	723-0253	marie
Qian, Jianliang	723-6006	qian
Rickett, James	723-6007	james
Sava, Paul	723-0463	paul
Sun, Yalei	723-5911	yalei
Symes, William	725-1334	symes
Vaillant, Louis	723-1250	louis
Zhao, Yi	723-0253	yi

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

The current login name for the sponsors' area on our world-wide web server (<http://sepwww.stanford.edu>) is sepsponsor2. We have set the NEW password to be ss2&sep2. The old login name/password will be disabled on May 10, 1999.

Research Personnel

Tariq Alkhalifah received his B.S. in geophysics from the University of Petroleum and Minerals in Saudi Arabia and his M.S. and Phd in geophysics from Colorado school of Mines. Tariq's current research interests include ray tracing, velocity inversion, and imaging, especially in anisotropic media. From 1997 to 1998 he was at SEP for a post-doc.



James G. Berryman received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on sound waves in rocks – at AT&T Bell Laboratories (1978-81) and in Earth Sciences at Lawrence Livermore National Laboratory (1981-). He is currently a physicist in the Earth and Environmental Sciences Directorate at LLNL. Continuing research interests include seismic and electrical methods of geophysical imaging and waves in porous media containing fluids. He is a member of APS, AGU, SEG, and EEGS.



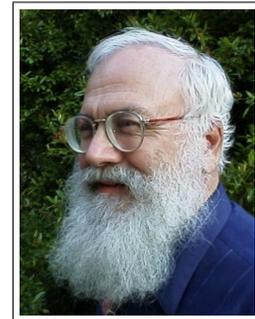
Biondo L. Biondi graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987 he worked as a Research Geophysicist for TOTAL, Compagnie Française des Pétroles in Paris. After his Ph.D. at Stanford 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 consulting company devoted to parallel seismic applications. Biondo is presently full time at SEP as Acting Associate Professor, and leading SEP efforts in 3-D imaging. He is a member of SEG and EAEG.



Morgan Brown received a B.A. in Computational and Applied Mathematics from Rice University in 1997. He is beginning his third year in the Ph.D. program at SEP. Morgan worked as a Research Geophysicist for Western Geophysical (1997) in Houston. He is interested in all aspects of signal processing, and has worked recently on signal/noise separation and pattern recognition. His outside interests include skiing, flyfishing, and backpacking.



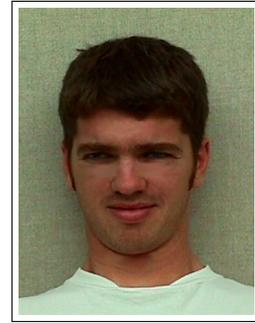
Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Consulted with Chevron (1967-73). Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields*. Honorary member and SEG 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. Published three books: *Fundamentals of Geophysical Data Processing*, 1976; *Imaging the Earth’s Interior* 1985, and *Processing versus Inversion* 1992. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG’s highest award.



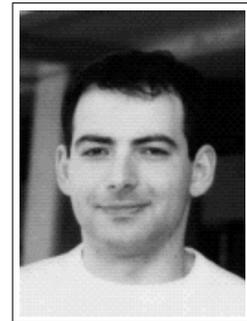
Robert Clapp received his B.Sc.(Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and is working towards a Ph.D. in geophysics. He is a member of the SEG and AGU.



Sean Crawley Graduated from Arizona, joined SEP, and best of all got married in 1994. He received his Masters in 1996. His current research interests are interpolation and multiple suppression.



Sergey Fomel received his Diploma (with Honors) in geophysics from Novosibirsk University in 1990. From 1990 to 1994 he worked at the Institute of Geophysics (Siberian Branch of the Russian Academy of Sciences) in Novosibirsk. He joined the Stanford Exploration Project in September 1994 and is currently working toward a Ph.D. in geophysics. During six months of 1998 Sergey worked at Schlumberger/Geco-Prakla in England. He is a member of SEG and SIAM.



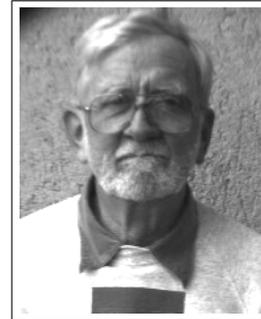
Antoine Guitton received his MSc in geophysics from Université de Strasbourg, France in 1996. He also received (1996) his "Diplome d'ingenieur de L'Ecole de Physique du Globe de Strasbourg". Assistant research geophysicist at the Institut Français du Pétrole (Paris-1996/97) working on well seismic imaging. Assistant research geophysicist at CGG Houston (1997-98) working on multiples elimination. He joined SEP in September 1998. Current research topics are nonlinear inversion and OBC data processing.



Stewart A. Levin was acting director of the Stanford Exploration Project during Jon Claerbout's 1993-4 sabbatical year. Stew is now with Landmark in Denver, developing software tools in areas such as 3D migration, AVO and attribute inversion.



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.



Marie Prucha received her B.Sc.(Hons.) in Geophysical Engineering from Colorado School of Mines in May 1997. She joined SEP in September 1997 and received her MS in June 1999. She is currently working towards a Ph.D. in geophysics. She is a member of SEG.



Jianliang Qian graduated from Harbin Institute of Technology, China. He received his B.S. degree in 1990 and M.S. degree in 1993 with major in applied mathematics, both from Harbin Institute of Technology. He joined TRIP (The Rice Inversion Project, Rice University) in 1996 and is now working towards his Ph.D. degree under the supervision of Prof. Bill Symes'. Currently he is a visiting student at SEP. His research interest includes numerical methods for wave propagation in isotropic and anisotropic media, such as eikonal solvers for traveltimes, prestack Kirchhoff depth migrations and inversions, velocity analysis by inversion. Member: SIAM, SEG.



James Rickett graduated with a B.A. in Natural Science from Cambridge University in 1994, and obtained an M.Sc. in Exploration Geophysics from Leeds University in 1995. He is currently working towards a Ph.D. in geophysics at Stanford, and is a member of the SEG and the EAGE.



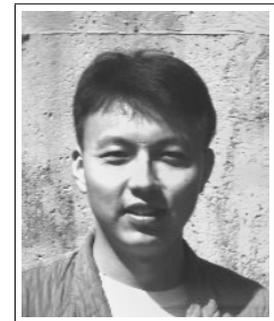
Paul Sava graduated in June 1995 from the University of Bucharest, with an Engineering Degree in Geophysics. From March 1995 he worked as a Software Support Geoscientist and Log Analyst with Schlumberger GeoQuest. He joined SEP in 1997, received his M.Sc. in 1998, and continues his work towards a Ph.D. in Geophysics. He is a member of SEG, EAGE and the Romanian Society of Geophysics.



Mamta Sinha received her B.Sc.(Hons) in 1996 in Geological Sciences and M.Sc. in 1998 in Exploration Geophysics from Indian Institute of Technology, Kharagpur. She joined SEP in September, 1998 and is working towards a Ph.D. in Geophysics.



Yalei Sun graduated from Tsinghua University (Beijing, China). He received his B.S. degree in 1993 and M.S. degree in 1995 with major in solid mechanics. He joined SEP in 1995 and is now working towards his Ph.D. degree. His research interest includes 3-D multiple attenuation, 3-D true-amplitude inversion, fast migration in the wavelet domain, and travelttime calculation in the complex structure.



William W. Symes received degrees in Mathematics from the University of California at Berkeley and Harvard University, and has held faculty and research positions at the University of British Columbia, the University of Wisconsin, and Michigan State University. Since 1983 he has been Professor of Computational and Applied Mathematics at Rice University, where he also directs The Rice Inversion Project, an industry-university consortium project on mathematical and computational problems in seismic imaging. During '98-'99 he is visiting SEP on sabbatical leave from Rice, working with SEPerS on velocity analysis, imaging in the presence of strong refraction, and nonlinear estimation.



Louis Vaillant graduated from the Ecole des Mines de Paris (M.S.) and joined SEP in December 1998 for his military obligation, supported by Elf Exploration Production. He spent one summer with Elf UK at the Geoscience Research Centre in London (1997), and worked on earthquake rupture processes at the Ecole Normale Supérieure de Paris (1998). His current research is on wave-equation migration.



Yi Zhao received his B.S. degree in 1995 and M.S. degree in 1998 in Geophysics from University of Science and Technology of China (Hefei, China).



Sponsors of the Stanford Exploration Project, 1998-99