

**P AND SV POLARIZATION FILTERING
OF A MULTICOMPONENT VECTOR VSP**

A Thesis

Presented to

the Faculty of the Department of Geosciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By Katarina Jovanovic

December 2004

**P AND SV POLARIZATION FILTERING
OF A MULTICOMPONENT VECTOR VSP**

Katarina Jovanovic

APPROVED:

Dr. Kurt Marfurt, Chairman
Department of Geosciences

Dr. Robert Wiley
Department of Geosciences

Dr. Hua-wei Zhou
Department of Geosciences

Dr. Steve Roche
Veritas DGC

Dean, College of Natural
Sciences and Mathematics

ACKNOWLEDGMENTS

This thesis would not have been possible without the constructive ideas, guidance, support and enthusiasm of my advisor, Dr. Kurt Marfurt. I also thank Dr. Robert Wiley, Dr. Hua-wei Zhou and Dr. Steve Roche for their suggestions and review of this thesis. I sincerely appreciate the contribution of my colleagues from Allied Geophysical Laboratories and the Department of Geosciences at the University of Houston.

I want to express my gratitude to Dr. Robert and Margaret Sheriff, David Worthington and the SEG Foundation for their financial support and confidence in my work. I am also indebted to Output Exploration, LLC (OPEX) for providing the data, and to Schlumberger Geoquest, GX Technology, Geocenter and Paradigm Geophysical for providing the software I used in this project.

My deepest gratitude goes to the ones who always believed in me, my parents Mirjana and Daroslav and my brother Ivan, for their constant encouragement, enormous love and total support. My special thanks go to Patrick Jonke for his love, optimism and help with editing this thesis. I also would like to thank my family and friends for all their support, encouragement and assistance during this work.

**P AND SV POLARIZATION FILTERING
OF A MULTICOMPONENT VECTOR VSP**

An Abstract of a Thesis

Presented to

the Faculty of the Department of Geosciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By Katarina Jovanovic

December 2004

ABSTRACT

The separation of P and S wavefields is required to migrate PP and PS reflections acquired by a multicomponent 3D VSP. However, 3D VSP data suffer from aliasing, such that both f - k and conventional Radon transform separation techniques usually give poor results. Therefore, I performed P and SV separation using the high resolution Radon transform and the polarization characteristics of a particular wave.

Using the fact that stronger seismic reflections are typically coherent, I applied semblance derived model weights that construct stronger, most coherent events first. Additionally, in the case of the equally coherent P and SV events with the same slope and intercept, I favor the mode with a higher energy. I implemented these time variant constraints via a time domain weighted conjugate gradient solution for nonlinear least-squares equations.

Separation results depend on the spatial window size. Since we know impedances within the well, we can reliably estimate the major reflectors and form spatial windows based on this information, avoiding the poor representation of discrete waveforms and weak semblance estimates at major discontinuities.

In addition to the P and SV waves separation, this algorithm can be used not only for filtering upgoing-downgoing wavefields, but for muting of any event, which can be particularly helpful for the extraction of salt face reflections.

Compared to the classical VSP separation techniques, this method successfully separates major linear events, having fewer filtering artifacts. Unfortunately, the linear Radon transform does not accurately fit the hyperbolic moveout of the salt face reflections, which remains as a filtering challenge.

CONTENTS

LIST OF SYMBOLS	ix
1. INTRODUCTION	1
1.1. Motivation	1
1.2. Previous work.....	2
1.3. Objective	4
1.4. Content	5
2. MODELING AND PROCESSING OF THE VINTON DOME VSP DATA	
USING THE <i>F-K</i> AND MEDIAN FILTERS	7
2.1. Geology of the Vinton Dome	7
2.2. Acquisition of the Vinton Dome survey	11
2.3. Processing of the Vinton Dome VSP data	12
2.4. Modeling of the Vinton Dome and generation of the synthetic VSP data.....	14
2.5. Separation of synthetic VSP data using the classical processing approach	25
2.5.1. P and S wave separation using the <i>f-k</i> filter	26
2.5.2. P and S wave separation using the median filter.....	29
2.5.3. P and S wave separation using combined <i>f-k</i> and median filter.....	31
2.6. Separation of Vinton Dome VSP data using the classical processing approach....	40
2.7. Limitations of the classical mode separation approaches	42
3. THEORY AND DEVELOPMENT OF THE HIGH RESOLUTION DISCRETE	
RADON TRANSFORM VECTOR VSP SEPARATION ALGORITHM	43
3.1. Propagation and polarization of VSP plane P and S waves	43
3.2. Decomposition of vector VSP wavefield into the scalar compressional and shear	

waves	48
3.3. Radon transformation	50
3.4. Aliasing	52
3.5. Least-square linear discrete Radon transform	54
3.5.1. Preconditioning of the data	55
3.6. P and SV wave separation of vector VSP data by conjugate gradient method	57
4. APPLICATIONS OF THE VECTOR VSP SEPARATION ALGORITHM	64
4.1. Algorithm structure and requirements	64
4.1.1. Three approaches for algorithm application based on the size of a spatial window	65
4.1.2. Separation of the upgoing and downgoing waves	67
4.2. Algorithm testing on synthetic data	68
4.2.1. A simple model test	68
4.2.1.1. A simple test model with added random noise	81
4.2.2. Vinton Dome test model	88
4.2.2.1. A fixed, large spatial window	88
4.2.2.2. A small, running spatial window	89
4.2.2.3. Fixed, sequential windows	90
4.2. Algorithm testing on the Vinton Dome VSP data	97
4.2.1. Tests of constraints	97
4.2.2. Spatial window test	98
4.2.3. Offset test	99
4.2.4. Data separation by filtering in τ - p space	99

5. CONCLUSIONS	115
REFERENCES	117
APPENDIX A: CONJUGATE GRADIENT METHOD	121
A.1. The method of steepest descent.....	122
A.2. The Method of Conjugate Gradients	124

LIST OF SYMBOLS¹

φ - the angle between the north direction and shot-receiver direction.

λ, μ - Lamé's coefficients.

ρ - density (g/cm³).

t - time (s).

u - a wave displacement vector.

s - a ray propagation vector, i.e. slowness (s/m).

v_p - the P-wave velocity (m/s).

v_s - the S-wave velocity (m/s).

x_n - the n -th surface receiver location.

z_n - the n -th depth (VSP) receiver location.

Δx - distance between surface geophones (m).

Δz - distance between depth (VSP) geophones (m).

Δt - a time difference in plane waves arrivals (s).

ψ_p - the angle of P-wave propagation, i.e. the angle between a ray propagation vector and VSP receiver array (rad).

ψ_s - the angle of S-wave propagation, i.e. the angle between a ray propagation vector and VSP receiver array (rad).

$\hat{\mathbf{x}}$ - the unit vector in the horizontal direction.

$\hat{\mathbf{z}}$ - the unit vector in the vertical direction.

s_{xp} - magnitude of the P-wave slowness vector in the x direction (s/m).

¹ Lower case letters - scalars; lower case, bold letters- vectors; lower case, bold letters with ^ symbol - unit vectors; upper case, bold letters- matrices.

s_{xs} - magnitude of the S-wave slowness vector in the x direction (s/m).

s_{zp} - magnitude of the P wave slowness vector in the z direction (s/m).

s_{zp} - magnitude of the S wave slowness vector in the z direction (s/m).

ω - temporal frequency (rad/s).

\mathbf{a}_p - P-wave amplitude vector.

\mathbf{a}_s - S-wave amplitude vector.

a_{px} - P-wave amplitude x-component.

a_{pz} - P-wave amplitude z-component.

a_{sx} - S-wave amplitude x-component.

a_{sz} - S-wave amplitude z-component.

\mathbf{u}_p - P-wave displacement vector.

\mathbf{u}_s - S-wave displacement vector.

u_{px} - P-wave particle displacement in the x direction.

u_{pz} - P-wave particle displacement the z direction.

u_{sx} - S-wave particle displacement in the x direction.

u_{sz} - S-wave particle displacement in the z direction.

p - ray parameter (s/m).

$\hat{\mathbf{a}}_p$ - P-wave polarization unit vector.

$\hat{\mathbf{a}}_s$ - S-wave polarization unit vector.

τ - intercept time (s).

f - frequency (Hz).

k - wavenumber (rad/m).

x - source-receiver offset (m).
 f_{Nyq} - Nyquist frequency (Hz).
 k_{Nyq} - Nyquist wavenumber (rad/m).
 Δp - the ray parameter interval (s/m).
 x_r - range of offset values (m).
 f_{max} - maximum frequency in the data (Hz).
 $d(t,x)$ - the observed VSP data.
 $m(\tau,p)$ - the τ - p transform for intercept τ and ray parameter p .
 \mathbf{L} - a transformation-polarization matrix.
 ε - a small number introduced to assure numerical stability.
 \mathbf{I} - the identity matrix.
 σ - the semblance function.
 \mathbf{M} - model preconditioning matrix.
 $\rho(\tau)$ - the time domain 'rho filter'.
 $\tilde{\rho}(\omega)$ - the Fourier transform of $\rho(\tau)$.
 \mathbf{W} - data weighting matrix.
 \mathbf{d}_{xobs} - the observed horizontal data space vector VSP.
 \mathbf{d}_{zobs} - the observed vertical data space vector VSP.
 \mathbf{m}_p - P-wave model space vector.
 \mathbf{m}_s - S-wave model space vector.
 \mathbf{T} - Radon transform matrix.
 \mathbf{T}_p - P-wave Radon transform vector.

\mathbf{T}_s - S-wave Radon transform vector.

\mathbf{A} - polarization matrix made by unit polarization vectors $\hat{\mathbf{a}}_p$ and $\hat{\mathbf{a}}_s$.

\mathbf{d}_{xn} - reconstructed x component data after iteration n .

\mathbf{d}_{zn} - reconstructed z component data after iteration n .

$\Delta\mathbf{d}_{xn}$ - x component data reconstruction vector at iteration n .

$\Delta\mathbf{d}_{zn}$ - z component data reconstruction vector at iteration n .

\mathbf{r}_{xn} - x component data misfit (residual) vector at iteration n .

\mathbf{r}_{zn} - z component data misfit (residual) vector at iteration n .

\mathbf{R} - the 'rho filter' matrix.

\mathbf{g}_{pn} - P-wave gradient vector at iteration n .

\mathbf{g}_{sn} - S-wave gradient vector at iteration n .

\mathbf{h}_{pn} - P-wave conjugate gradient vector at iteration n .

\mathbf{h}_{sn} - S-wave conjugate gradient vector at iteration n .

$\boldsymbol{\sigma}_p$ - P-wave semblance vector.

$\boldsymbol{\sigma}_s$ - S-wave semblance vector.

$\boldsymbol{\zeta}_p$ - P-wave semblance based taper vector used in model preconditioning matrix \mathbf{M} .

$\boldsymbol{\zeta}_s$ - S-wave semblance based taper vector used in model preconditioning matrix \mathbf{M} .

$\boldsymbol{\mu}_p$ - P-wave masking vector used for model preconditioning matrix \mathbf{M} .

$\boldsymbol{\mu}_s$ - S-wave masking vector used for model preconditioning matrix \mathbf{M} .

ϕ - the semblance relaxation criteria.

α_n - the model correction steplength at iteration n .

$J(\alpha_n)$ - the objective function in the numerical search of α_n .

β_n - the conjugate direction steplength at iteration n .

ξ_{\min} - the user specified convergence tolerance.

1. INTRODUCTION

1.1. Motivation

Vertical seismic profiling (VSP) is a geophysical technique where seismic energy is generated at the surface of the earth and recorded with the receivers down in the well. VSPs provide a direct relationship between surface seismic data and borehole data, including logs and cores. The VSP technique was originally developed to provide a direct measurement of time-to-depth conversion, but it also gives information useful in data processing including interval velocities, attenuation values, zero-phase reflectivity, multiple prediction and reflectors generating shear conversion. The uniqueness of VSPs is that the seismic energy is measured as it propagates down the earth. This provides a good evidence of different wave types and possibilities for better understanding of wave propagation phenomena. VSPs also help in improving the definition of the subsurface velocity model.

Integrated with surface seismic, VSP can significantly improve both vertical and lateral resolution of the seismic image around the borehole. In addition, VSPs unambiguously tie seismic reflections to depth, and via well logs to lithology and stratigraphic boundaries (Hardage 1983, Stewart and Disiena, 1989, Chopra *et al.*, 2004).

Most VSP surveys have been either 1-D where the shots are close to the well head (zero-offset VSP), or 2-D where the shots are along a line away from the borehole. However, due to the fact that the Earth and its properties are 3-D, we now also use 3-D VSP techniques. In the 3-D VSP technique, the seismic signal is generated from an areal distribution of surface shots and detected using the geophones in the well. 3-D reverse VSPs, where the shot is in the hole and (less expensive) geophones placed on the surface,

promise to make 3-D VSP acquisition affordable if we can generate sufficient signal without damaging the well. Early 3-D VSPs were shoot by AGIP/Schlumberger in 1986 (Stewart et al., 1998).

The most valuable 3D VSP data are obtained using the multicomponent geophones. The data are acquired using three orthogonal geophones providing a complete record of both P and S wavefields (Cornish *et al.*, 2000, Roche *et al.*, 1999). By the decomposition of such a vector wavefield, we can separate P and S waves from full wavefield. For use in imaging and amplitude analysis, the motivation for the P and SV wavefields separation is the improvement of:

- VSP processing,
- event identification and picking for P-traveltime and S-traveltime tomography,
- PP and PS AVO processing with long offsets,
- estimation of both P-wave and S-wave anisotropy, and possibly
- determination if faults/fractures are fluid filled or tight.

1.2. Previous work

Recorded multicomponent VSP wavefields consist of a multitude of interfering compressional and shear waves, which are recorded on each of the vertical, north, and east geophone components. Our difficulty is in separating these modes of propagation while retaining their amplitude and phase characteristics.

If we know the arrival times of the wave along the array as well as compressional and shear velocities of the formation, we can decompose the vector wavefield into P and S waves using the plane-wave expansion in the frequency-wavenumber ($f-k$) domain (Devaney and Oristaglio, 1986). A velocity based separation filter can be applied in the f -

k domain. The data are then inverted to obtain a result in the time-depth domain (Dankbaar, 1987). Similar techniques for separation by polarization dependant velocity filtering in the frequency-wavenumber domain were used by Morozov et al. (1997). For a given plane wave type and propagation angle, slowness and polarization vectors can be computed for each plane wave. The scalar plane wave amplitudes can then be obtained by solving a linear system for each frequency (Scott et al., 1989). Assuming the downgoing data as a superposition of P and SV waves, and using the local P and S velocities and angles of incidence as model parameters, Esmersoy (1990) minimized the least square error between modeled and data waveforms in the frequency domain. Unfortunately, separation of P and S waves using the f-k transform suffers from spatial aliasing inherent in VSP data acquisition.

Several authors have applied the Radon transform to the VSP wavefield separation problem. Foster and Gaiser (1986) rotated the coordinates for P and S waves in the forward Radon Transformation step, after which the P and S wavefields were obtained using the inverse Radon transform. Boelle et al. (1998a) decomposed the data in τ - p space using a least-square inversion followed by polarization analysis over a running window and obtained polarization and azimuth angles used in the calculation of P and S waves velocities. Using these velocities, they generate mutes to separate P and S wavefields in the τ - p domain. Leaney (2002) developed a very efficient least-squares vector wavefield separation technique for 3-D VSP data. For a given plane wave type and propagation angle, he computed slowness and polarization vectors for each plane wave. He solved the linear system at each frequency to yield the scalar plane wave amplitudes. Again, muting of these scalar plane waves in the τ - p domain allows reconstruction of

separated P and S wavefields. The methods above are all sensitive to aliasing when implemented in the frequency domain. An unconstrained frequency domain Radon transform results in very efficient system of equations having a Toeplitz matrix structure. Adding our *a priori* constraints couples each frequency component to the others, thereby destroying this efficient form. In this work, I will present applications of a high resolution discrete Radon transform in the time domain to the vector wavefield separation problem for the Vinton Dome data set.

1.3. Objective

One of the research objectives of the Allied Geophysical Laboratories at the University of Houston is to develop new processing and interpretation techniques for both surface and vector VSP data acquired at Vinton Dome, Louisiana. The joint processing effort is to improve imaging of the salt dome. My research work and the objective of this thesis is P and SV wave separation of vector VSP data.

A classical VSP processing approach is VSP data decomposition to downgoing and upgoing wavefields by picking particular wave arrivals (Hardage, 1983). The same processing technique can be applied to obtain separate P and SV wavefields. Unfortunately, due to complex geological structures and very long offsets, wave fields of interest may arrive horizontally, such that the classical processing approach is not an efficient wave separation technique. For this reason, it is necessary to develop a new and specific separation algorithm.

The Vinton Dome 3-D VSP source geometry is irregular and acquired data are highly aliased. Separation of such aliased VSP data by an $f-k$ transform will not give correct results. The alternative $\tau-p$ transform suffers from the same aliasing artifacts

(Marfurt et al., 1996) when implemented in the frequency domain. The main goal of this research is to exploit some concepts developed for high resolution Radon transform in the time domain and extend them to the wavefield separation problem, in particular:

- development of a high resolution vector Radon transform P and SV waves separation algorithm,
- calibration of the algorithm using the full waveform synthetic VSPs,
- quantification of the separation by comparison of P and SV synthetic VSP wavefields obtained by classical separation processing and by the implementation of my developed algorithm, and
- application of the high resolution vector Radon transform separation algorithm to the Vinton Dome multicomponent VSP survey.

1.4. Content

Following this introduction, in Chapter 2, I briefly describe the geology of the Vinton Dome area and explain in more details modeling and generation of a full wavefield synthetic VSP data. I separate different VSP modes using the classical processing techniques. Obtained results show capabilities of applied techniques. More important, I use separated VSP modes to compare and evaluate my new vector Radon transform separation approach.

In Chapter 3, I review the general theory for the discrete Radon transform and aliasing, as well as wavefield propagation and plane wave decomposition concepts for the vector VSP data. I also explain the seismic inversion techniques applied in this research.

Chapter 4 is a detailed explanation of development and calibration of the vector VSP separation algorithm. Here I apply the P and SV waves separation algorithm first to

a couple of simple models, then to the synthetic VSP data set described in Chapter 2 and, finally, to the Vinton Dome VSP seismic data. I discuss quality of the separated P and SV data sets and compare results obtained by the new separation algorithm with the results obtained by classical VSP separation techniques.

I conclude with a summary of the advantages and limitations of the method.

2. MODELING AND PROCESSING OF THE VINTON DOME VSP DATA USING THE *F-K* AND MEDIAN FILTERS

In order to evaluate the high resolution discrete Radon transform VSP separation code, I will compare results from the Radon transform separation algorithm with those from the classical VSP separation processing. Separation is performed on both synthetic and real Vinton Dome VSP data. In order to generate a useful synthetic data set, I will make an appropriate model based on the geology of the research area.

2.1. Geology of the Vinton Dome

Vinton Dome is situated in south-west Louisiana in the Calcasieu Parish. It is a mature oil field with the main production from Tertiary sands. The field has produced 140 million barrels of oil since 1900 (Constance, 2003).

Vinton Dome complex geology is the result of the Cenozoic structural evolution of the northern Gulf of Mexico Basin where different tectono-stratigraphic provinces (Figure 2.1) describe regions of contrasting structural styles and ages (Diegel et al., 1995). Geology of the Vinton Dome is characterized by stratigraphy, structure and salt tectonics.

The most important Tertiary stratigraphic units of the western Gulf Coast are the Vickburg, Frio and Anahuac formations. The Vickburg stage belongs to the lower Oligocene and consists of calcareous and fossiliferous shales. The Anahuac and Frio Stages have long been known as prolific sources of hydrocarbons and may be divided into three units based on paleontology (Warren, 1957).

The Oligocene Frio formation consists of dark to very dark silty shales and massive to thin-bedded strata of sand. The upper and lower sandy units are separated by a

middle shale unit. Sediments and fossils of the lower Frio represent a regressive to transgressive phase, those of the middle Frio indicate an inundative phase, while the sediments of the upper Frio suggest the regressive to transgressive phase (Warren, 1957). In the Vinton Dome area, the Frio consists of five 30 m thick sand units. These are deltaic sands coarsening upward. Well-defined, 30 m thick shale separates the sands. Underlying the upper Frio is the Hackberry unit of the middle Frio, with the gas producing sand channels on both sides of the Vinton Dome (Constance, 2003).

The Anahuac Formation in the eastern part of Louisiana consists of shales with some sand and limestones and is said to be either upper Oligocene or lower Miocene age (Warren, 1957). The Anahuac shale provides an important seal to migrating hydrocarbons of the Vinton Dome Frio formation. It is deep-water, 214 m thick shale unit (Constance, 2003).

In southwestern Louisiana, the strata immediately above the Anahuac formation is a more or less massive sand unit. The Vinton Dome post-Anahuac shallow Miocene unit consists of unconsolidated sands inter bedded with shales. Most of the production from this unit is found close to the crest of the dome (Constance, 2003).

Though previously seen as a passive margin with vertical rooted salt stocks and massifs with steep growth faults, the northern Gulf of Mexico basin is a complex mosaic of diachronous detachment fault systems and variously deformed allochthonous salt sheets (Diegel et al., 1996). The Oligocene-Miocene detachment province covers the modern slope and coastal onshore of Louisiana and Texas (Figure 2.1). The region is characterized by the down-to-the-basin normal listric growth faults that sole on the regional detachment and by a very thick package of deltaic sediments above the

detachment. The northern edge of the Oligocene-Miocene detachment is the Salt-Dome Minibasin Province. During the progradation, extensive salt withdrawal from tabular salt bodies on the slope formed isolated salt minibasins bounded by the large arcuate counter-regional growth faults. The counter-regional growth faults are also normal faults, but with the dips in the opposite direction to the listric normal faults (Figure 2.2). As the margin moved toward the basin, salt was moved along the listric fault. The Vinton Dome was formed when one portion of that salt was likely detached and pushed up the arcuate fault. The salt movement has pushed the gently dipping Oligocene-Miocene sediments upward, forming the dipping sediments near the flanks and radial faults around the dome.

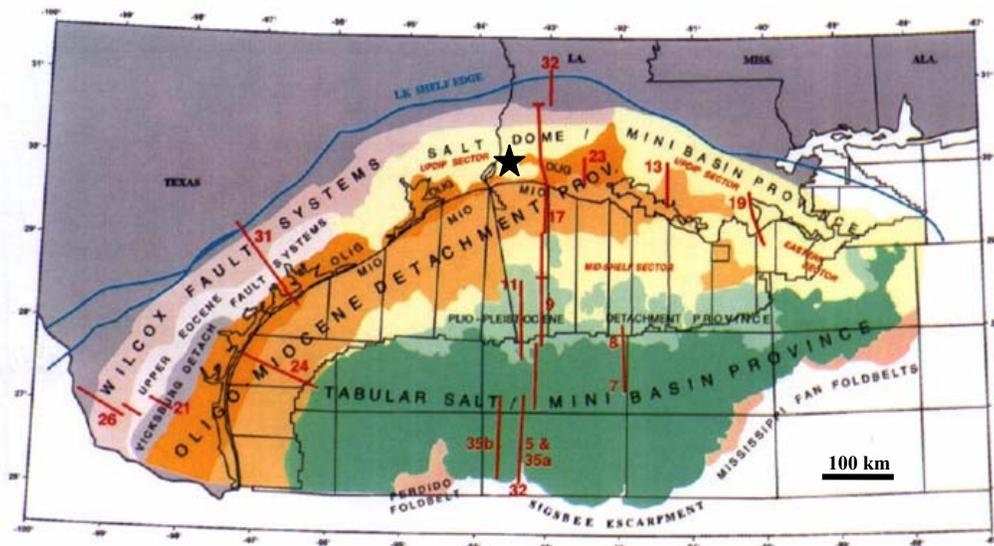


FIG. 2.1. Tectono-stratigraphic provinces of the northern Gulf of Mexico Basin. Location of the Vinton Dome is indicated by the star (after Diegel et al., 1996).

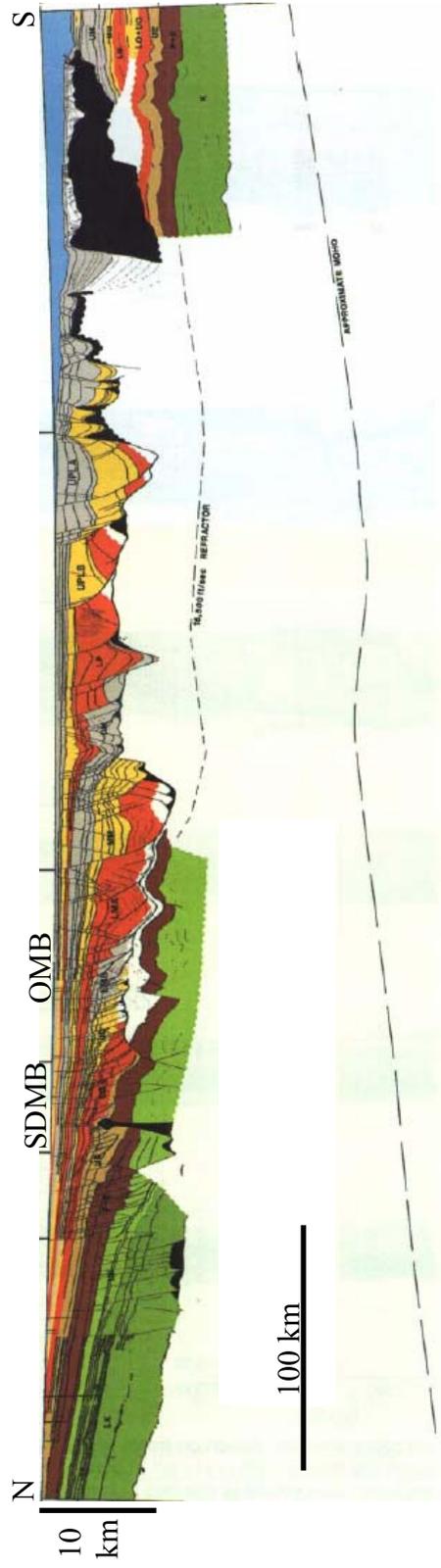


FIG. 2.2. Interpreted seismic profile across western Louisiana, from the lower Cretaceous margin to the abyssal plain. Salt Dome Minibasin province is marked by SDMB. Oligocene-Miocene detachment province is marked by OMB. Salt is shown in black (after Diegel et al., 1996).

2.2. Acquisition of the Vinton Dome survey

A seismic survey over the Vinton Dome was conducted by Output Exploration, LLC (OPEX) in 1997. Surface seismic 3-D data and VSP data were acquired simultaneously. The seismic sources were 2.5 kg pentolite charges at 18 m depth. Seismic energy was recorded by vertical geophones in a conventional surface spread and by two downhole multicomponent arrays deployed in two abandoned field wells (Roche et al., 1999).

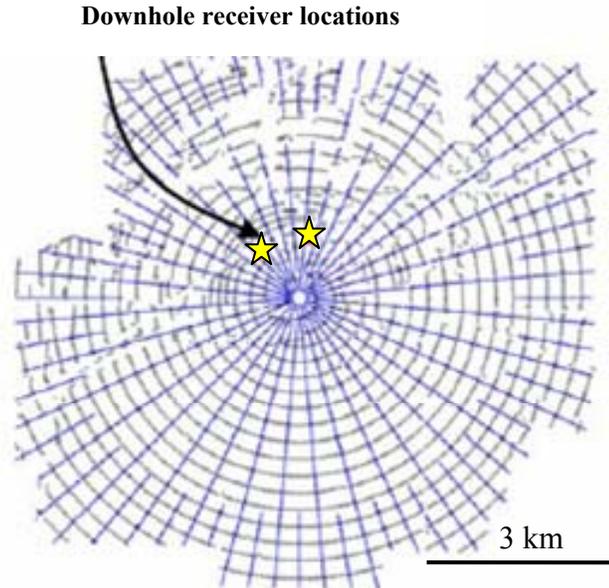


FIG. 2.3. Data acquisition grid showing the radial receiver lines centered over the Vinton Dome structure. Source lines are in concentric circles around the dome. Positions of VSP arrays are indicated by the stars (after Roche et al., 1999).

Surface receiver lines were radially directed from the piercement salt dome. Short and long receiver lines were alternated such that the receiver line separation was approximately 5° or ranging 275 to 365 m (Figure 2.3). The receiver station spacing is 50

m along the radial receiver line with a six 10 Hz geophone array at each station. Source positions were along the approximately concentric circles with individual shot spacing 50 m along the arc. Bin size is 20 m (Roche et al., 1999).

Simultaneously with the acquisition of the northern half of the 3-D surface survey, VSP data were recorded in two abandoned wells using three-component geophone arrays. I will use VSP data recorded in the western G-23 well.

Well G-23 was instrumented with 80 multicomponent geophones. Receivers were cemented over the interval of 287 to 1492 m TVD. Bottom 18 levels were damaged during the deployment and were not used in the survey. The 3-C 3-D VSP was recorded with 61 active three-component geophones at 15.25 m increment. The first geophone is at 287 m and the last one at 1202 m (Roche et al., 1999). The record length is 8 s.

2.3. Processing of the Vinton Dome VSP data

After acquisition, the Vinton Dome VSP data are prepared for further processing. Each horizontal component pair from a geophone level deployed in the well has a different orientation during the acquisition. Since we do not know that orientation, we have to calculate orientation azimuths from the energy of the first arrivals.

Every sensor level is then rotated such that the two components are oriented in the same coordinate system: for example, one component toward north and the other toward east. This part of data processing, together with deconvolution and trace balancing was performed by Mariana Gherasim. Data from shot 3068 are shown in Figure 2.4.

In a 2-D problem definition, we assume that shot and receivers are in the same vertical plane. In order to apply the separation algorithm, I rotated north and east components to the radial and transversal direction:

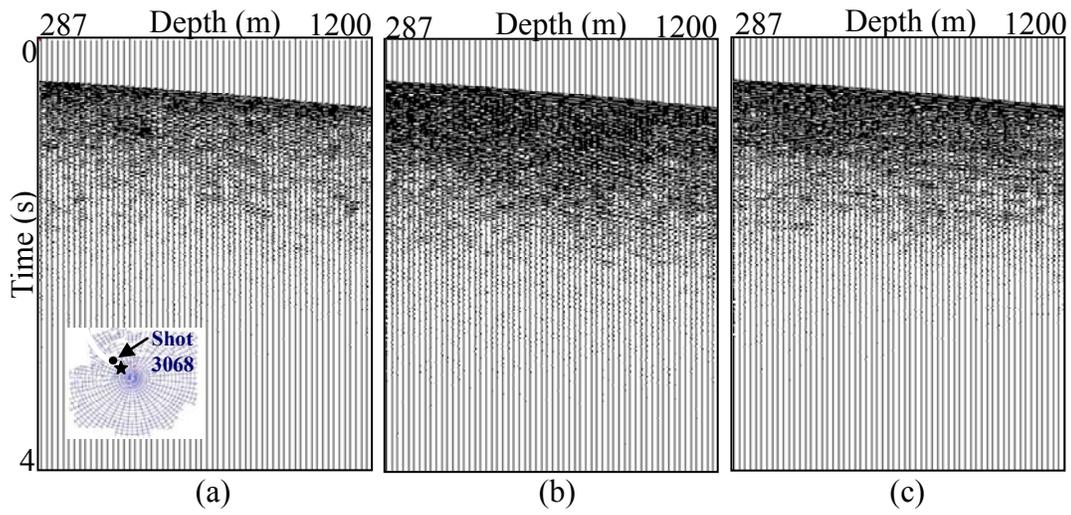


FIG. 2.4. East (a), north (b) and vertical (c) component, shot 3068, offset 600 m. Well G-23 and shot 3068 positions are shown in the small, captured figure.

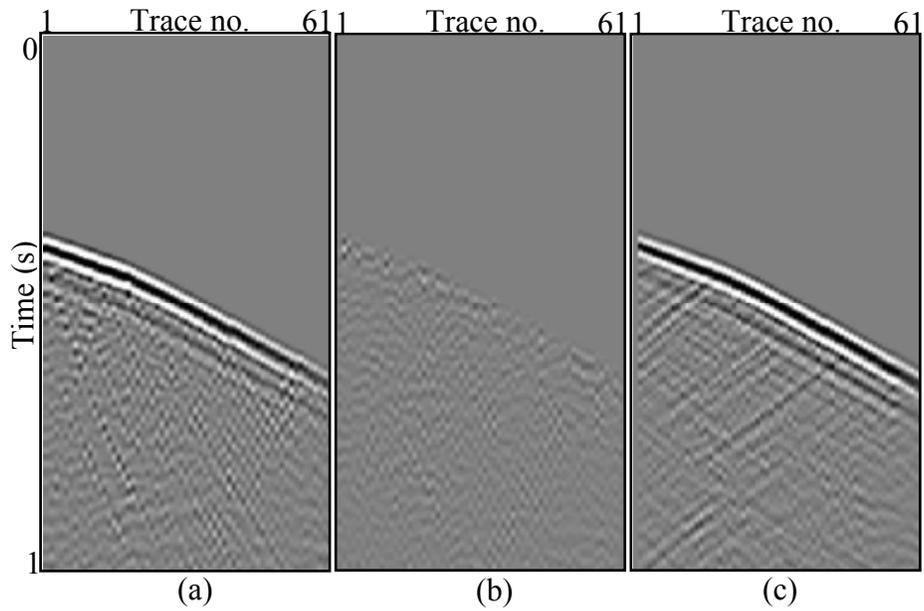


FIG. 2.5. Radial (a), transversal (b) and vertical (c) component, shot 3068, offset 600 m.

$$\mathbf{d}_{\text{radial}} = \sin\varphi \mathbf{d}_{\text{east}} + \cos\varphi \mathbf{d}_{\text{north}} \quad (2-1)$$

$$\mathbf{d}_{\text{transversal}} = \cos\varphi \mathbf{d}_{\text{east}} + \sin\varphi \mathbf{d}_{\text{north}} \quad (2-2)$$

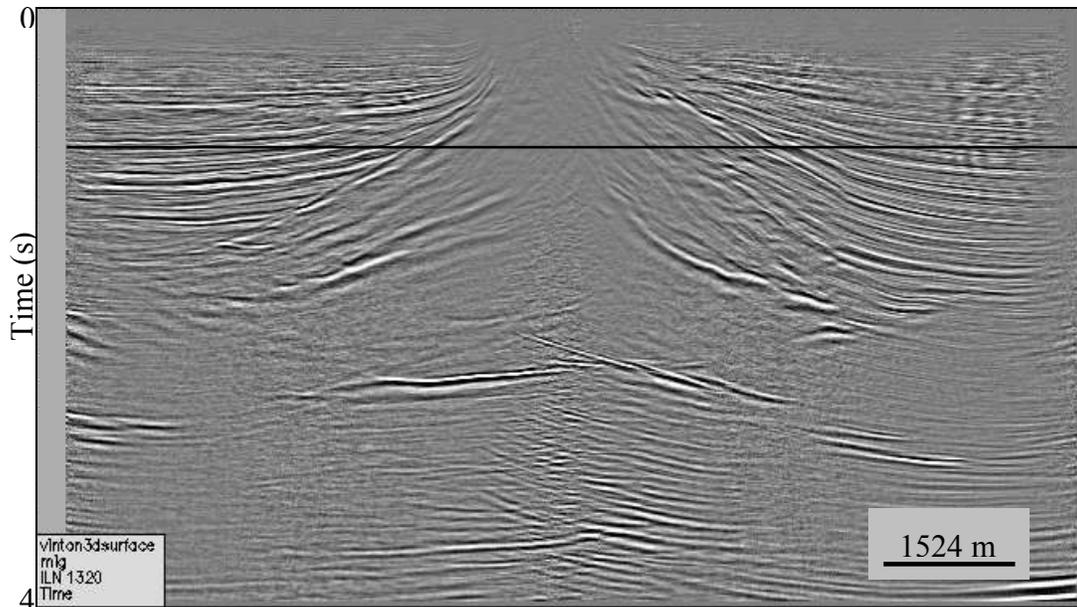
The angle φ is the azimuth angle between the north direction and shot-receiver direction. This angle is calculated from shot and well coordinates. One second of rotated data is shown on Figure 2.5.

As we can see, determination of the particular VSP arrivals is hard. In order to better understand this data set, I have modeled Vinton Dome using the surface seismic.

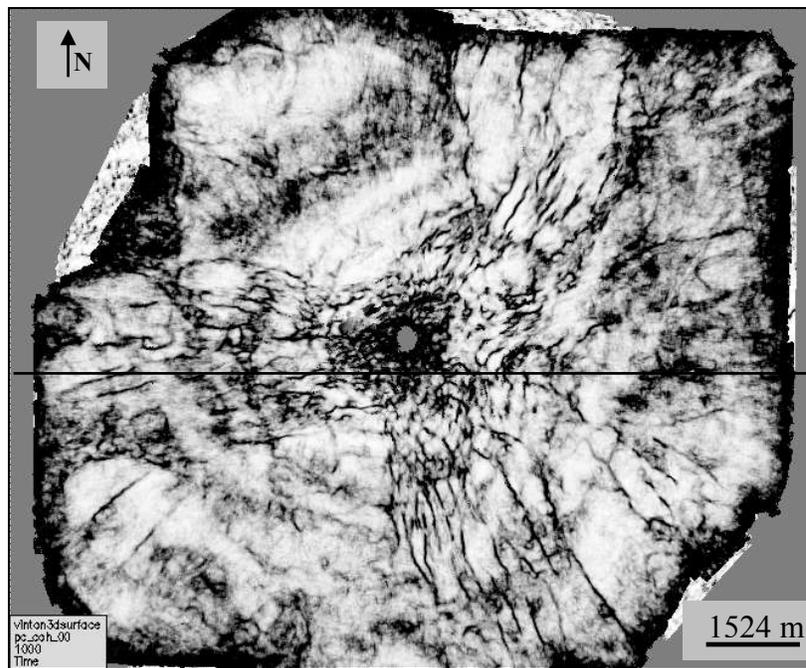
2.4. Modeling of the Vinton Dome and generation of the synthetic VSP data

The Vinton Dome survey recording geometry provided a wide range of offsets and azimuth fold coverage of the subsurface. While designed to avoid raypaths through the salt body, the positions of receivers and sources enabled a good fold in the area of the salt flanks and a better illumination of the steep dips (Figure 2.6a) and faults (Figure 2.6b) near the dome.

I interpreted surface seismic volume in order to understand the geology of the area and make an appropriate Vinton Dome model. Interpretation was performed with Warren Duncan, Marija Djordjevic and Sonja Kisin. We picked seven horizons: the Upmiocene, the Upmidmiocene, the Midmiocene, the Topanahuac, the A-sand and D-sand of the Upper Frio formation and the Hackberry of the Middle Frio unit. Horizons at CDP 340 are shown in Figure 2.7a. Together with the horizons, the major faults have been picked. Most of the horizons (especially upper geological units) were hard to track near the dome flanks (Figure 2.7b) because of steep dips and extensive radial faulting.

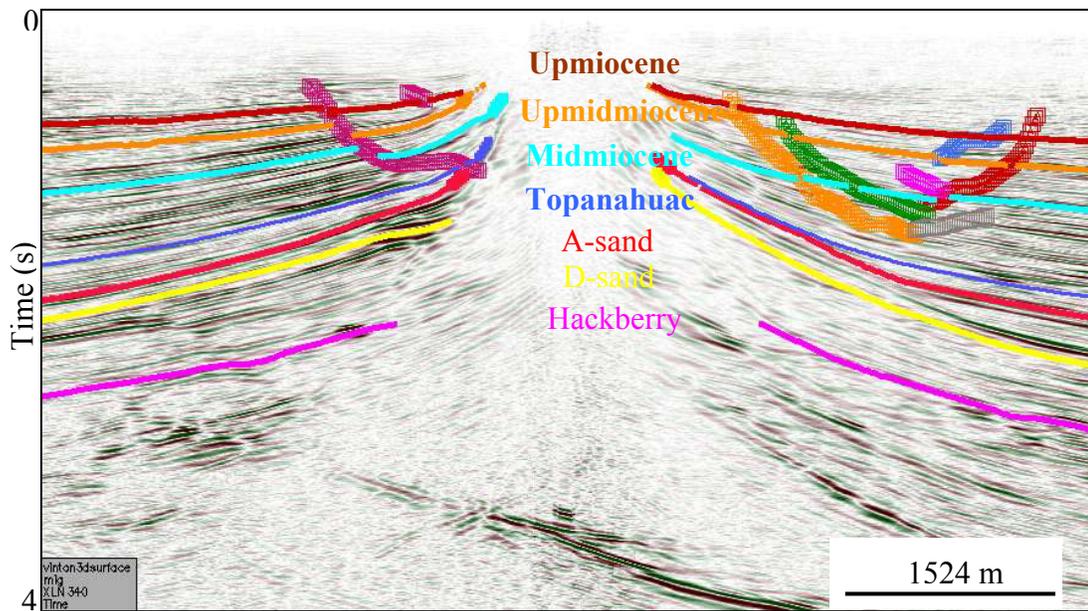


(a)

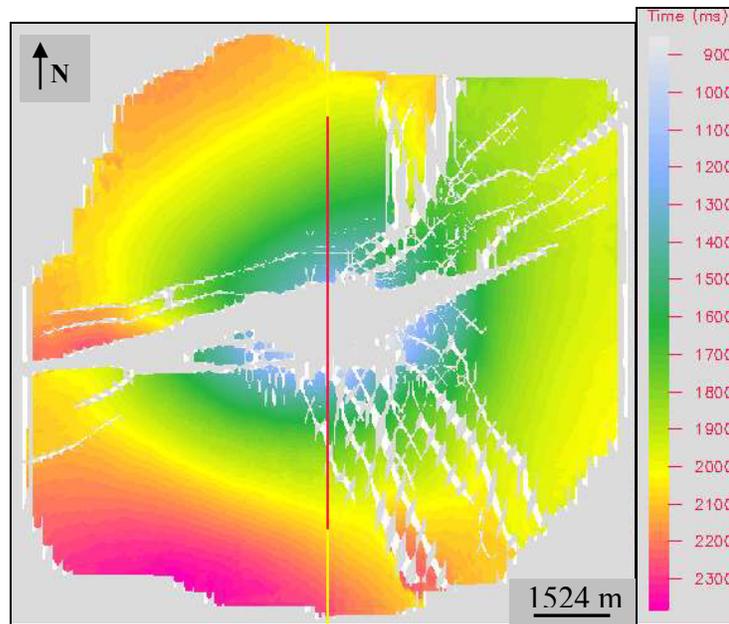


(b)

FIG. 2.6. Migrated Vinton dome surface seismic data. Line 1320 (W-E) crosses the central part of the Vinton Dome. The salt has moved sediments upwards near the dome flanks. Position of the time slice is marked with a black line (a). Principal coherence time slice at 1000 ms. Coherent events are white, while less coherent events are black. Extensive radial faulting occurs in three major zones: W, N-NE and S-SE. Position of Line 1320 is marked with a black line (b). Seismic data courtesy of OPEX.



(a)



(b)

FIG. 2.7. The Vinton Dome surface seismic data interpretation. CDP 340 (N-S) with seven Vinton Dome horizons: Upmiocene, Upmidmiocene, Midmiocene, Topanahuac, A-sand, D-sand, and Hackberry (a). A-sand horizon was hard to pick in the area of extensive faulting and salt. CDP 340 is marked with a red line (b). Seismic data courtesy of OPEX.

The salt body has pushed all horizons upwards, but the salt itself does not reach the surface. The dome has a cone shape, with a narrower part on the top and a larger diameter toward the bottom. The upper part of the dome is better imaged in the depth migration data (Figure 2.8). However, the deeper salt boundaries were hard to pick despite the high resolution of the seismic data volume.

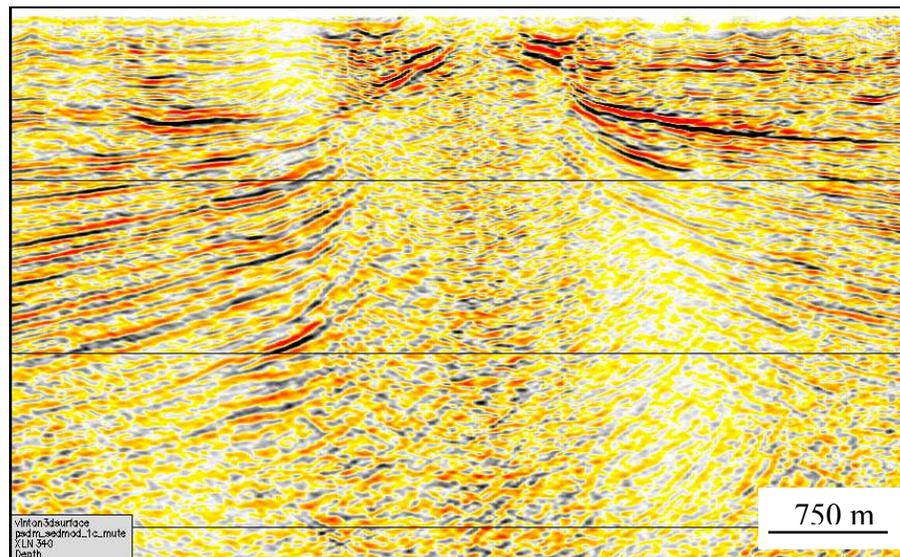
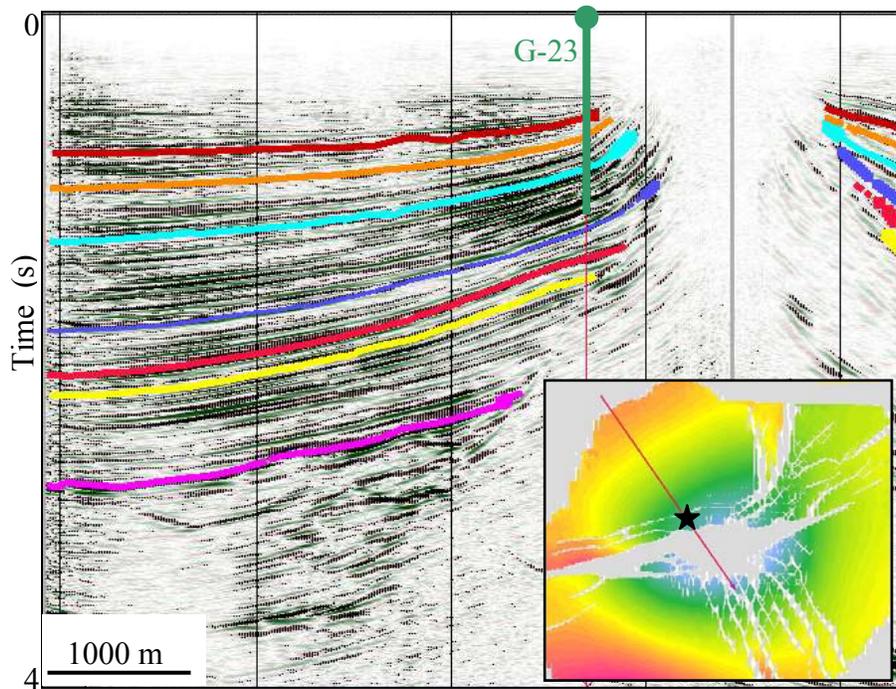
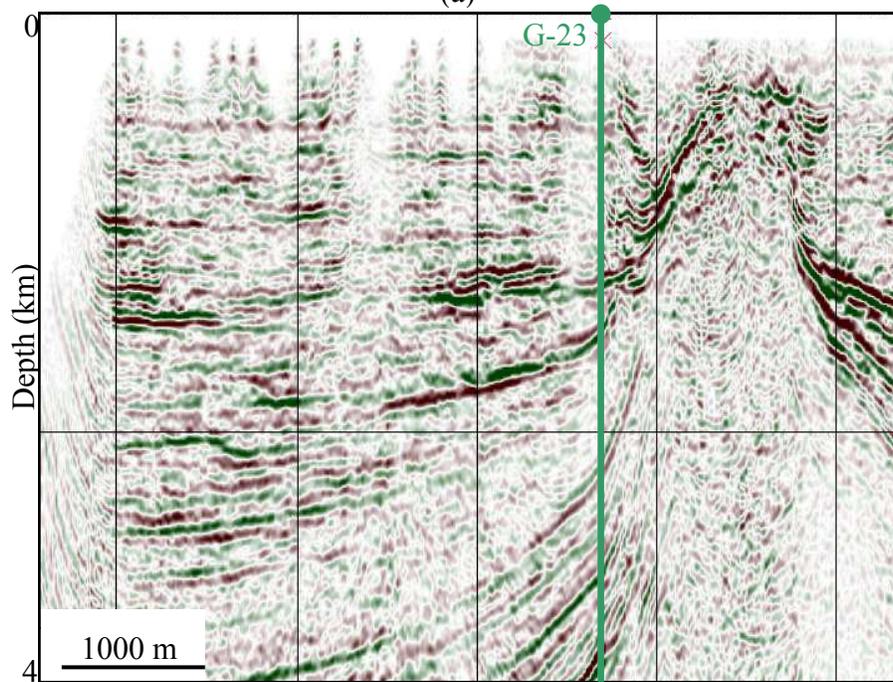


FIG. 2.8. CDP 340 (N-S) after depth migration (2440 m). Seismic data courtesy of OPEX.

Shots recorded by the VSP array are located in the northwestern area of the survey. Based on the previous interpretation and a radial profile through well G-23 (Figure 2.9), I constructed a model in GXII modeling software (Figure 2.10). The model has eight layers with the salt dome in the middle. P-wave velocities, V_p , are calculated from the sonic logs. S-wave velocities, V_s and densities are calculated automatically in GXII using the Poisson coefficient of $\sigma=0.25$ and Gardner coefficient value of 0.23.



(a)



(b)

FIG. 2.9. Seismic traverse in time (a) and depth (b) through well G-23. Position of the traverse and well is indicated with a star on captured figure (a). Seismic data courtesy of OPEX.

Positions of VSP receivers correspond to G-23 well receivers array. The shallow geophone is at 287 m and next 60 geophones are 15 m apart. There are 11 shots. The shot separation is 30 m. The closest shot has offset of 3 m, while the largest offset is 3356 m. Transmitted P and S waves and reflected PP, SS, PS and SP waves were recorded from the model as if using a vertical geophone. Shots 1, 6 and 11 are compared (Figure 2.11). The ray propagation geometry is different (captured figures) because of the different offsets.

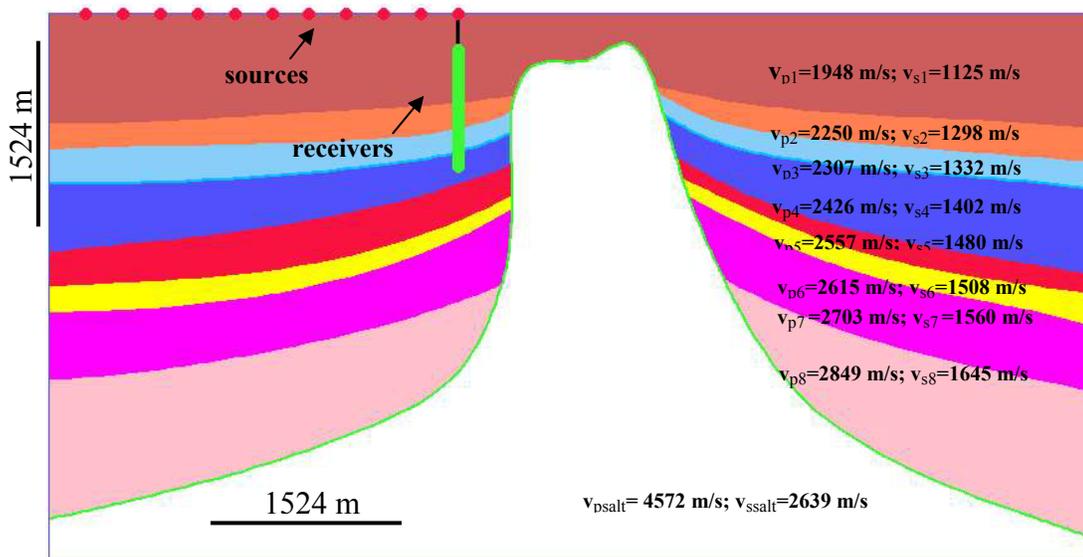


FIG. 2.10. The Vinton Dome model. The model is made in the GXII modeling software. P-wave velocities are calculated from the sonic logs. S-wave velocities and densities are calculated automatically in GXII. Receivers in well G-23 are marked in green and position of sources are marked in red.

The reflections and converted waves from five horizons, together with the strong salt reflections were recorded on shot 1. The color of the arrivals corresponds to the color of the Vinton Dome horizon interpretation colors. The reflections from all seven horizons are recorded on shots 6 and 11.

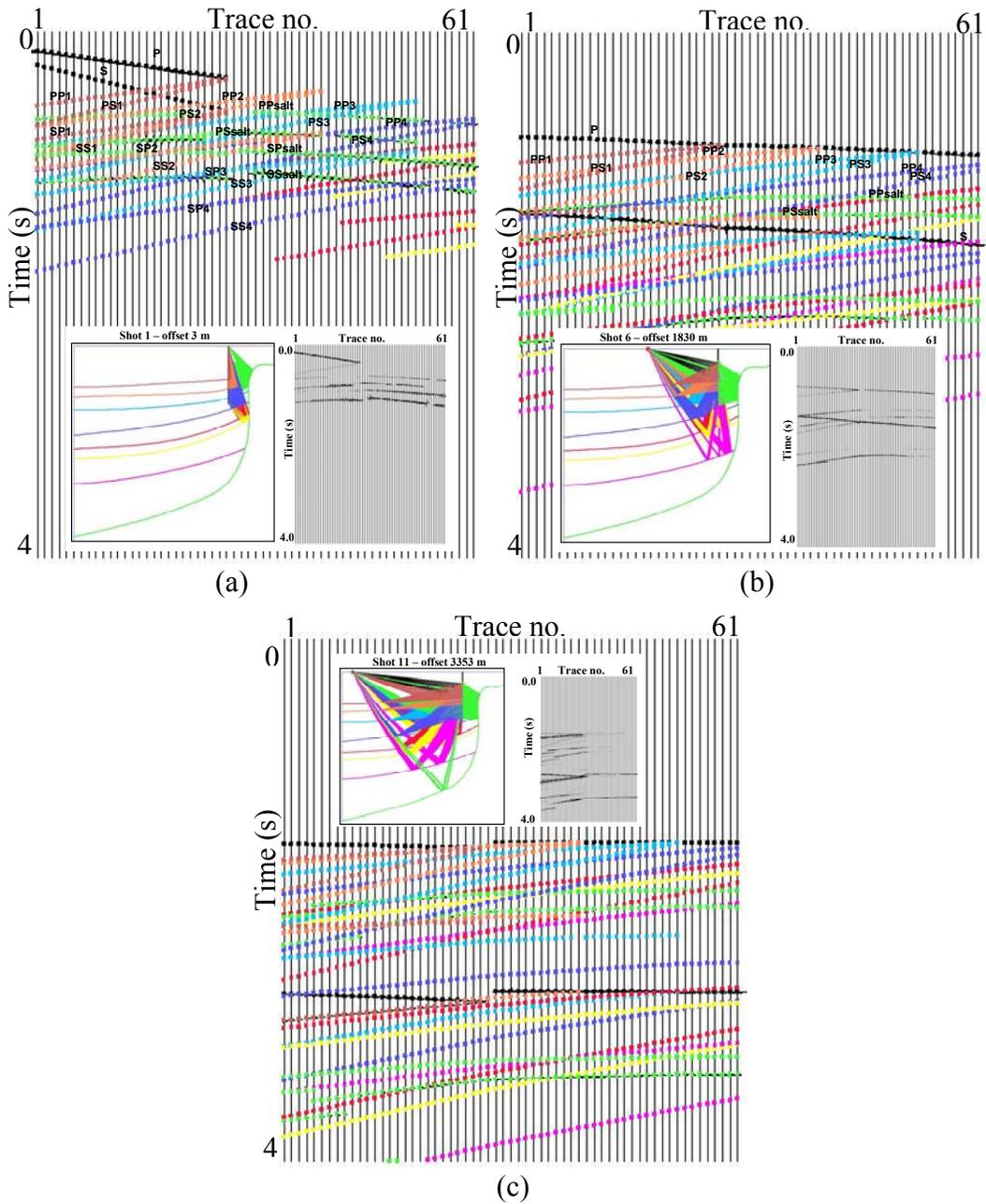


FIG. 2.11. (a) Shots 1, offset 3 m, (b) shot 6, offset 1830 m and (c) shot 11, offset 3353 m. Ray propagation geometry and recorded traces without a gain are displayed in the captured figures. Colors of VSP arrivals correspond to the Vinton Dome horizons interpretation colors. Reflections from salt are green. Some PP, PS, SP and SS arrivals are marked such that the numbers with reflected and converted waves correspond to the horizons numbers.

Because of a large offset, the downgoing arrivals of shot 11 have dips similar to those of the upgoing arrivals, making the upgoing-downgoing wavefield separation difficult.

I made a simple flat layer model to see the influence of the salt dome (Figure 2.12). Reflections from salt arrive at the same time as the other event reflections. The reflections from the salt show hyperbolic arrivals, while reflections from other horizons have more linear events. Salt reflection arrivals are more linear in the case of far offsets, making the separation of salt events particularly hard (linear trend of salt reflections is similar with the other events). In general, the separation of different VSP wave modes depends on two factors: offset and presence of salt.

Unfortunately, GXII modeling software is ray theoretical and some wave phenomena such as diffractions, head waves or surface waves are only approximately modeled. Downgoing refracted waves are not modeled at all. I used a 2-D Elastic Pseudospectral Modeling algorithm to generate a full wavefield synthetic data. 2-D Elastic Pseudospectral Modeling algorithm (EPS2D) developed by Kurt J. Marfurt and Kwangjin Yoon is a wave equation based algorithm. For user specified P and S wave velocities, densities and anisotropy parameters, EPS2D generates a suite of common shot gathers and snapshots. Geometry and trace headers are exported from GXII and used in EPS2D. Synthetic traces and arrivals generated in GXII can be used to judge the arrival times and different wave modes in the synthetic data generated by EPS2D. Data are recorded on the horizontal (Figure 2.13a) and on the vertical component geophone (Figure 2.13b). We can observe more different wave arrivals (downgoing refracted P and S waves) comparing to GXII synthetics (Figures 2.11).

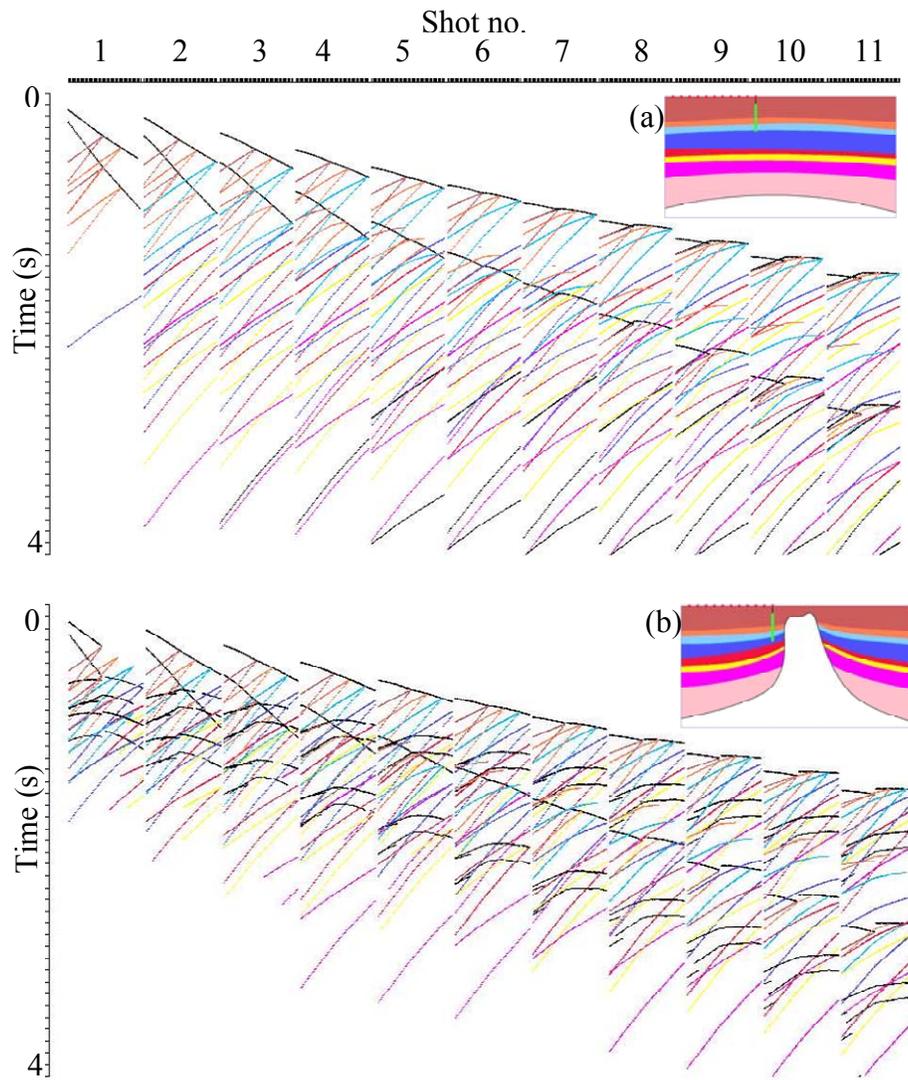


FIG. 2.12. A flat VSP model with recorded arrivals for 11 shots (a). A VSP model with a salt dome in the middle and recorded arrivals for 11 shots (b). Model values and recording geometry are the same. Reflections from salt are black.

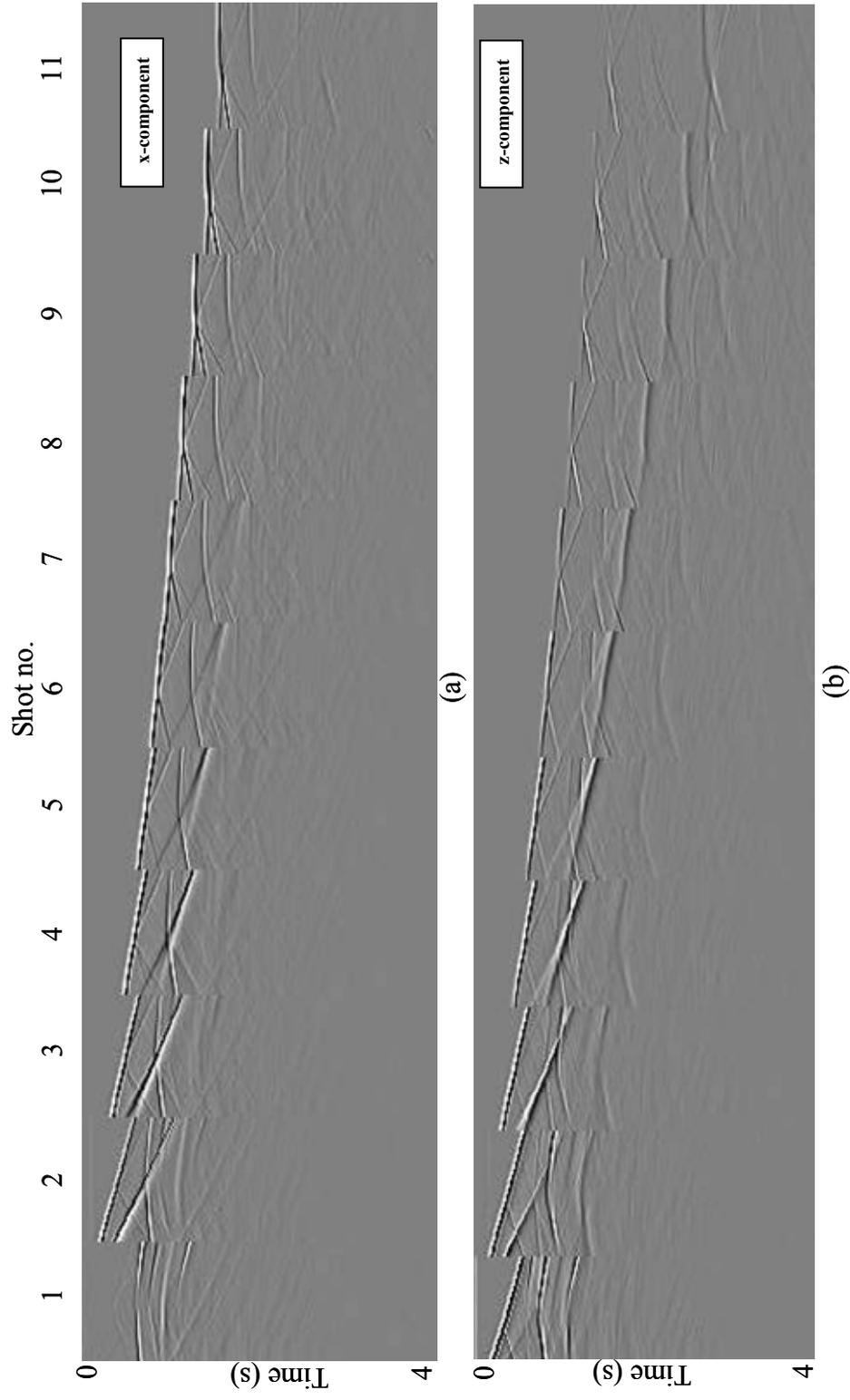


FIG. 2.13. EPS2D synthetic data for 11 shots: x-component (a) and z-component (b).

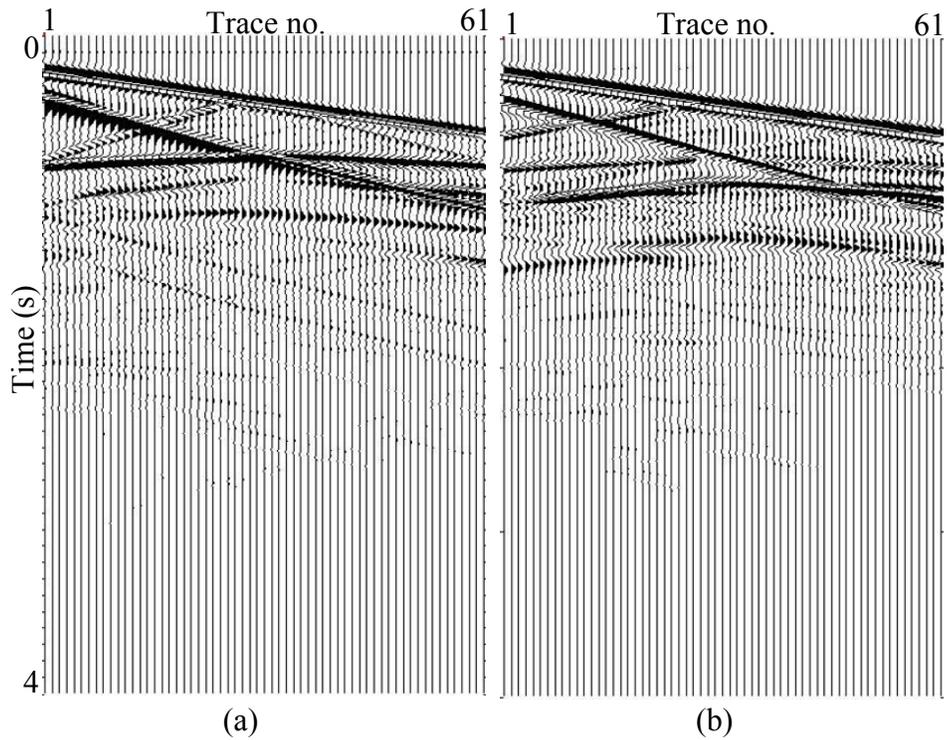


FIG. 2.14. EPS2D synthetic data for shot 2, offset 300 m. Horizontal component (a) records more of the downgoing and upgoing S arrivals and P-wave salt reflection. Vertical component (b) records more of the downgoing and upgoing P arrivals and S wave salt reflection.

The horizontal component of shot 2, offset 300 m (Figure 2.14), records more of the downgoing and upgoing S wave arrivals, while downgoing and upgoing P arrivals are recorded more on the vertical component. However, the opposite is true for the salt reflections: strong P salt reflections are recorded on the horizontal component and strong S salt reflections are recorded on the vertical component.

2.5. Separation of synthetic VSP data using the classical processing approach

Established VSP processing steps include:

1. Defining and editing the geometry;
2. Rotating the horizontal components into radial and transverse, or north and east components (in the case of multicomponent 3-D VSP). The original orientation of two horizontal geophones down in the borehole is determined using the polarization information measured on the first P arrivals from the source locations at varying azimuths (Boelle et al., 1998, Roche et al., 1999);
3. Calculating and applying surface statics. If there is surface seismic, the source static corrections for a 3-D VSP can be derived from the surface data volume and applied to the VSP gathers to bring them to a common datum (Chopra et al., 2004);
4. Picking of direct arrivals;
5. Separating upgoing and downgoing wavefields using the median, f - k or τ - p filters;
6. Deconvolving the upgoing wavefield to help balance the frequencies, improve the phase, and eliminate multiples. The downgoing waves are used for the construction of the deconvolution operator.
7. Applying the 3-D VSP-CDP transform. For dipping interfaces, the amplitudes on a single VSP trace are mapped onto several traces on the x - t plane, where x is a lateral distance of reflection points from the borehole. This ray tracing procedure is called the VSP-CDP transformation.
8. Applying VSP migration which maps the amplitudes along semi-elliptical trajectories whose focal points are the source and receiver locations. Superposition of all these trajectories yields the migrated section (Yilmaz, 2001).

In the case of my synthetic data, there are no edits or static corrections. My modeled VSP is two dimensional with only one horizontal component oriented in the source-receiver plane. Thus after generation, the synthetic VSP data are ready for mode separation.

2.5.1. P and S wave separation using the $f-k$ filter

I performed $f-k$ VSP mode separation using "Seisup", Geocenter's commercial processing package. The $f-k$ spectrum of the horizontal component shown in the previous figure reveals strong P and S wave downgoing arrivals (Figure 2.15). Additionally, we can observe strong salt reflections close to the zero wavenumber. Due to the hyperbolic nature of the salt reflections, these events are mapped both in the upgoing and downgoing part of the spectrum.

I extracted different wave modes using only the moveout component of the seismic data by an appropriate mute selection in the $f-k$ spectrum (Figure 2.16). However, my results suffer from several artifacts, including the frequency loss due to a problematic mute selection.

In order to enhance the salt reflection only, I passed nearly horizontal arrivals (Figure 2.17). Unfortunately, this approach allows other events with the small wavenumbers (including aliases of steeper events) to be extracted as well. More important, due to the parallel nature of the salt hyperbolic reflections, I cannot separate P from S salt reflections because they map in the same region of the $f-k$ spectrum.

Though it is easy to implement, the $f-k$ separation performs poorly on the Vinton Dome synthetic VSP data.

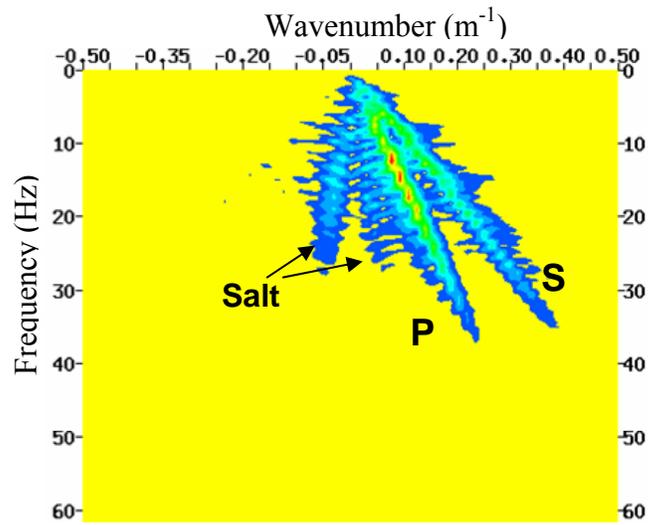


FIG. 2.15. The f - k spectrum of the horizontal component from Figure 2.14.

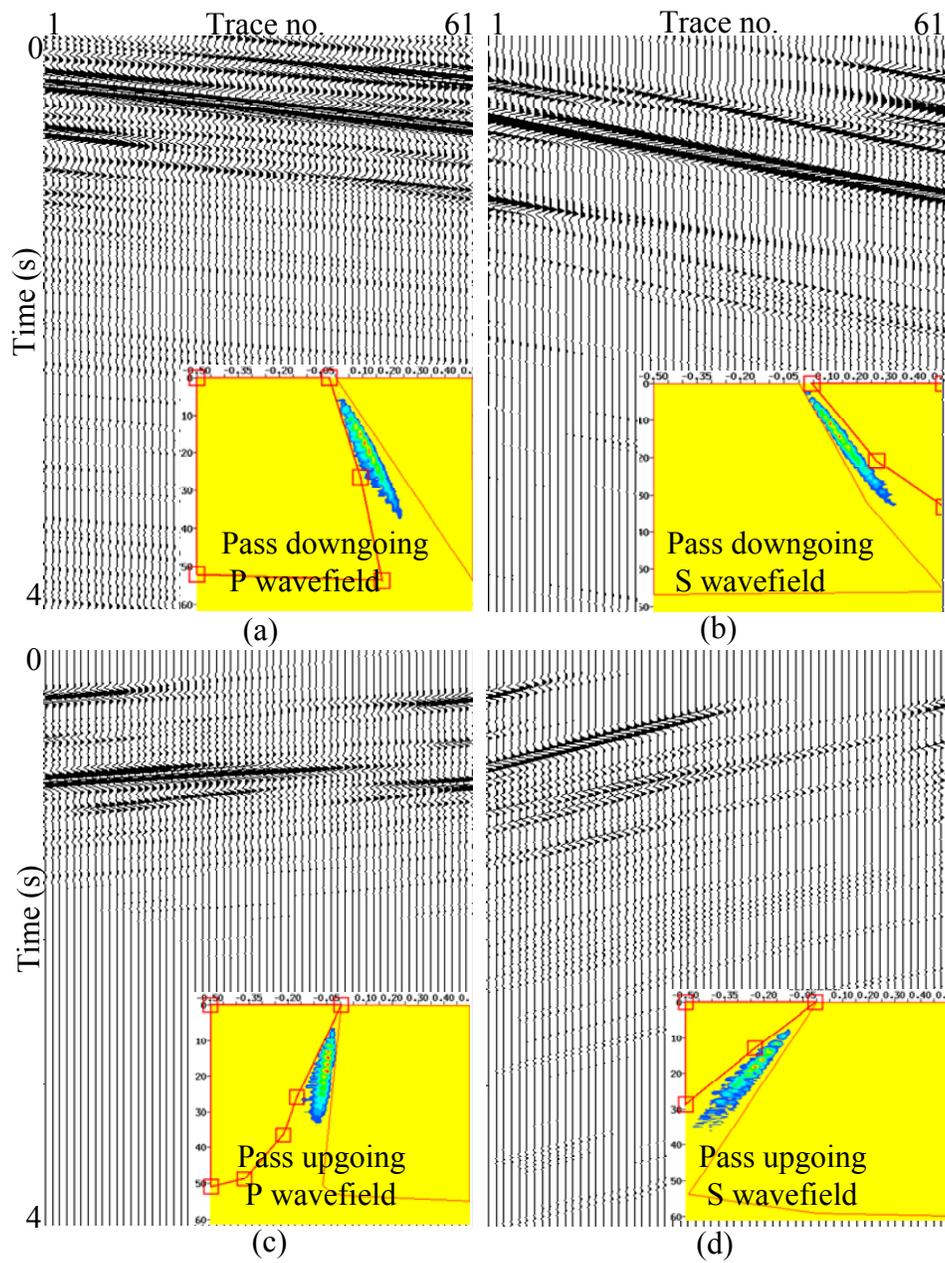


FIG. 2.16. Selecting the proper mute in the $f-k$ spectrum, downgoing P-waves (a), downgoing S-waves (b), upgoing P-waves (c) and upgoing S-waves (d) are separated from the rest of the wavefield.

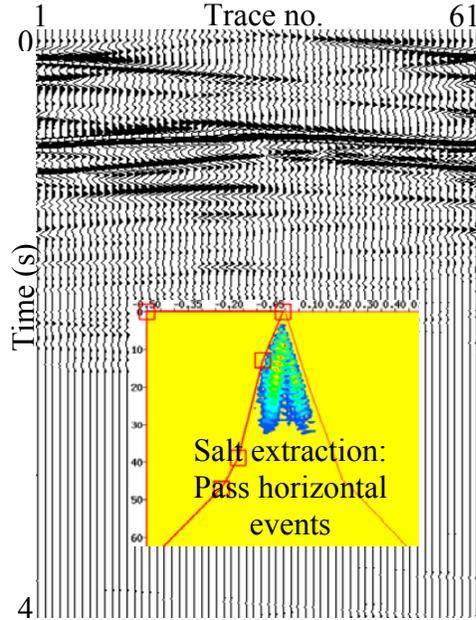


FIG. 2.17. Muting the events with the bigger wavenumbers, horizontal salt reflections are separated from the rest of the field. However, this approach cannot separate P from S salt arrivals.

2.5.2. P and S wave separation using the median filter

Separation of different wave modes can also be achieved using the median filter.

If we arrange N samples in ascending order of magnitude, then the median value is the sample in the $(N+1)/2$ position of the sequence. The median filter is useful in VSP data processing because it rejects noise spikes and because it passes the step functions without altering them (Hardage, 1983).

Separation by median filter is done on shot 2 in "Focus", Paradigm Geophysical processing software (Figure 2.18). Downgoing P arrivals are picked, flattened and extracted by median filter.

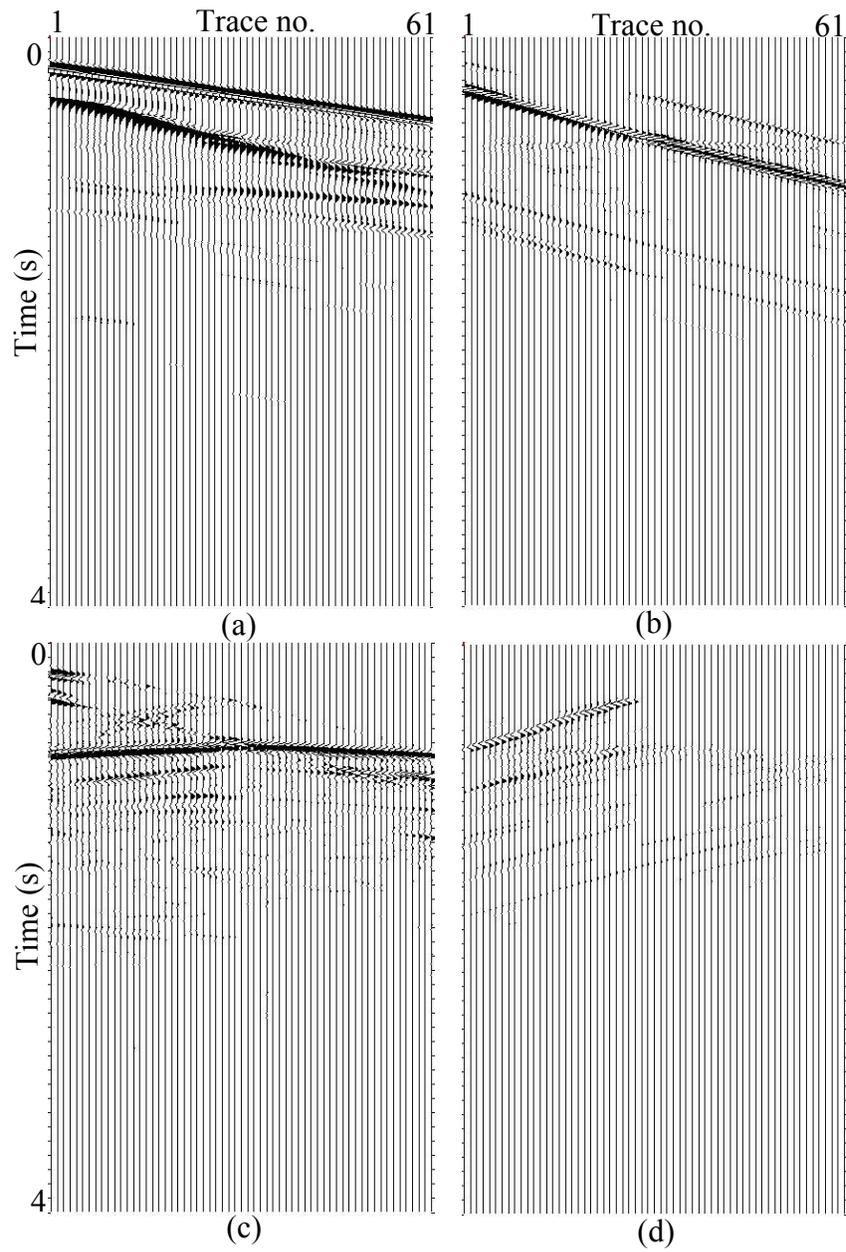


FIG. 2.18. The median filter separates x-component of the data shown in Figure 2.14 into downgoing P (a), downgoing S (b), upgoing P (c) and upgoing S (d) wavefields. The median filter separates salt reflections only partially.

After subtraction from the original data and after 'un-flattening', I obtained the downgoing S with the complete upgoing wavefield and separated downgoing P waves. Downgoing S is then picked, flattened and extracted by a median filter. The same procedure is applied to the upgoing S (stronger than the upgoing P on the horizontal component) such that I obtained the upgoing P wavefield.

The median filter successfully separates flat, linear arrivals such as downgoing S arrivals (Figure 2.18b). Unfortunately, median filter separates salt arrivals only partially, taking the upgoing or downgoing, i.e. flattened events (Figure 2.18a) of the hyperbolic salt arrivals.

2.5.3. P and S wave separation using combined f -k and median filter

In the combined separation approach, I first separated downgoing from the upgoing wavefield using the f -k filter. As we saw before (Figure 2.15), downgoing and upcoming waves map in a different half-plane in the f -k domain. I muted the negative wavenumber components and retained the positive wavenumber components to reconstruct downgoing wavefields for the horizontal and for the vertical component (Figure 2.19a and 2.20a). I reversed this process to reconstruct the upgoing wavefields in Figures 2.19b and 2.20b. Unfortunately, originally hyperbolic PP, PS, SP and SS salt reflections are divided apart with the muting, such that the salt moveout is on both upgoing and downgoing panels.

In the horizontal downgoing wavefield, I picked and flattened P-wave arrivals. Flattened arrivals are vertical in the f -k spectrum because the wavenumbers of the horizontal events are close to zero. Since most of the P and S salt dome reflections are parallel events in the t - x space (Figure 2.14), I was unable to separate P from S salt

reflections using a mute in the f - k spectrum. However, using the fact that we have mostly P-waves from the salt on the horizontal component and that all salt reflections have wavenumbers close to zero, I muted flattened downgoing P events together with both strong P and weak S salt reflections. After application of this filter, we are left with the slower S-wave events of the downgoing wavefield. Both downgoing P and downgoing S are then unflattened to obtain separated downgoing P-waves with all salt events and S-waves.

Separation of upgoing P and S wavefields is similar, except that upgoing P-waves can be flattened by adding the previously picked downgoing P-wave events. As in the case of downgoing separation, flattened upgoing arrivals are extracted using the f - k filter. The upgoing P wavefield includes strong P-waves and weak S-waves reflections again. Separated downgoing and upgoing P and S waves for the horizontal component are shown in Figure 2.21.

The same procedure is applied for the vertical component wave mode separation (Figure 2.22). The vertical component records more of the S salt reflections, so I separated downgoing and upgoing S-wave arrivals together with the weak P and strong S salt reflections.

Alternatively, instead of downgoing-upgoing separation, I could pick the downgoing P-waves first, flat and mute, and repeat the whole procedure for each of the events.

This application procedure is valid for the near offset shots only. In the case of long offsets, selection of the appropriate upgoing-downgoing filters in the f - k analysis is

almost impossible because of the similar dips of downgoing and upgoing waves (Figure 2.12 and 2.13).

In order to evaluate f - k filtering separation results, I compared wavefields with the input data for both horizontal and vertical component (Figure 2.23 and 2.24). By summing the downgoing and upgoing P waves, I reconstruct a P wavefield (Figure 2.23b) and by summing the downgoing and upgoing S waves, I reconstruct an S wavefield (Figure 2.23c). Reconstructed shot 2 (Figure 2.23d) is the sum of P and S wavefields. A difference between the original shot 2 and reconstructed data (Figure 2.23e) is due to the incorrect filtering in the f - k separation procedure. A similar comparison is shown for the vertical component separation (Figure 2.24).

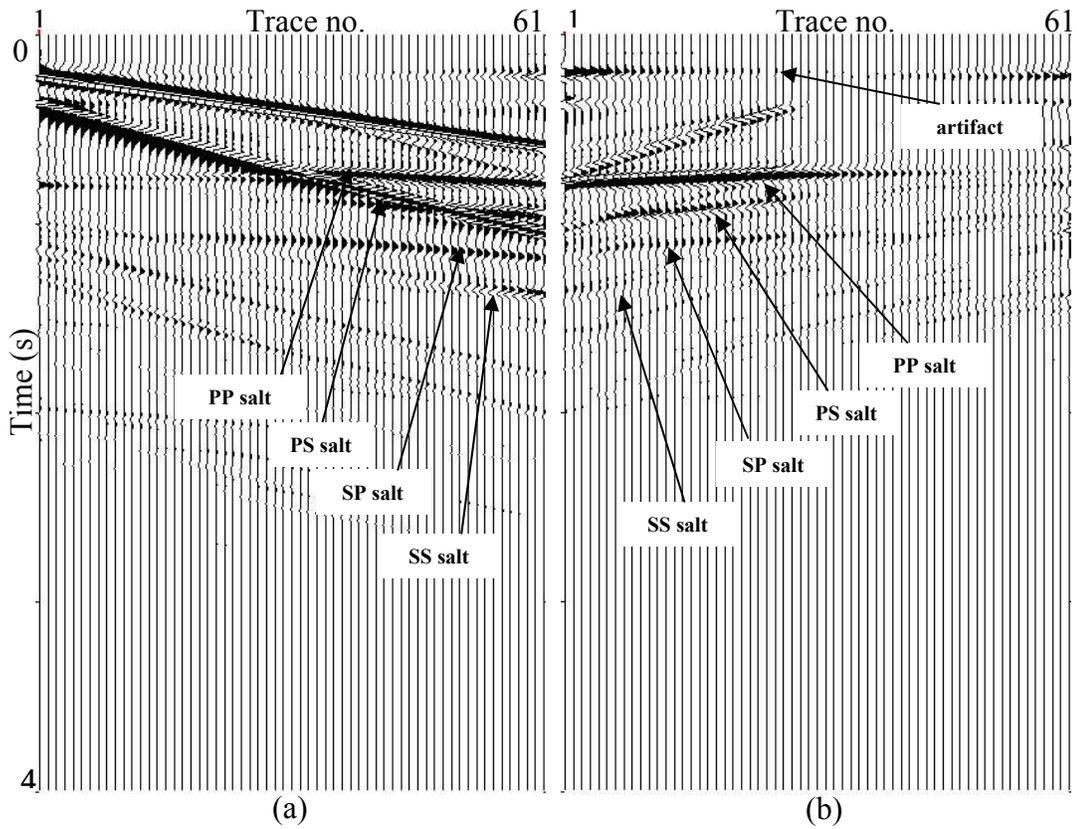


FIG. 2.19. f - k separation of the x-component of the data shown in Figure 2.14 into downgoing (a) and upgoing (b) wavefields. Due to the hyperbolic nature of salt reflections, this mute divides all reflections from the salt such that upgoing parts of PP, PS, SP and SS hyperbolas are separated together with the upgoing wavefield, while downgoing parts of PP, PS, SP and SS hyperbolas are separated together with the upgoing wavefield.

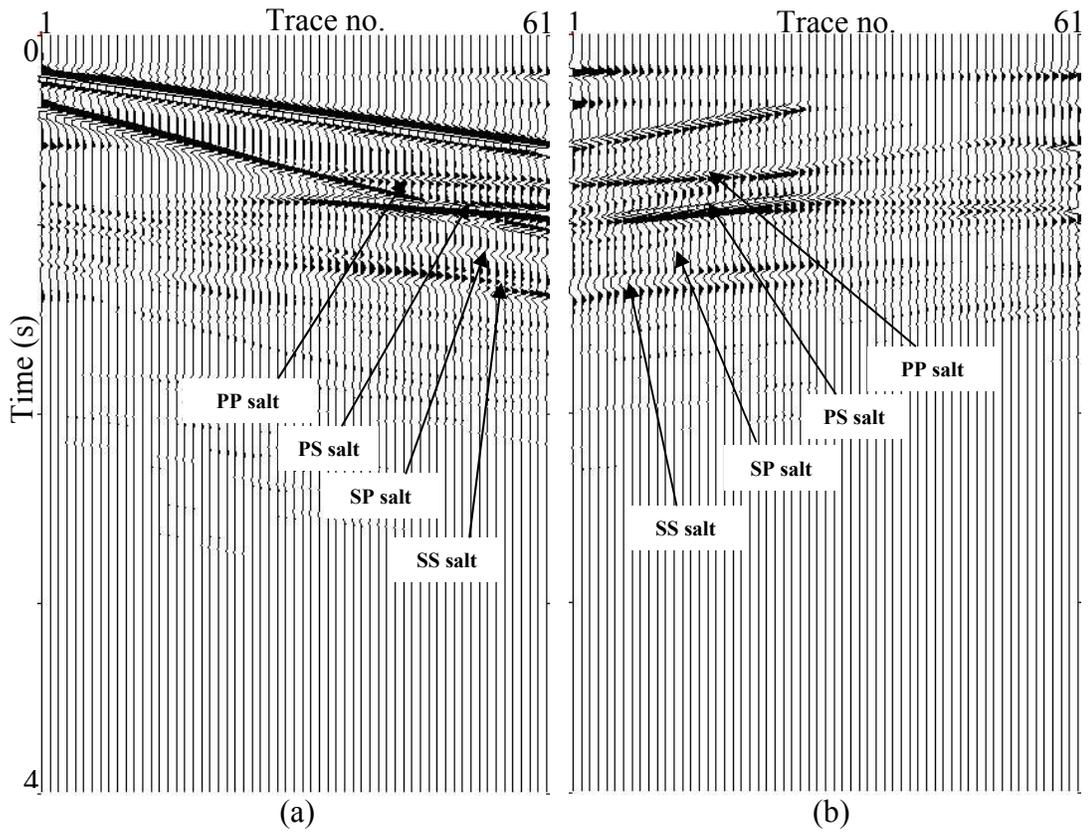


FIG. 2.20. f - k separation of the z -component of the data shown in Figure 2.14b into downgoing (a) and upgoing wavefields (b). Due to the hyperbolic nature of salt reflections, this mute divides all reflections from the salt such that upgoing parts of PP, PS, SP and SS hyperbolas are separated together with the upgoing wavefield, while downgoing parts of PP, PS, SP and SS hyperbolas are separated together with the upgoing wavefield.

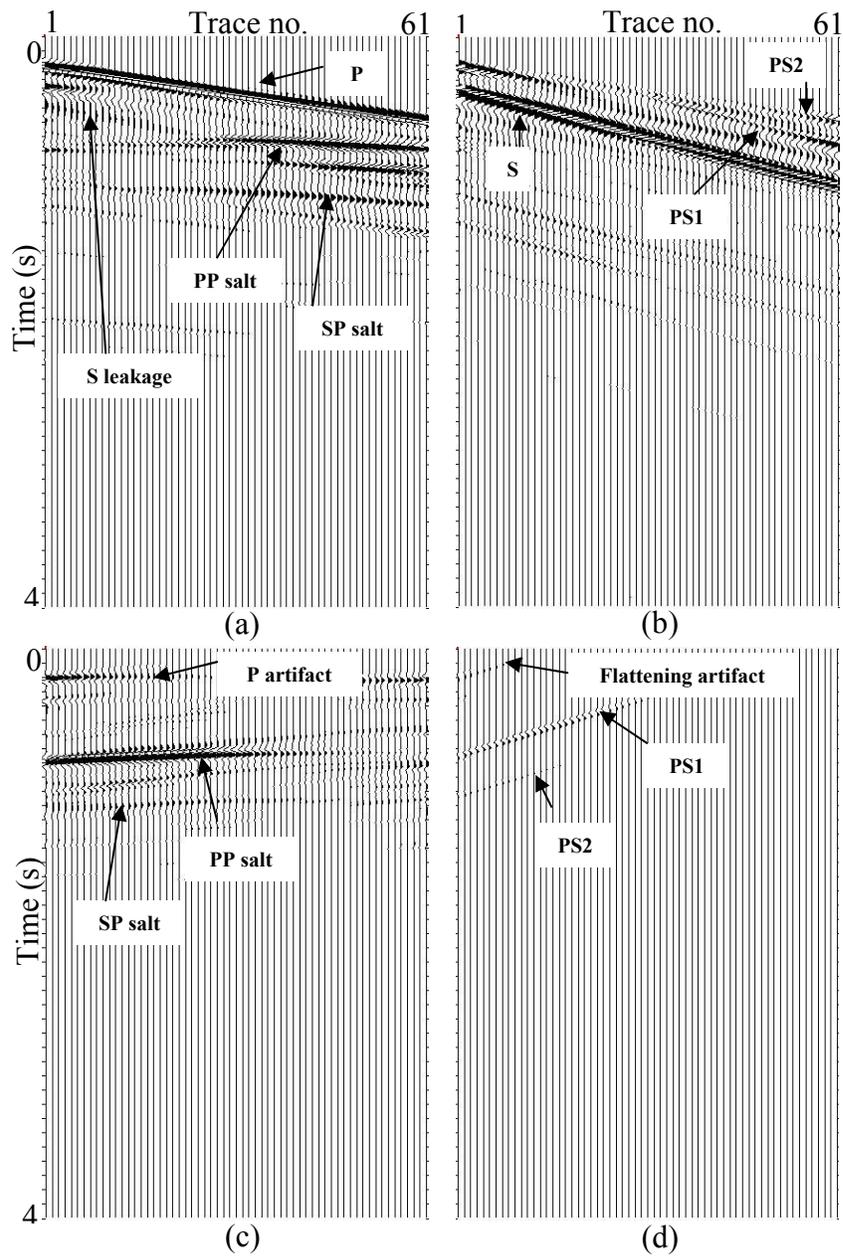


FIG. 2.21. Separated waves for the x-component from Figure 2.14a: downgoing P with both P and S salt events (a), downgoing S (b), upgoing P with both P and S salt events (c) and upgoing S (d). Since I could not separate P from S salt events, I extracted both strong P and weak S salt reflections with P-waves assuming that x-component records more of the P-waves salt reflection.

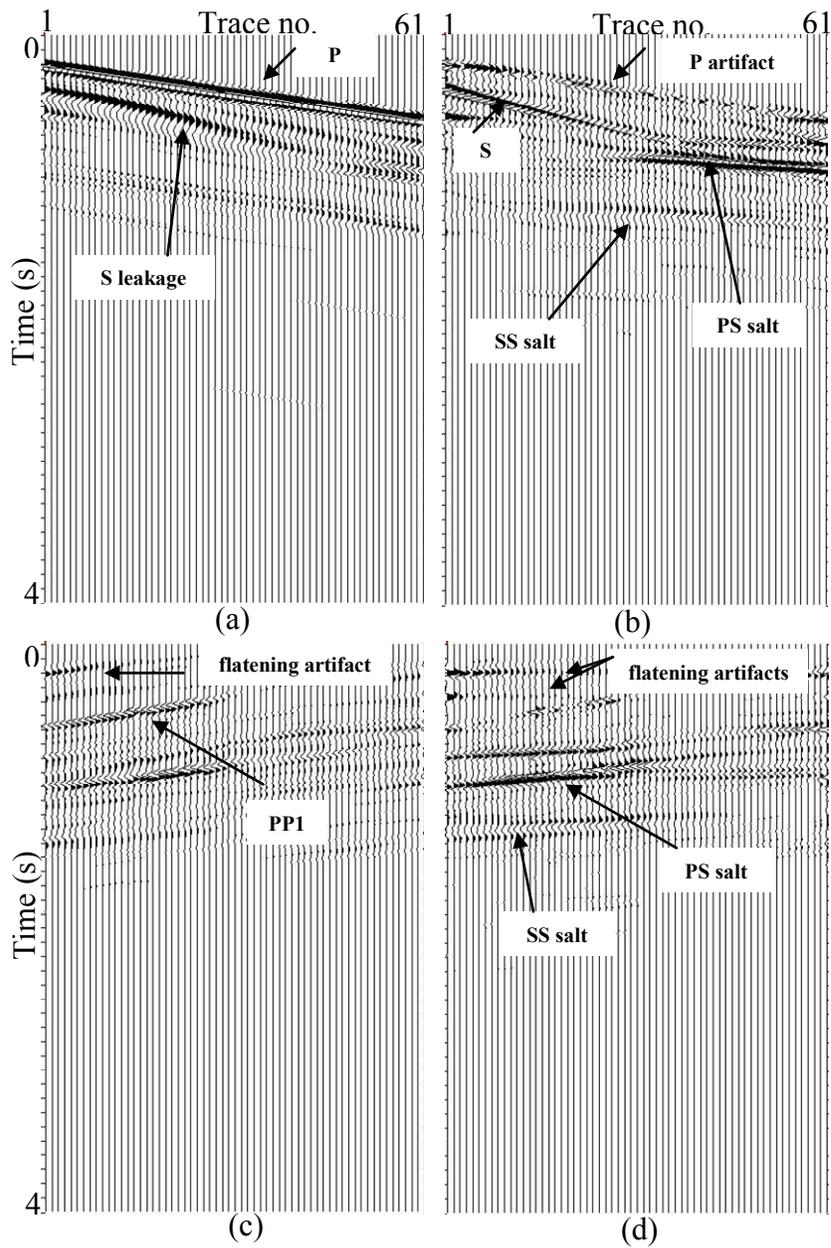


FIG. 2.22. Separated waves for the z-component from Figure 2.14b: downgoing P (a), downgoing S with both P and S salt events (b), upgoing P (c) and upgoing S with both P and S salt events (d). Since I could not separate P from S salt events, I extracted both weak P and strong S salt reflections with S-waves assuming that z-component records more of the S-waves salt reflections.

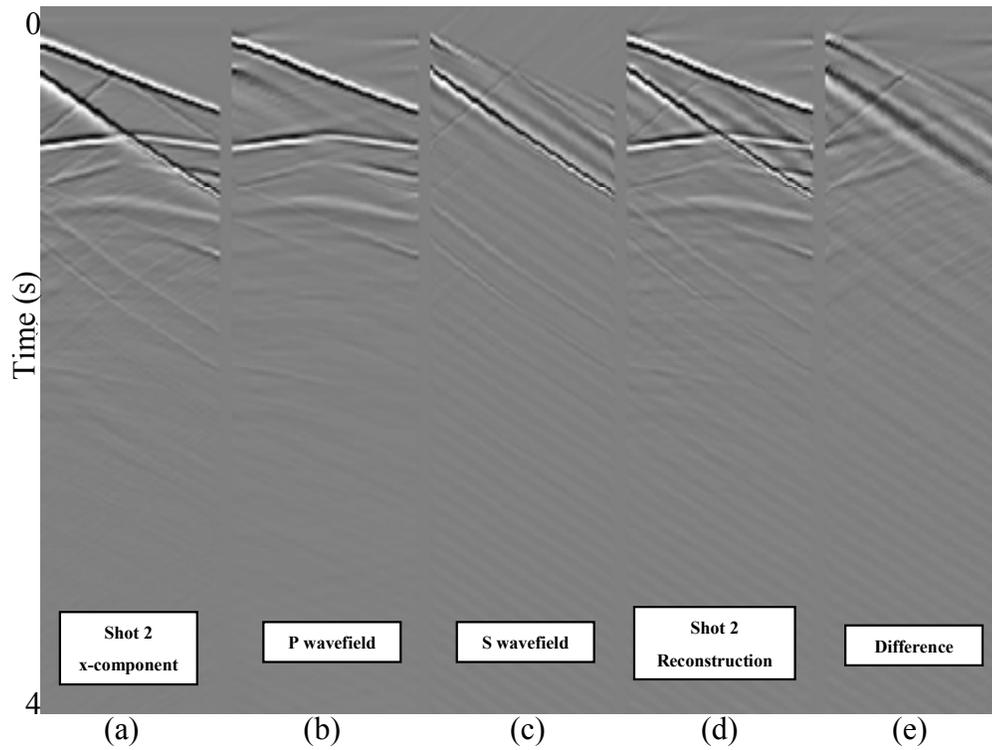


FIG. 2.23. Comparison of the original x-component data (a) and separated wavefields. Summed downgoing and upgoing P waves (b) and summed downgoing and upgoing S waves (c). P wavefield (b) includes all salt reflections for the horizontal component shown in Figure 2.21. Reconstructed data from separated components (d) and difference between the input data and reconstructed data (e).

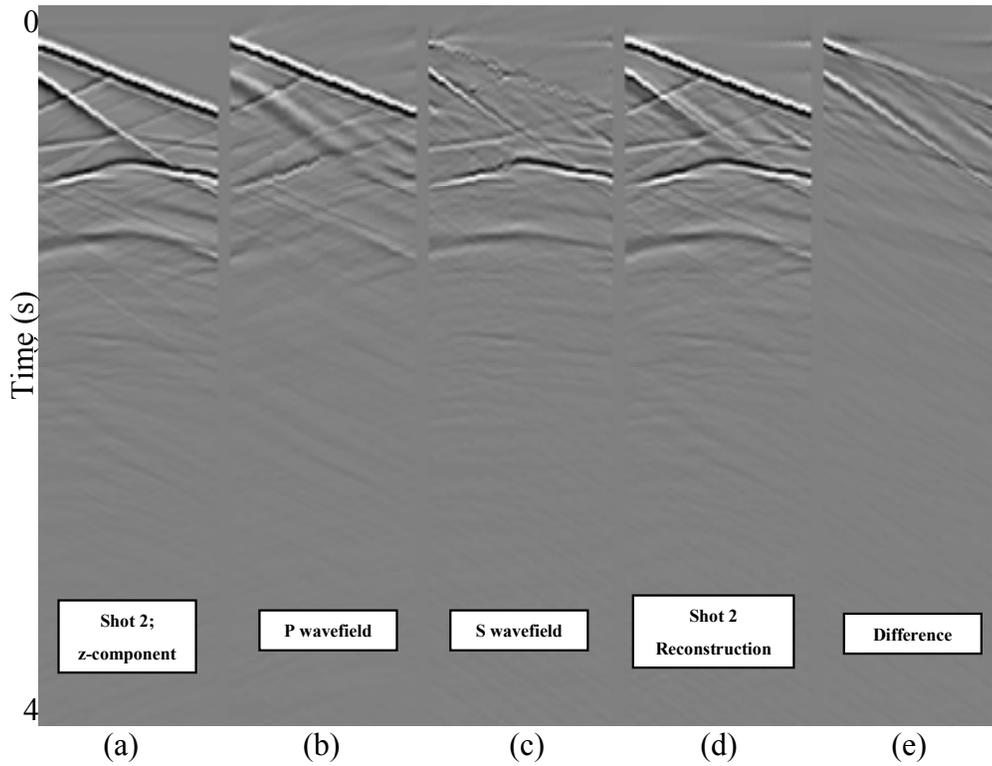


FIG. 2.24. Comparison of the original z-component data (a) and separated wavefields. Summed downgoing and upgoing P waves (b) and summed downgoing and upgoing S waves (c). S wavefield (c) includes all salt reflections for the vertical component shown in Figure 2.22. Reconstructed data from separated components (d) and difference between the input data and reconstructed data (e).

2.6. Separation of Vinton Dome VSP data using the classical processing approach

Though it was shown on the synthetic data that classical separation approaches give poor results in the case of long offsets and presence of salt, I tried the $f-k$ separation on the real data shown in Figure 2.5. We can see from the $f-k$ spectrum of the vertical component that data are highly aliased (Figure 2.25). Beside the downgoing and upgoing arrivals, it is hard to differentiate any particular P or S arrival. For this reason, I limited myself to separate downgoing from upgoing arrivals only. The higher frequencies of the upgoing arrivals are mixed with the aliased part of the downgoing waves (Figure 2.26).

From my previous modeling, I know that most of the events recorded on the vertical component are P waves. However, we record strong S-wave salt reflection on the vertical component too. In order to extract salt arrivals, I tried a method already applied to the synthetic data (Figure 2.17). My results are not unique, and, in addition, I was not able to recover the full frequency spectrum.

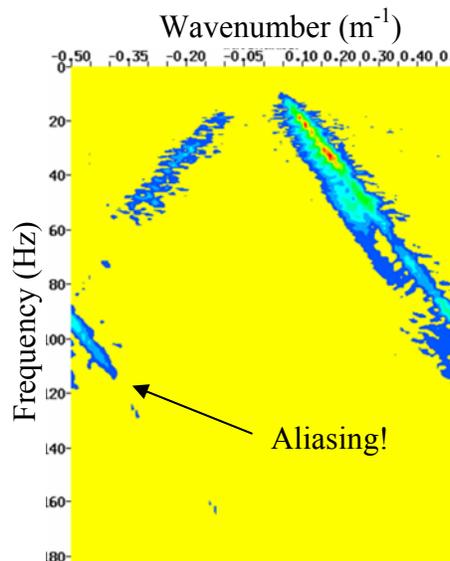


FIG. 2.25. The $f-k$ spectrum of the Vinton Dome data, shot 3068, vertical component, offset 600 m (Figure 2.5). Data are highly aliased.

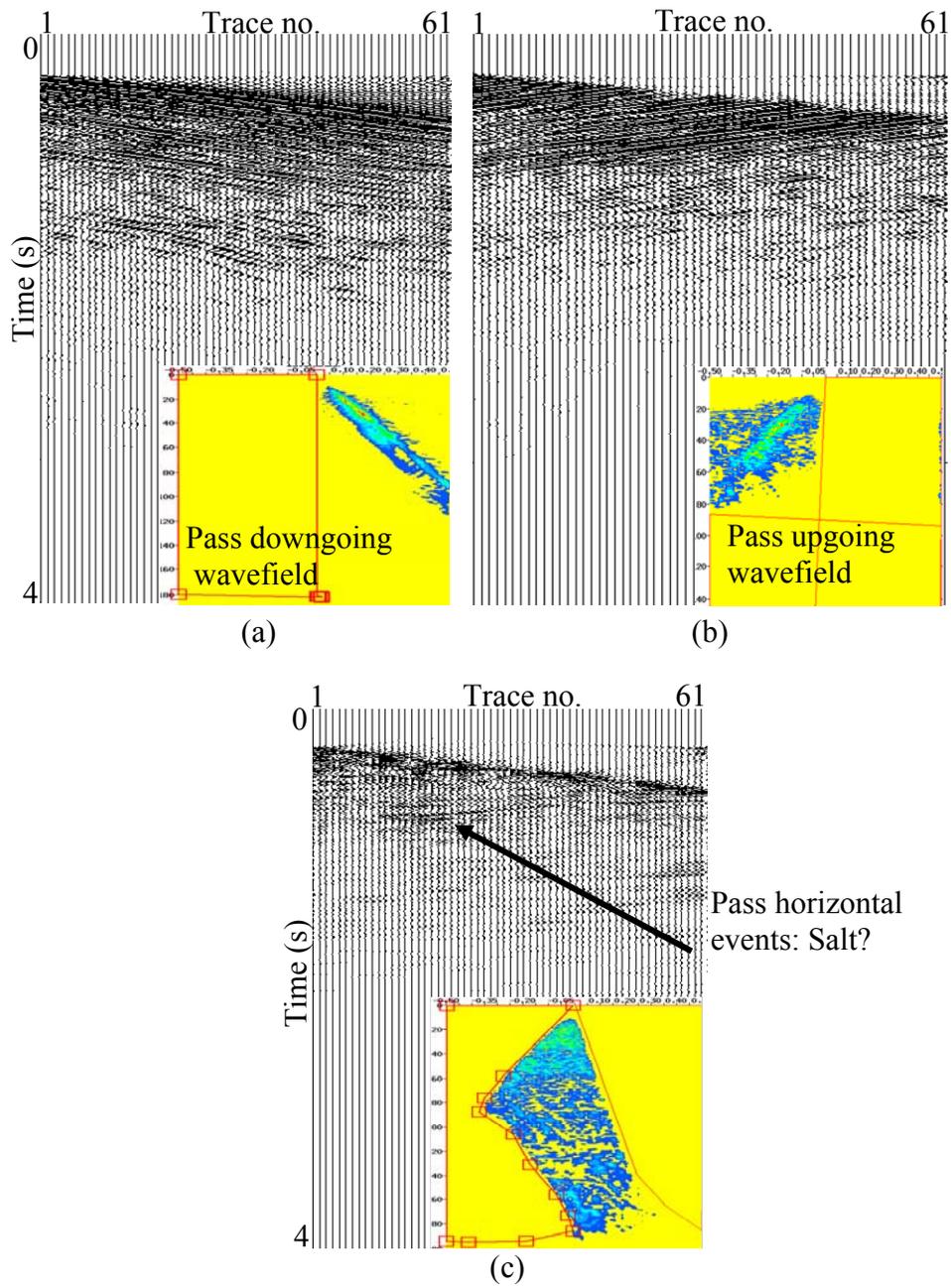


FIG. 2.26. Separated downgoing wavefield (a), upgoing wavefield (b) and horizontal arrivals with possible salt extraction (c).

2.7. Limitations of the classical mode separation approaches

Compared to real data, synthetic data are not aliased and the $f-k$ filter behaves reasonably well. However, incorrect mute in the $f-k$ domain produces significant edge effects that can be seen as artifacts in the $t-x$ space. Besides that, both P and S salt reflections have a more or less similar reflection shape such that these arrivals cannot be distinguished in the $f-k$ space. The salt face reflections arrive about $k \sim 0$ having a wavelength similar to the flattened arrivals wavelength. This complicates correct filtering of the flattened arrivals.

The geophones are cemented in the well, so we do not have bad spikes in the data. Because of that, there is no need to use a median filter which preserves the signal in the case of the abrupt discontinuities in the data. Additionally, the median filter partially separates hyperbolic salt face reflections. If we want to extract whole P or S salt reflection, we need to consider both velocity (moveout) and polarization information content of the different waves.

The high resolution discrete Radon transform successfully exploits velocity, polarization, aliasing and irregular sampling, without subjective mute selection or tedious and time consuming event picking.

3. THEORY AND DEVELOPMENT OF THE HIGH RESOLUTION DISCRETE RADON TRANSFORM VECTOR VSP SEPARATION ALGORITHM

3.1. Propagation and polarization of VSP plane P and S waves

A wave is a disturbance that propagates through a medium without involving net movement of the material (Sheriff, 2002). The relation of the spatial and time dependence of such a disturbance is given by the wave equation. The wave equation is derived from Hook's law and Newton's second law and can be written in vector form as:

$$(2\mu + \lambda)\nabla(\nabla \cdot \mathbf{u}) - \mu\nabla \times (\nabla \times \mathbf{u}) = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}, \quad (3-1)$$

where λ and μ are Lamé's coefficients, ρ is density, t time and vector \mathbf{u} wave displacement. If $\nabla \cdot \mathbf{u} = 0$, the solution of the vector wave equation is an S-wave, and if $\nabla \times \mathbf{u} = 0$, the solution is a P-wave.

Propagation of elastic waves, and their relation with the elastic media they travel through, was first analyzed by Green in 1839, but complete analytical solutions that determined amplitudes of the reflected and refracted waves were independently given by Knott in 1899 and by Zoeppritz in 1907 (Aki and Richards, 1981).

If the wave is sufficiently far away from the source, than the wavefronts incoming to some point can be treated as plane waves (Figure 3.1). A wave propagating in the direction \mathbf{s} with a speed v is a plane wave if:

1. At the fixed time, the quantity that propagates is spatially unchanged over each plane normal to the vector \mathbf{s} , and if
2. The plane associated with a particular value of the quantity moves with speed v in the direction \mathbf{s} (Aki and Richards, 1981).

Figure 3.1 shows a plane wave recorded at a geophone depth, z_1 . In order to understand plane waves, we will set our coordinate system about the surface geophone (x_1, z_1) . The same wave will arrive at depth geophone (x_1, z_2) and surface geophone (x_2, z_1) after some time Δt . The velocity of this arrival is called a phase velocity and it depends on the medium velocity, v , and the angle of propagation, ψ . The horizontal phase velocity is then $\frac{\Delta x}{\Delta t} = \frac{v}{\sin \psi}$ and the vertical phase velocity is $\frac{\Delta z}{\Delta t} = \frac{v}{\cos \psi}$. The inverse of the phase velocity is called the slowness (Yilmaz, 2001).

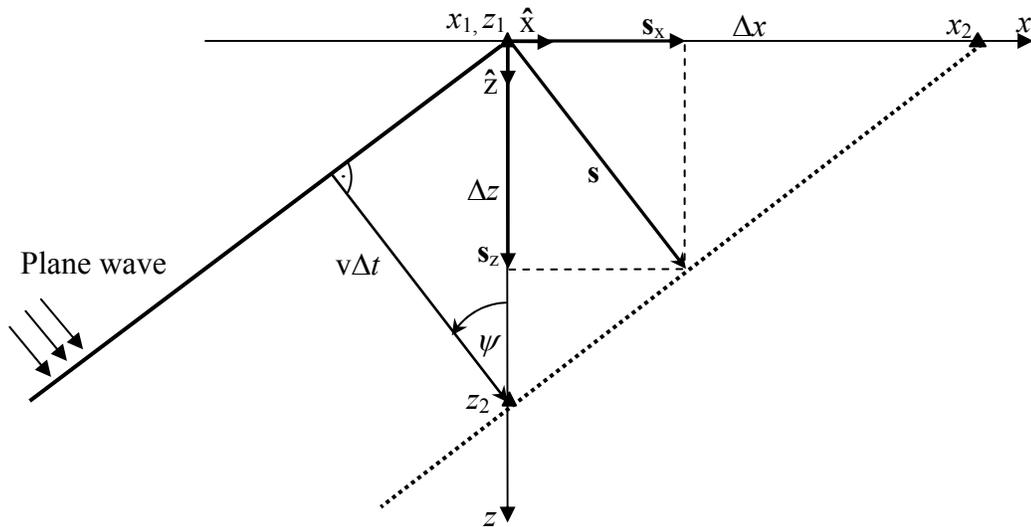


FIG. 3.1. Propagation of a plane wave with the slowness \mathbf{s} . The wave arrives to the VSP geophone array parallel to the unit vector $\hat{\mathbf{z}}$ in a medium with velocity V . The angle between the ray, \mathbf{s} , and geophone array, $\hat{\mathbf{z}}$, is ψ .

The reason for using the slowness rather than velocity to summarize the speed and direction of a wave is that the slownesses may be added vectorially $\mathbf{s} = \hat{\mathbf{x}}\mathbf{s}_x + \hat{\mathbf{z}}\mathbf{s}_z$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are unit vectors in the horizontal and vertical direction (Figure 3.1), while

velocities may not (Aki and Richards, 1981). The magnitude of the horizontal component of the slowness vector is $s_x = \frac{\sin \psi}{v}$, while the magnitude of the vertical component of the

slowness vector is $s_z = \frac{\cos \psi}{v}$. For a P-wave with velocity v_p and angle of propagation ψ_p

(Figure 3.2), the slowness components are $s_{px} = \frac{\sin \psi_p}{v_p}$ and $s_{pz} = \frac{\cos \psi_p}{v_p}$. For an S-wave

with velocity v_s and angle of propagation ψ_s (Figure 3.2), the slowness components are

$$s_{sx} = \frac{\sin \psi_s}{v_s} \text{ and } s_{sz} = \frac{\cos \psi_s}{v_s}.$$

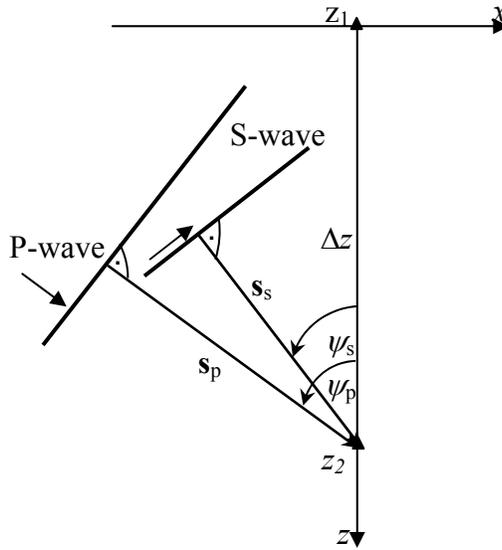


FIG. 3.2. Propagation and arrival of P and S waves to VSP receiver array.

Using Fourier analysis, we can represent any waveform as the superposition of harmonic waves. We define the wave displacement \mathbf{u} as $\mathbf{u} = \mathbf{a} e^{i\omega(\mathbf{s} \cdot \mathbf{x} - t)}$, where \mathbf{a} is the polarized amplitude, ω is frequency, \mathbf{s} is slowness and \mathbf{x} is particle motion of a particular wave. In an isotropic homogeneous medium, a P-wave particle motion is in the direction

of propagation, while an S-wave particle motion is perpendicular to the direction of propagation. In our convention, z is positive downward and x is positive to the right, such that P-wave displacements are positive along the direction of travel, while S-wave displacements are positive to the left of the direction of travel. The angle of approach, ψ , is the angle calculated counter clockwise from the positive direction of z axis (Figure 3.3). We can now write two equations for the wave displacements of P-wave and S-wave:

$$\mathbf{u}_p = \mathbf{a}_p e^{i\omega\left(\frac{\sin\psi_p}{V_p}x + \frac{\cos\psi_p}{V_p}z - t\right)}, \text{ and} \quad (3-2a)$$

$$\mathbf{u}_s = \mathbf{a}_s e^{i\omega\left(\frac{\sin\psi_s}{V_s}x + \frac{\cos\psi_s}{V_s}z - t\right)}. \quad (3-2b)$$

Finally, we resolve the displacements and polarization of both waves in the x -direction and z -direction. This gives four equations for P-wave and S-wave amplitudes in two directions:

$$u_{px} = \sin\psi_p |a_p| e^{i\omega\left(\frac{\sin\psi_p}{V_p}x + \frac{\cos\psi_p}{V_p}z - t\right)}, \quad (3-3a)$$

$$u_{pz} = \cos\psi_p |a_p| e^{i\omega\left(\frac{\sin\psi_p}{V_p}x + \frac{\cos\psi_p}{V_p}z - t\right)}, \quad (3-3b)$$

$$u_{sx} = \cos\psi_s |a_s| e^{i\omega\left(\frac{\sin\psi_s}{V_s}x + \frac{\cos\psi_s}{V_s}z - t\right)}, \text{ and} \quad (3-3c)$$

$$u_{sz} = -\sin\psi_s |a_s| e^{i\omega\left(\frac{\sin\psi_s}{V_s}x + \frac{\cos\psi_s}{V_s}z - t\right)}. \quad (3-3d)$$

Both P and S plane waves can be characterized by the value of their common slowness component along a reflector or measurement surface, or ray parameter, p (Aki and Richards, 1981). The ray parameter p is derived from Snell's law. Since we record

only the vertical component of slowness in a vertical VSP, the ray parameter in our VSP geometry is:

$$p = s_{pz} = s_{sz} = \frac{\cos \psi_p}{v_p} = \frac{\cos \psi_s}{v_s}. \quad (3-4)$$

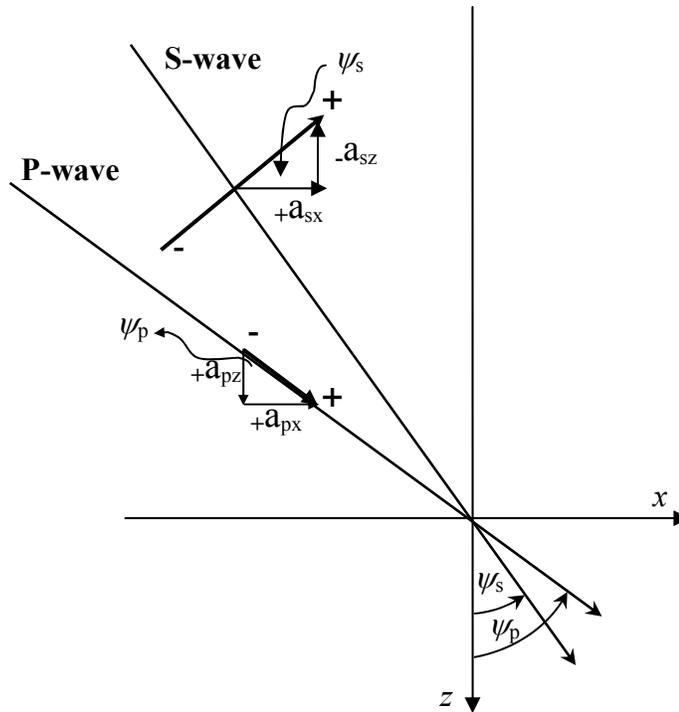


FIG. 3.3. Polarization of P and S waves. In this algorithm, the S-wave amplitude is positive to the left of the direction of propagation, while the P-wave amplitude is positive along the direction of propagation. We measure the angle of approach ψ counter clockwise from the positive direction of z-axis.

Along our vertical measurement direction z , horizontal component x is zero, such that the four equations for particle displacement of P and S waves along the horizontal and vertical component become:

$$u_{px} = \sin \psi_p |a_p| e^{i\omega(pz-t)}, \quad (3-5a)$$

$$u_{pz} = \cos \psi_p |a_p| e^{i\omega(pz-t)}, \quad (3-5b)$$

$$u_{sx} = \cos \psi_s |a_s| e^{i\omega(pz-t)}, \text{ and} \quad (3-5c)$$

$$u_{sz} = -\sin \psi_s |a_s| e^{i\omega(pz-t)}. \quad (3-5d)$$

We use multicomponent receivers to record the full wavefield including P and S waves. Multicomponent receivers have three orthogonal geophones which record the vector sum of each polarized event. If we want to estimate the amplitudes a_p and a_s of the P and S waves, we need to decompose the recorded wavefields.

3.2. Decomposition of vector VSP wavefield into the scalar compressional and shear waves

For a nonzero offset VSP geometry and complex structure, the recorded wavefields become quite complex, with conversion of P-waves to S-waves at each geological interface. The differences in the amplitude of the propagating waves measured on any given receiver component depend on reflection and transmission coefficients, angle of incidence and type of polarization. Figure 3.4 shows a 2-D salt dome model and snapshots of the data recorded by two geophone components: horizontal or radial (oriented in the source-receiver plane) and vertical. For an isotropic medium, P waves are polarized in the direction of propagation and thus the salt face reflections in Figure 3.4a

are recorded more on the horizontal component. Similarly, S waves polarized in the vertical plane in an isotropic medium are polarized perpendicular to the direction of propagation, such that the salt face reflections in Figure 3.4b are recorded more on the vertical component. A multicomponent elastic wave field can be separated into P and S waves by projecting each plane wave component onto appropriate vectors from Figure 3.3.

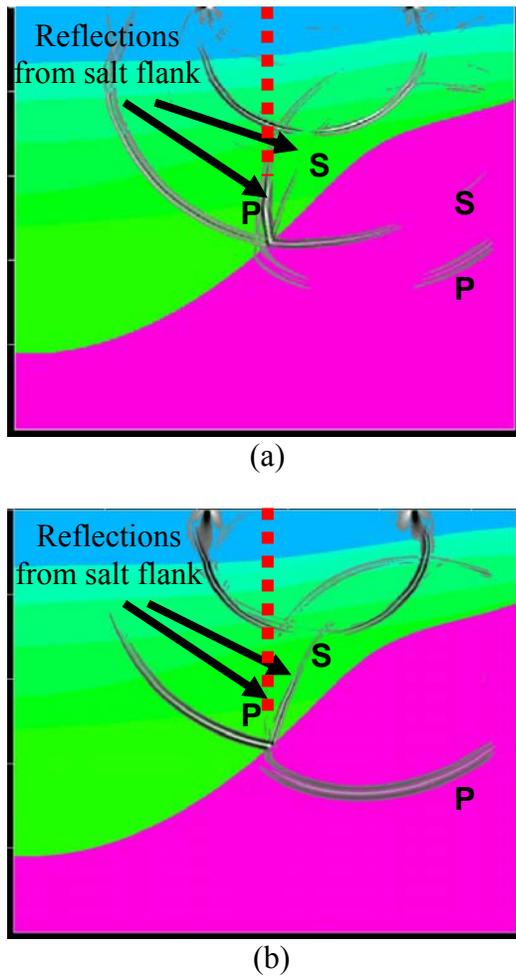


FIG. 3.4. Horizontal (a), and vertical components (b) of wavefield propagation for the Vinton salt dome model and polarization of waves. VSP location is indicated by the dashed line.

We will represent particle motion polarization directions of P and S vectors in the source-receiver plane by the unit vectors $\hat{\mathbf{a}}_p$ and $\hat{\mathbf{a}}_s$:

$$\hat{\mathbf{a}}_p = \begin{pmatrix} \mathbf{a}_{px} \\ \mathbf{a}_{pz} \end{pmatrix} = \begin{pmatrix} \sin \psi_p \\ \cos \psi_p \end{pmatrix} \quad (3-6a)$$

$$\hat{\mathbf{a}}_s = \begin{pmatrix} \mathbf{a}_{sx} \\ \mathbf{a}_{sz} \end{pmatrix} = \begin{pmatrix} \cos \psi_s \\ -\sin \psi_s \end{pmatrix}. \quad (3-6b)$$

Using the particle motion polarization vectors, $\hat{\mathbf{a}}_p$ and $\hat{\mathbf{a}}_s$, incidence angles, ψ_p and ψ_s , and medium velocities, we can uniquely model the vector response of P and SV plane waves.

Our assumption is that an array of waveforms can be modeled as the superposition of a small number of direction-dependent wave fields. In the f - k domain, the difference in time arrival becomes a different linear phase shift for each frequency, such that the data at a given frequency, ω , can be modeled as the sum of N plane waves. Polarization coefficients are then applied and the data inverted back to obtain P-waves and S-waves in the time-depth domain (Devaney and Oristaglio, 1986; Dankbaar, 1987; Esmersoy, 1990). Unfortunately, separation of P and S waves using the f - k transform suffers from spatial aliasing inherent in VSP data acquisition. For this reason we will separate the data in τ - p domain where we can add anti-alias constraints.

3.3. Radon transformation

Plane wave decomposition of a wavefield can be achieved by applying a linear moveout and summing amplitudes over the offset axis. This procedure is alternatively called slant stacking (Yilmaz, 2001), τ - p transformation or when applied in a least-squares sense, a linear Radon transform. Slant stack is basically a transformation of the

offset axes. A linear moveout is applied to the data through a coordinate transformation $\tau = t - px$, where p is the ray parameter, x is the offset, t is the two-way travel time, and τ is the intercept time at $x = 0$. The linear moveout converts all events sloping at a rate p in (x, t) space to horizontal events in (x, τ) space (Claerbout, 1985). The data are then summed over the offset axis by

$$m(p, \tau) = \sum_x d(x, \tau + px), \quad (3-7)$$

where $m(p, \tau)$ is the amplitude of the plane wave with ray parameter $p = \sin \psi / v$ and intercept time τ (Figure 3.5). This linear moveout correction is repeated for a range of p values and followed the summation in equation (3-7), thereby constructing a τ - p gather. The mapping from the t - x domain to the τ - p domain is reversible, so using a inverse linear moveout correction $t = \tau + px$, one can approximately reconstruct data in the t - x domain:

$$d(x, t) = \sum_p m(p, t - px). \quad (3-8)$$

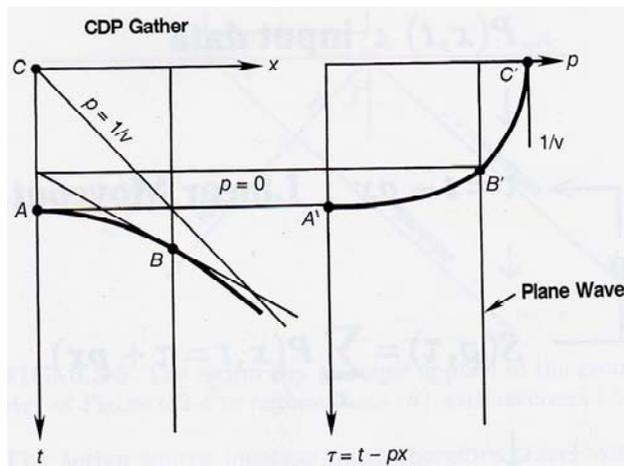


FIG. 3.5. Linear Radon Transformation (after Yilmaz, 2001).

If we want to restore amplitudes properly, i.e. to compensate for the frequency and phase distortion as a result of the summation of discrete samples, we need to apply a *rho* filter (Claerbout, 1985) in the τ - p domain before the inverse transform. The *rho* filtering is achieved by the multiplying amplitude spectrum of each τ - p trace by the absolute value of the frequency. The *rho* factor does not depend on p , so we may apply it either before or after the integration over p .

Once in the τ - p domain, the easiest way to separate P and S waves is to apply polarization coefficients (equation 3-6) in τ - p domain and then apply mutes to separate P-waves from S-waves (Boelle, 1998a; Leaney, 2002). However, when data are aliased, events corresponding to one p value leak into its aliases, such that when we mute our desired (aliased) signal, un-wanted (aliased) noise is preserved and muting does not give the correct result.

3.4. Aliasing

In order to record all frequencies, we need to sample our data with the sampling frequency which is at least twice the highest frequency in our data, giving us two or more samples per period. When there are fewer than two samples per period, two different frequencies can express the same information. This undersampling can lead to an ambiguity known as aliasing. To avoid aliasing, frequencies above the Nyquist frequency

$f_{Nyq} = \frac{1}{2\Delta t}$ need to be removed by an antialiasing filter. However, a seismic wavefield is

a function of both time and space. Like the Nyquist frequency, the Nyquist wavenumber

is defined as $k_{Nyq} = \frac{2\pi}{2\Delta x}$, where Δx is the spatial sampling interval. Data can be aliased if

the spatial sampling interval is too large; for VSP's this happens when the receivers are too far apart. Low frequency data are usually less aliased than high frequency data (Yilmaz, 2001). Even if the input data are not aliased while recording, aliasing can occur due to an inaccurate sampling in the τ - p domain. Undersampling in both τ and p directions can lead to aliasing. Sampling problems in the τ direction are the same as the problems encountered in the sampling of any time series. As the result of an undersampling in the p direction, the data from one trace can be placed on multiple traces back in the t - x domain (Turner, 1990). If we want to avoid aliasing, the ray parameter interval Δp has to be

$$\Delta p < \frac{1}{2x_r f_{\max}}, \quad (3-9)$$

where x_r is the range of offset values and f_{\max} is the maximum frequency in the data (Turner, 1990). In the case of one way travel time, the spatial sampling interval should be less than or equal 1/2 the wavelength of the highest frequency, because spatial aliasing occurs when the wavefront separation equals half the dominant period (Yilmaz and Taner, 1994).

In the f - k domain, aliased events appear to "wrap-around" the left and right ends of the k -axis. In the τ - p transform, aliasing is usually seen as the dispersion of otherwise localized seismic events. The usual way to remove these aliases is to mute the data in f - k or conventional τ - p domain. Unfortunately, the useful signal is very often removed together with the aliased noise, so that these transforms can suffer from severe inverse transform artifacts. At the same time, the useful signal can be contaminated with the aliased noise because aliasing affects filtering two ways: aliased noise can leak into the

signal pass band and aliased signal may leak into the noise reject band (Marfurt et al., 1996). One of the approaches to solve the aliasing problem is to use the preconditioned least-square discrete Radon transform (DRT) instead of the conventional one.

3.5. Least-square linear discrete Radon transform

The integral version of the Radon transform gives the plane-wave decomposition exactly. The discrete τ - p transform given by equation (3-8) can be seen as the map from model space $\mathbf{m}(\tau, p)$ to data space $\mathbf{d}(t, x)$. Using the inverse theory, one can write:

$$\mathbf{d} = \mathbf{Lm} , \quad (3-10)$$

where \mathbf{L} is the linear forward modeling operator. To solve equation (3-10) for \mathbf{m} in a least-square approach, one first forms the normal equations:

$$\mathbf{L}^T \mathbf{d} = \mathbf{L}^T \mathbf{Lm} \quad (3-11)$$

and then inverts to obtain

$$\mathbf{m} = (\mathbf{L}^T \mathbf{L} + \varepsilon \mathbf{I})^{-1} \mathbf{L}^T \mathbf{d}, \quad (3-12)$$

where the term ε and identity matrix, \mathbf{I} , insure a minimum energy solution (Menke, 1984). The discrete Radon transform represents the result of a least-mean-squares solution of a set of equations. These equations are set up to minimize the differences between the inverse and the original data set. Such a decomposition is not unique because solutions can be different based on the limits of the ray parameter p (Yilmaz and Taner, 1994). This classical minimum energy solution generally behaves poorly on aliased data (Marfurt et al., 1996). In the frequency domain, a discrete Radon transform decouples into matrices having a Toeplitz structure which results in a computationally efficient algorithm (Kostov, 1990). Unfortunately, the algorithm fails in the case of aliased data

because at a fixed frequency, f , the array responses for two components that are separated in slowness by $N/(f\Delta x)$ are indistinguishable (Yilmaz and Taner, 1994). One of the ways to deal with the aliasing is to use semblance as the preconditioning of the data.

3.5.1. Preconditioning of the data

The semblance is a function which can be written as:

$$\sigma = \frac{\sum_{k=1}^N d_k}{N \sum_{k=1}^N d_k^2}, \quad (3-13)$$

where d_k is the sample of the k th trace which lies along a trajectory through the $x-t$ plane and W is a time window centered about the trajectory (Stoffa et al., 1981). The value of the semblance is independent of the amplitude of the arrival to be detected. It can have a maximum value of 1 and a minimum value of 0. Aliased energy of strong arrivals is sometimes stronger than the primary energy of weak arrivals. In this case, a weight based on the semblance can pass the amplitudes of weak arrivals provided that they are not dominated by interfering coherent arrivals (Stoffa et al., 1981). Yilmaz and Taner (1994) used semblance computed along a range of slant directions to calculate the slope of coherent arrivals at each point. The slope of the maximum semblance gives the major slope at that sample, and thus limits analysis to coherent portions of $\tau-p$ space. We will follow Yilmaz and Taner (1994), but instead of calculating a semblance along short running windows in $\tau-x$ space, we will calculate semblance $\sigma(\tau,p)$ in the $\tau-p$ space over short windows that are defined by our velocity structure. We will design weights based

on semblance for each τ - p pair to generate a masking matrix \mathbf{M} which is applied in the model space. Each semblance value is calculated over a running time window.

For the proper reverse τ - p transform, we have calculated a *rho* filter. The *rho* filter for frequency ω is (Claerbout, 1985):

$$\tilde{\rho}(\omega) = \frac{|\omega|}{2\pi}, \quad (3-14)$$

where $\tilde{\rho}(\omega)$ is the Fourier transform of $\rho(\tau)$. The rho filter is applied in the frequency domain to \mathbf{m} and data are transferred back to the time domain.

Preconditioning of the input data is also performed. We calculate frequency tapers for application in the Fourier transform. In order to minimize edge effects, we also calculate both spatial and temporal taper weights for the input data. The product of these two tapered weights gives a data weighting matrix \mathbf{W} . We also use data weighting matrix \mathbf{W} to mute the data prior to the first arrivals by setting $\mathbf{W}_j = 0, j=1, \dots, k-1$ where k is the first break sample.

With the preconditioning, the modeling equation (3-10) becomes:

$$\mathbf{d} = \mathbf{LMm}. \quad (3-15)$$

To do inversion, we define a cost or objective function, which is a mathematical expression that measures the degree of data misfit. The optimization problem is:

$$\text{minimize } \|\mathbf{Mm}\|^l \text{ subject to } \|\mathbf{W}(\mathbf{d} - \mathbf{LMm})\|^m = r, \quad (3-16)$$

where r is a weighted residual or difference between the original data and data obtained from the proposed model, while l and m indicate that different norms can be applied to

measure the norm of vectors (Trad et al., 2003). Minimization of the objective function given by equation (3-16), with $l=2$ and $m=2$, produces the following system of equations:

$$\mathbf{M}^T \mathbf{L}^T \mathbf{W}^T \mathbf{W} \mathbf{d} = \mathbf{L}^T \mathbf{W}^T \mathbf{W} \mathbf{L} \mathbf{M} \mathbf{m} . \quad (3-17)$$

This system of equation can be solved using the conjugate gradient method described in Appendix A1.

3.6. P and SV wave separation of vector VSP data by conjugate gradient method

For a vector VSP data with recorded horizontal and vertical components $\mathbf{d}_x(t,x)$ and $\mathbf{d}_z(t,x)$, we generalize the forward modeling problem given by equation (3-10) to obtain:

$$\begin{pmatrix} \mathbf{d}_x \\ \mathbf{d}_z \end{pmatrix} = \begin{pmatrix} \mathbf{L}_{xp} & \mathbf{L}_{xs} \\ \mathbf{L}_{zp} & \mathbf{L}_{zs} \end{pmatrix} \begin{pmatrix} \mathbf{m}_p \\ \mathbf{m}_s \end{pmatrix}, \quad (3-18)$$

where \mathbf{L} is a vector transformation-polarization matrix and $\mathbf{m}_p(\tau,p)$ and $\mathbf{m}_s(\tau,p)$ are amplitudes of P and S plane waves. The vector transformation-polarization matrix, \mathbf{L} , consists of Radon transform matrix, \mathbf{T} , and polarization matrix, \mathbf{A} :

$$\mathbf{L} = \mathbf{T} \mathbf{A} \quad (3-19)$$

We write a Radon transform matrix \mathbf{T} as

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_p & 0 \\ 0 & \mathbf{T}_s \end{pmatrix}, \quad (3-20)$$

and we form a polarization matrix \mathbf{A} using the unit vectors $\hat{\mathbf{a}}_p$ and $\hat{\mathbf{a}}_s$ (3-6)

$$\mathbf{A} = \begin{pmatrix} \sin \psi_p & \cos \psi_s \\ \cos \psi_p & -\sin \psi_s \end{pmatrix}, \quad (3-21)$$

to obtain the transformation-polarization matrix \mathbf{L}

$$\mathbf{L} = \begin{pmatrix} \mathbf{T}_p & 0 \\ 0 & \mathbf{T}_s \end{pmatrix} \begin{pmatrix} \sin \psi_p & \cos \psi_s \\ \cos \psi_p & -\sin \psi_s \end{pmatrix}. \quad (3-22)$$

With the conjugate gradient method, we will predict the initial model \mathbf{m}_0 and then iterate until the difference between \mathbf{d}_{obs} and \mathbf{d} is acceptable. Our numerical reconstruction at the n^{th} iteration is a sum of a series of correction vectors $\Delta \mathbf{d}_l$:

$$\begin{bmatrix} \mathbf{d}_x \\ \mathbf{d}_z \end{bmatrix}_n = \sum_{l=0}^n \begin{bmatrix} \Delta \mathbf{d}_{xl} \\ \Delta \mathbf{d}_{zl} \end{bmatrix}. \quad (3-23)$$

The data misfit vector or residual, \mathbf{r}_n is defined as $\mathbf{r}_n = \mathbf{W}(\mathbf{d}_{\text{obs}} - \mathbf{d}_n)$ or $\mathbf{r}_n = \mathbf{W}(\mathbf{r}_{n-1} - \Delta \mathbf{d}_n)$. The data correction vector $\Delta \mathbf{d}_n$ is the difference between two calculated values for the data vector or the difference between two residuals $\Delta \mathbf{d}_n = \mathbf{d}_n - \mathbf{d}_{n-1} = \mathbf{r}_{n-1} - \mathbf{r}_n$. We begin with iteration $n=0$, with the model and data correction vectors,

$$\begin{bmatrix} \mathbf{m}_p \\ \mathbf{m}_s \end{bmatrix} = 0, \quad (3-24)$$

$$\begin{bmatrix} \Delta \mathbf{d}_{x0} \\ \Delta \mathbf{d}_{z0} \end{bmatrix} = 0, \quad (3-25)$$

the augmented data misfit vector,

$$\begin{bmatrix} \mathbf{r}_{x0} \\ \mathbf{r}_{z0} \end{bmatrix} = \mathbf{W} \begin{bmatrix} \mathbf{d}_{x\text{obs}} \\ \mathbf{d}_{z\text{obs}} \end{bmatrix} \quad (3-26)$$

and the gradient vector \mathbf{g}_0 (A1-21)

$$\begin{bmatrix} \mathbf{g}_{p0} \\ \mathbf{g}_{s0} \end{bmatrix} = \mathbf{M}_0^T \mathbf{R} \mathbf{L}^T \mathbf{W}^T \mathbf{W} \begin{bmatrix} \mathbf{r}_{x0} \\ \mathbf{r}_{z0} \end{bmatrix}, \quad (3-27)$$

where \mathbf{R} is the *rho* filter matrix. Equation (3-24) is performed in the forward τ - p transform.

We also initialize the conjugate gradient vector \mathbf{h}_0 before we start the first iteration

$$\begin{bmatrix} \mathbf{h}_{p0} \\ \mathbf{h}_{s0} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{p0} \\ \mathbf{g}_{s0} \end{bmatrix}. \quad (3-28)$$

In general, a forward τ - p transform is an inversion of the equation (3-10), so we can write

$$\mathbf{m} = \mathbf{L}^{-1} \mathbf{d} = \mathbf{A}^{-1} \mathbf{T}^{-1} \mathbf{d}. \quad (3-29)$$

In order to avoid calculation of the inverse of the transformation matrix, \mathbf{T}^{-1} , we will approximate the inverse with the transpose matrix \mathbf{T}^T . We will not approximate the inverse of the polarization matrix, \mathbf{A}^{-1} , because it is easy to calculate this matrix analytically, and thus use the correct value in the transform:

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{\sin \psi_p}{\cos(\psi_p - \psi_s)} & \frac{\cos \psi_s}{\cos(\psi_p - \psi_s)} \\ \frac{\cos \psi_p}{\cos(\psi_p - \psi_s)} & \frac{-\sin \psi_s}{\cos(\psi_p - \psi_s)} \end{pmatrix}. \quad (3-30)$$

For a range of p values, we calculate incident angles of P-waves and S-waves using equation (3-4). The requirement is that we know the velocities of both P and S waves, \mathbf{v}_p and \mathbf{v}_s within the borehole.

In order to assure the correct forward and inverse Radon transforms, we precalculate each value of the data space $\mathbf{d}(t,x)$ and the model space $\mathbf{m}(\tau,p)$ from the values of the five surrounding points using the Lagrange interpolation polynomial.

The input data $\mathbf{d}_x(t,x)$ and $\mathbf{d}_z(t,x)$ can be low pass filtered to generate unaliased semblances $\sigma_p(\tau,p)$ and $\sigma_s(\tau,p)$, which are calculated during the τ - p transform. The model

preconditioning matrix \mathbf{M} can have a minimum value of 0 and a maximum value of 1. We calculate matrix \mathbf{M}_n with the each iteration of our conjugate gradient solver based on the semblance of the residual vector \mathbf{r}

$$\mathbf{M}_n = \begin{bmatrix} \mathbf{M}_{pn} \\ \mathbf{M}_{sn} \end{bmatrix} = \begin{bmatrix} \zeta_{pn} \boldsymbol{\mu}_{pn} \\ \zeta_{sn} \boldsymbol{\mu}_{sn} \end{bmatrix}. \quad (3-31)$$

Functions ζ_p and ζ_s are calculated from the semblances local maxima, either $\sigma_{p\max}(\tau, p)$ or $\sigma_{s\max}(\tau, p)$, to taper the masking matrices $\boldsymbol{\mu}_p$ and $\boldsymbol{\mu}_s$. Masking matrices $\boldsymbol{\mu}_p$ and $\boldsymbol{\mu}_s$ originally have value 1 or 0. In the case when we have intersection of both P and S coherent arrivals, we will use masking matrices $\boldsymbol{\mu}_p$ and $\boldsymbol{\mu}_s$ to prevent leakage of the two coherent slant stacks with different polarizations. We will favor the slant stack with a higher energy (numerator in the equation 3-13).

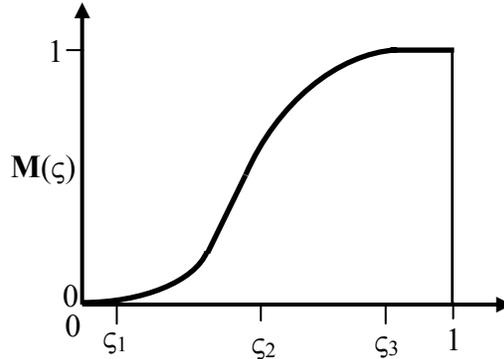


FIG. 3.6. The model preconditioning matrix \mathbf{M} .

The model preconditioning matrix \mathbf{M} (Figure 3.6) is then recalculated from the semblances $\sigma_p(\tau, p)$ and $\sigma_s(\tau, p)$ of the updated data misfit vector \mathbf{r} with each iteration, n , of the conjugate gradient solver. With the user specified relaxation value ϕ , we calculate semblance based taper corners ξ_1 and ξ_2 such that we compare maximum P and S

semblance values to allow modeling of the most coherent events at the beginning and less coherent events with further iterations:

$$\zeta_1 = \phi \zeta_2 \quad (3-32)$$

$$\zeta_2 = \sigma(\tau, p) \phi^n . \quad (3-33)$$

After initialization, we iterate for values $n = (0, 1, 2, \dots, N)$. With each iteration, the conjugate gradient vector \mathbf{h}_n will be weighted with \mathbf{M}_n and then reverse transformed from the τ - p domain to the t - x space to obtain the data correction vector $\Delta \mathbf{d}_n$ (A1-22)

$$\begin{bmatrix} \Delta \mathbf{d}_{xn} \\ \Delta \mathbf{d}_{zn} \end{bmatrix} = \mathbf{W} \mathbf{L} \begin{bmatrix} \mathbf{M}_p \mathbf{h}_{pn} \\ \mathbf{M}_s \mathbf{h}_{sn} \end{bmatrix}. \quad (3-34)$$

Following equation A1-23, a new model steplength α is calculated from data correction vectors and gradient vectors

$$\alpha = \frac{\mathbf{g}_p^2 + \mathbf{g}_s^2}{\Delta \mathbf{d}_x^2 + \Delta \mathbf{d}_z^2}, \quad (3-35)$$

such that we can update the model vector

$$\begin{bmatrix} \mathbf{m}_{p,n+1} \\ \mathbf{m}_{s,n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{m}_{p,n} \\ \mathbf{m}_{s,n} \end{bmatrix} + \alpha \begin{bmatrix} \mathbf{M}_{p,n} \mathbf{h}_{p,n} \\ \mathbf{M}_{s,n} \mathbf{h}_{s,n} \end{bmatrix}, \quad (3-36)$$

and data misfit vector

$$\begin{bmatrix} \mathbf{r}_{x,n+1} \\ \mathbf{r}_{z,n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{x,n} \\ \mathbf{r}_{z,n} \end{bmatrix} - \alpha_n \begin{bmatrix} \Delta \mathbf{d}_{x,n} \\ \Delta \mathbf{d}_{z,n} \end{bmatrix}. \quad (3-37)$$

In order to accelerate convergence, we will numerically search for a value of the model steplength α_n . First, we will calculate the L2 norm of equation (3-37)

$$J(\alpha_n) = \begin{bmatrix} \mathbf{r}_{xn} - \alpha \Delta \mathbf{d}_{xn} \\ \mathbf{r}_{zn} - \alpha \Delta \mathbf{d}_{zn} \end{bmatrix}^T \mathbf{W}^T \mathbf{W} \begin{bmatrix} \mathbf{r}_{xn} - \alpha \Delta \mathbf{d}_{xn} \\ \mathbf{r}_{zn} - \alpha \Delta \mathbf{d}_{zn} \end{bmatrix}. \quad (3-38)$$

The minimum value for the function $J(\alpha_n)$ is found by comparison of adjacent function values for a range of α . The value of α for which function $J(\alpha_n)$ has a minimum value, corresponds to a wanted steplength α_n .

We calculate a new gradient vector \mathbf{g}_{n+1} using the updated misfit vector \mathbf{r}_{n+1} and new preconditioning weights \mathbf{M}_{n+1} obtained from semblance calculation for each particular iteration n

$$\begin{bmatrix} \mathbf{g}_{p,n+1} \\ \mathbf{g}_{s,n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{p,n+1}^T \\ \mathbf{M}_{s,n+1}^T \end{bmatrix} \mathbf{R} \mathbf{L}^T \mathbf{W}^T \mathbf{W} \begin{bmatrix} \mathbf{r}_{x,n+1} \\ \mathbf{r}_{z,n+1} \end{bmatrix}. \quad (3-39)$$

With a new gradient vector \mathbf{g}_{n+1} , we can calculate a new conjugate direction steplength β

$$\beta_n = \frac{\|\mathbf{g}_{n+1}\|^2}{\|\mathbf{g}_n\|^2}, \quad (3-40)$$

and an updated conjugate direction vector \mathbf{h}_{n+1}

$$\begin{bmatrix} \mathbf{h}_{p,n+1} \\ \mathbf{h}_{s,n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{p,n+1} \\ \mathbf{g}_{s,n+1} \end{bmatrix} - \beta_n \begin{bmatrix} \mathbf{h}_{p,n} \\ \mathbf{h}_{s,n} \end{bmatrix} \quad (3-41)$$

Separated P and S waves $\mathbf{m}_p(\tau,p)$ and $\mathbf{m}_s(\tau,p)$ are obtained in the τ - p domain as the result of the last iteration and then transformed to the t - x domain in the inverse τ - p transform step using equation (3-18).

The criteria for stopping iteration are calculated based on the user specified convergence tolerances ξ_{\min} and defined maximum possible number of iterations N. If the ratio between the difference in the current iteration residual value and previous residual value and the starting residual value is smaller than a user specified value ξ_{\min} ,

$$\frac{\|\mathbf{r}_{n+1} - \mathbf{r}_n\|}{\|\mathbf{r}_0\|^2} < \xi_{min} , \quad (3-42)$$

the iteration process will be stopped.

The other stopping criteria depend on the number of performed iterations n and are used to stop the algorithm even if the desired convergence in residual calculation is not achieved. If the number of performed iterations n exceeds the user allowed number of iterations N , such that $n > N$, the iteration process will be terminated.

4. APPLICATIONS OF THE VECTOR VSP SEPARATION ALGORITHM

4.1. Algorithm structure and requirements

The multicomponent VSP wavefield separation code is shown in the form of an algorithm in Figure 4.1. Input to the code are multicomponent VSP shot gathers and velocity file, with both v_p and v_s values specified for each receiver position. The user must also specify a semblance relaxation criterion ϕ and the algorithm termination parameters: convergence tolerances ξ_{\min} and maximum possible number of iterations N .

Though multicomponent VSP data are usually recorded on three components, the code is currently made to work with the horizontal \mathbf{d}_x and vertical \mathbf{d}_z component. It is important to know the frequency spectrum of the data in order to specify input frequency values used in data filtering. Values for the recording length, sample rate and offsets are read in from the input gathers together with the recording geometry data.

The range of p values used in the code is user specified based on velocity information such that $p_{\min} = -1/v_{s\min}$ and $p_{\max} = 1/v_{s\max}$. The sample rate, Δp , is calculated from a user specified range of offsets, x_r , and a maximum specified frequency such that it satisfies the Nyquist criteria (3-9).

Data filtering by muting in the τ - p domain is optional (Figure 4.1). However, if one wants to mute the data, the ray parameter values for muting, p_{\min_mute} and p_{\max_mute} , are required as input parameters. In addition to that, one can specify range of incident angles, ψ_p and ψ_s , to be either muted or extracted from the rest of the wavefield.

Previously described high resolution discrete Radon transform P and S wave separation, together with all conjugate gradient calculations, is performed in the forward

step of this algorithm. We obtain partially separated modes in the τ - p domain with every iteration. The completely separated P and S waves, \mathbf{m}_p and \mathbf{m}_s , are the output of the last iteration. We then reversely transform the last \mathbf{m}_p and \mathbf{m}_s values to get separated P and S waves, \mathbf{d}_p and \mathbf{d}_s , in the t - x domain. The whole procedure is repeated for every input gather.

4.1.1. Three approaches for algorithm application based on the size of a spatial window

There are three approaches for τ - p filtering depending on the spatial window size, i.e. how many traces from one shot are input to this algorithm,:

- 1) Use a fixed, large spatial window, which includes all traces. Unfortunately, discrete waveform discontinuities at major reflectors are not well represented by only a few τ - p parameters. In addition, the semblance estimates are weak since the wavefronts have some curvature (Figure 4.2a).
- 2) Use a small trace running window, for example 11 traces. The semblance will be better, but we will still span discontinuities and have artifacts in those areas (Figure 4.2b).
- 3) Use fixed, sequential windows. Since we know velocities within the well, we can reliably estimate the major velocity breaks and form spatial windows based on this information (Figure 4.2c). This approach will cost less computation time than the previous one and it will avoid artifacts.

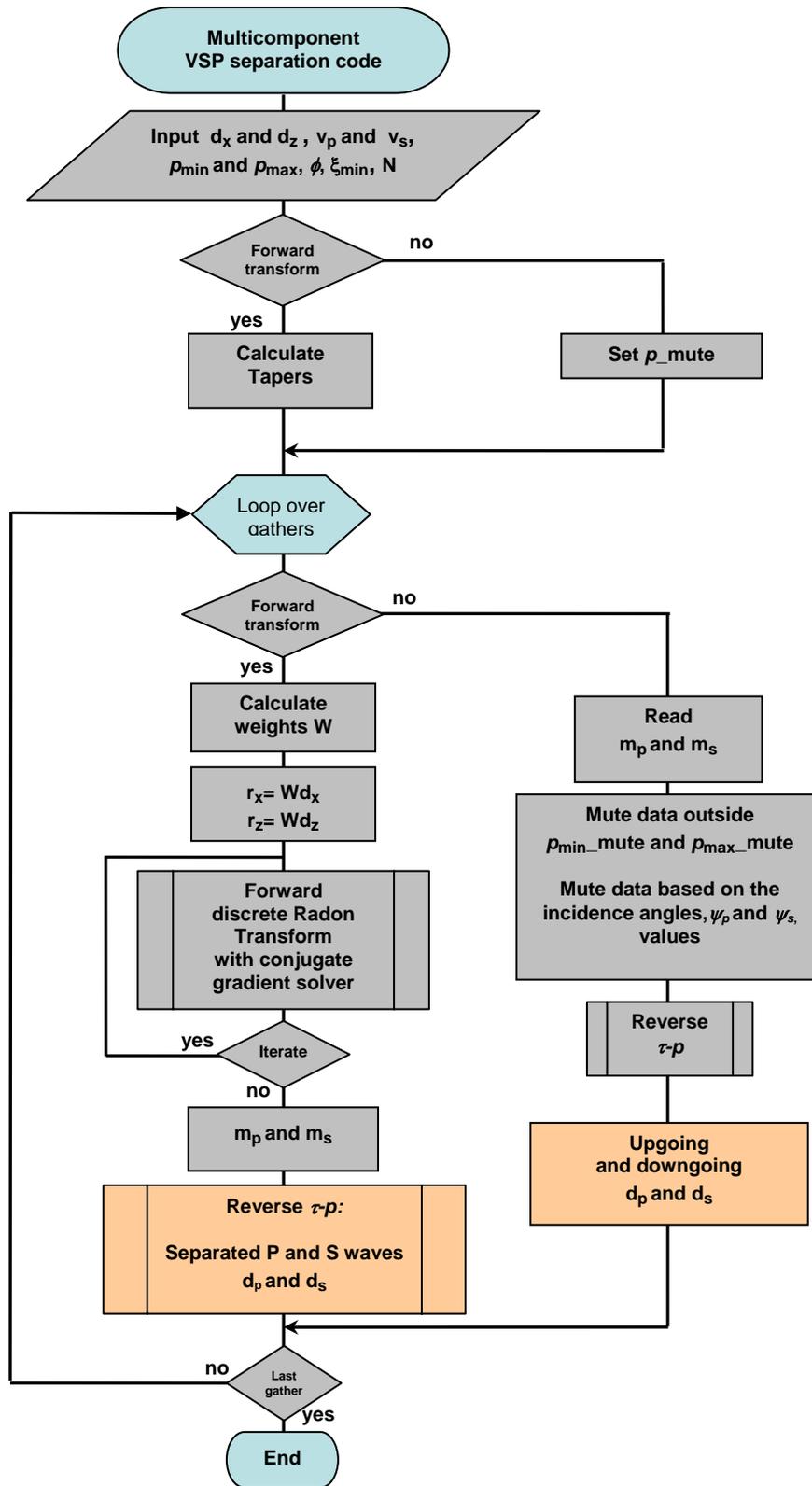


FIG. 4.1. Multicomponent VSP separation algorithm.

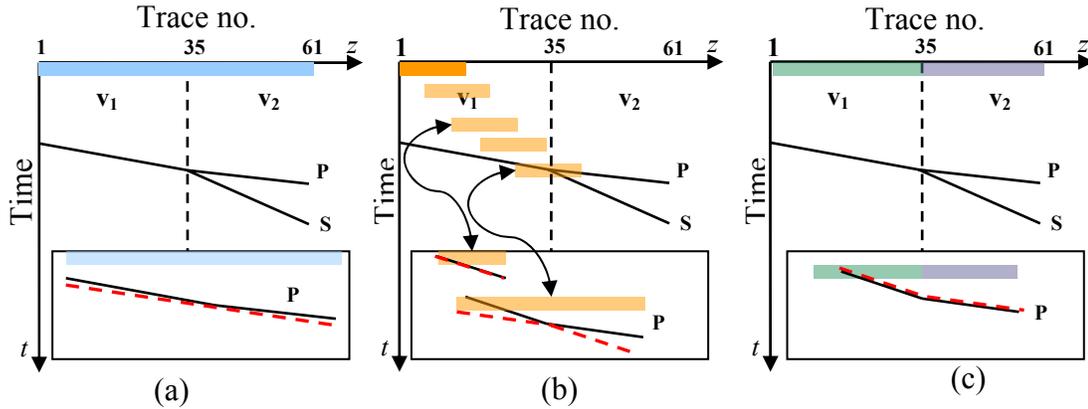


FIG. 4.2. Three approaches for τ - p filtering: (a) A fixed, large spatial window, (b) a small, running spatial window and (c) a velocity based fixed sequential window (c). Different colours on this scheme show positions and sizes of spatial windows. Reconstruction quality is approximated with dashed red lines.

4.1.2. Separation of the upgoing and downgoing waves

Beside vector VSP's P and S waves separation, this code can be used for separation of upgoing and downgoing waves. The separation procedure is achieved by simple muting of $\mathbf{m}_p(\tau, p)$ and $\mathbf{m}_s(\tau, p)$ components in τ - p domain. The upgoing VSP waves have a negative ray parameter p , while the downgoing VSP waves have a positive ray parameter p . Setting the value of the events with the negative or positive ray parameter p to zero, leaves us with the downgoing or upgoing VSP wavefield respectively. A separation of the upgoing and downgoing waves combined with the wave mode separation will give us four different wavefields: a downgoing P, a downgoing S, an upgoing P and an upgoing S wavefield. These wavefields can be used for further VSP processing.

4.2. Algorithm testing on synthetic data

In order to evaluate this method for the vector VSP separation, we tested our algorithm on the Vinton Dome synthetic data set and compared the results to the classical separation results from Chapter 2. Prior to this testing, I performed basic calibration on a simple model.

4.2.1. A simple model test

I generalized a two-layer model using GXII for basic algorithm testing (Figure 4.3). Receiver positions are the same as in the Vinton Dome model: the upper geophone is at 287 m and deepest is at 1202 m. The boundary between the two layers is at 610 m. I recorded two zero offset shots. The first shot was recorded with horizontal geophones, thus recording S waves only. The second shot was recorded with vertical geophones and it recorded P waves only. Having the waves already separated, I know the answer in advance and I easily test my algorithm.

In the first test, I included 22 receivers to record downgoing waves within the first layer only (Figure 4.4). Data are lowpass filtered to limit the frequency range such that the highest frequency is 5 Hz. Since I have only one velocity, I applied a fixed, large spatial window that includes all 22 traces. Maximum and minimum ray parameter values, $p_{max} = 0.0016$ s/m and $p_{min} = -0.0016$ s/m, are higher than $1/v_{smin}$ and lower than $-1/v_{smin}$ to avoid truncation effects. I set my relaxation parameter to $\phi = 0.7$ and the maximum number of iterations to 10.

Zero iteration residuals, \mathbf{r}_{x0} and \mathbf{r}_{z0} (Figure 4.5), are produced from the input data (3-26) by application of a weighting matrix \mathbf{W} . First, gradients calculated from these residuals, \mathbf{g}_{p0} and \mathbf{g}_{s0} (Figure 4.6), are used to obtain semblances \mathbf{s}_{p0} and \mathbf{s}_{s0} (Figure 4.7),

and model preconditioning matrices \mathbf{M}_{p0} and \mathbf{M}_{s0} (Figure 4.8). Conjugate gradients \mathbf{h}_{p0} and \mathbf{h}_{s0} (Figure 4.9) are the gradients \mathbf{g}_{p0} and \mathbf{g}_{s0} (Figure 4.6) with the preconditioning matrices \mathbf{M}_{p0} and \mathbf{M}_{s0} (Figure 4.8) applied.

Once the initial ($n = 0$) vectors are calculated, the iteration process starts. At each iteration, updated residuals, \mathbf{r}_{xn} and \mathbf{r}_{zn} , (Figure 4.10), show the difference between the original data and our current data estimation. The better we model the data, the smaller the residuals (Figure 4.10b and Figure 4.10d).

I update model vectors \mathbf{m}_{pn} and \mathbf{m}_{sn} (Figure 4.11) to obtain data reconstruction vectors \mathbf{d}_{xn} and \mathbf{d}_{zn} at each iteration step (Figure 4.12). If I am satisfied with the current reconstruction, i.e. if the residuals satisfy our convergence test, I will terminate the iteration process at that step. The separated P-waves and S-waves will be the most recently updated model vectors \mathbf{m}_{pn} and \mathbf{m}_{sn} (Figure 4.11 b and Figure 4.11d).

New gradients, \mathbf{g}_{pn} and \mathbf{g}_{sn} , significantly diminish after one iteration (Figure 4.13), since I have only one linear event that is easy to model. I calculate new semblances \mathbf{s}_{pn} and \mathbf{s}_{sn} (Figure 4.14) and model preconditioning matrices \mathbf{M}_{pn} and \mathbf{M}_{sn} (Figure 4.15). Preconditioning matrices allow modeling of the coherent arrivals at the beginning (Figure 4.15a) and all other arrivals at the end of the iteration process (Figure 4.15b). At the end of each iteration step, I update conjugate gradients \mathbf{h}_{pn} and \mathbf{h}_{sn} (Figure 4.16) and continue the iteration process from the beginning. At the end of the iteration process, we transform separated P-waves, \mathbf{m}_{p10} , and S-waves, \mathbf{m}_{s10} , from the τ - p domain to the t - x domain to obtain \mathbf{d}_p and \mathbf{d}_s wavefields (Figure 4.17).

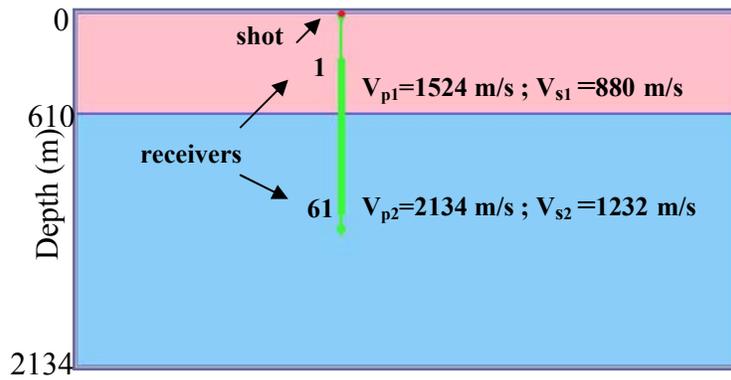


FIG. 4.3. A simple, two-layer model.

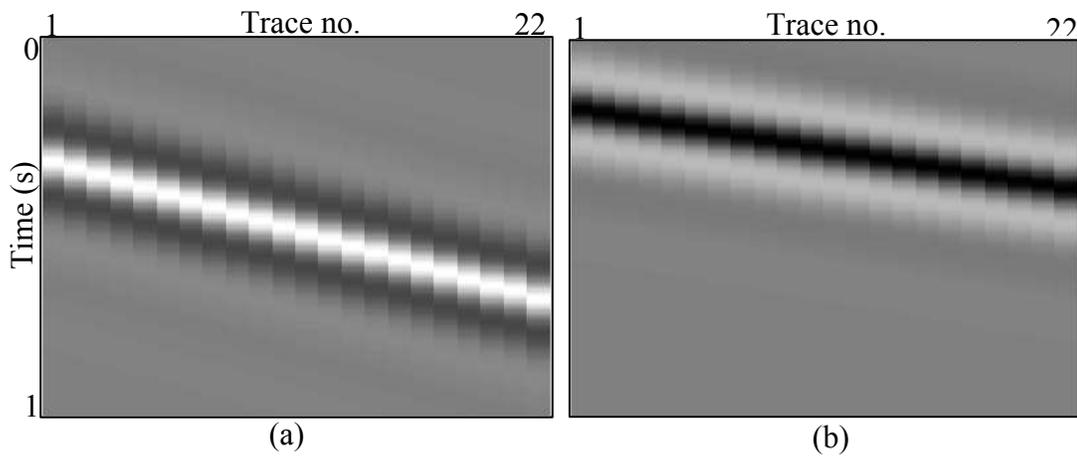


FIG. 4.4. Synthetic data from a simple one-layer model. Data are recorded within the first layer using 22 receivers. Data highest frequency is 5 Hz and recording length is 1 s. The horizontal component $\mathbf{d}_{x_{obs}}$ (a), records an S downgoing wave, while the vertical component $\mathbf{d}_{z_{obs}}$ (b), records a P downgoing wave.

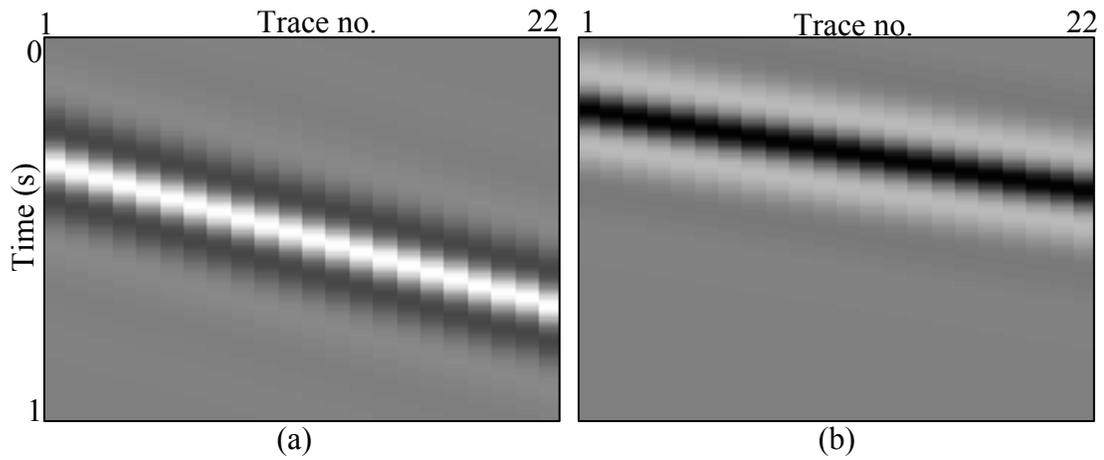


FIG. 4.5. A simple, one-layer model synthetic data x-component residual \mathbf{r}_x (a) and z-component residual \mathbf{r}_z (b) at iteration $n = 0$. Residuals are weighted input data (Figure 4.4).

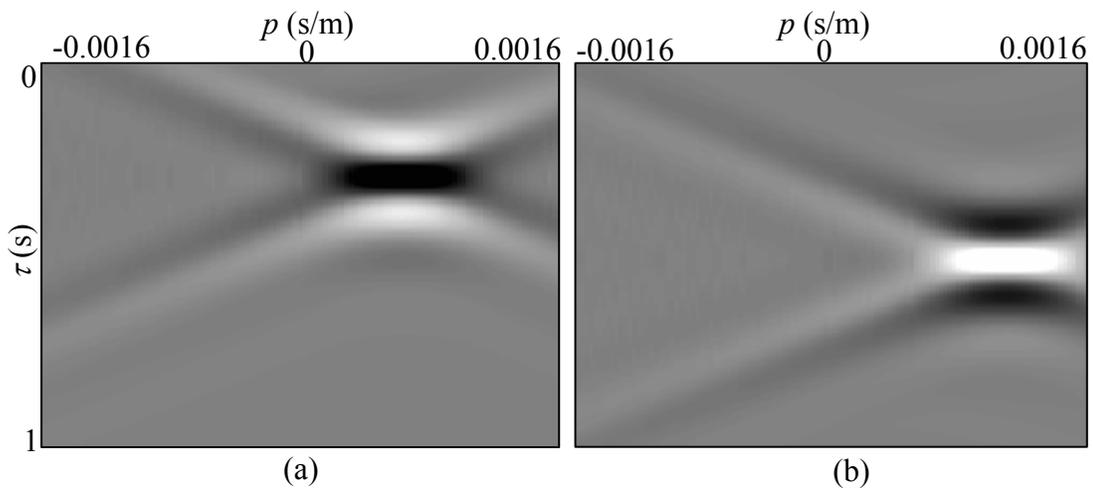


FIG. 4.6. A simple, one-layer model synthetic data P-wave gradient \mathbf{g}_p (a) and S-wave gradient \mathbf{g}_s (b) at iteration $n = 0$.

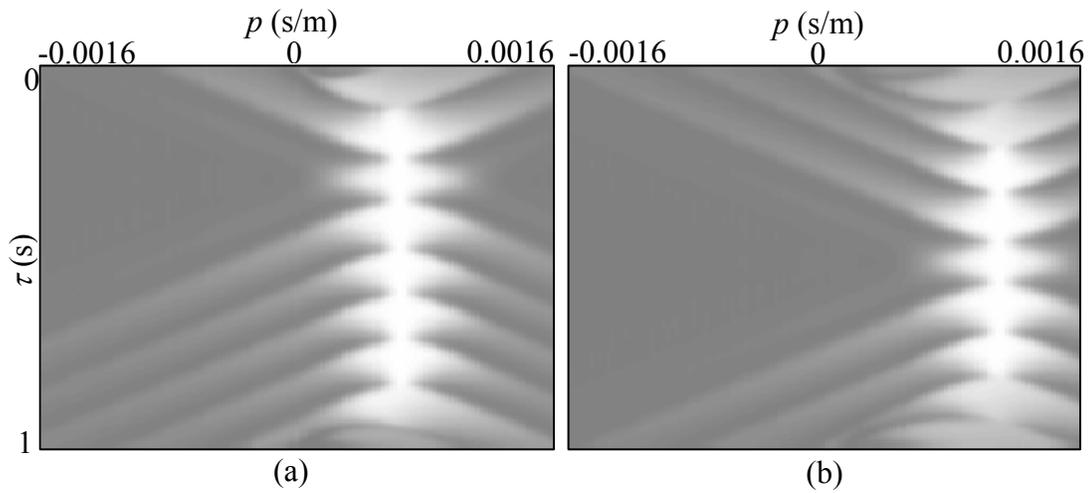


FIG. 4.7. A simple, one-layer model synthetic data P-wave semblance s_p (a) and S-wave semblance s_s (b) calculated from gradients (Figure 4.6) at iteration $n = 0$.

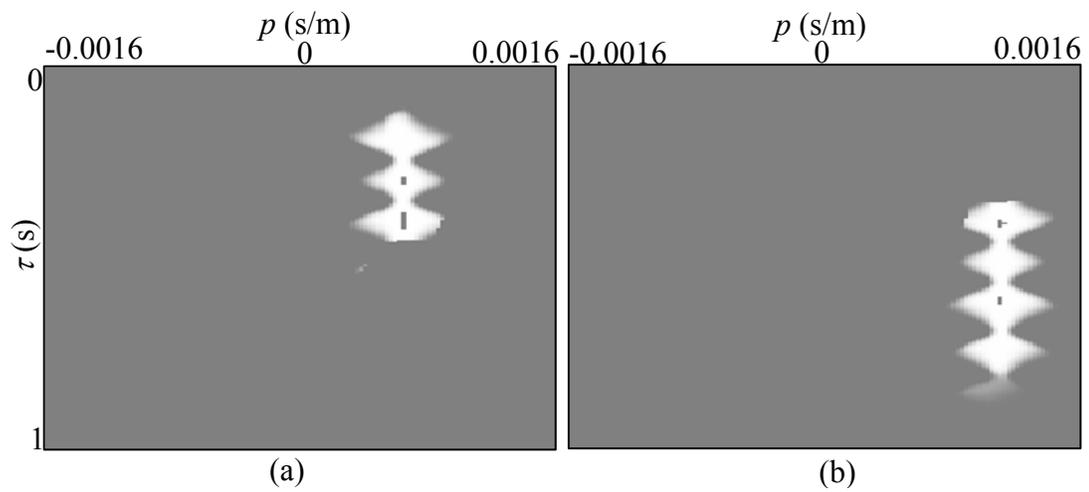


FIG. 4.8. A simple, one-layer model synthetic data P-wave model preconditioning matrix \mathbf{M}_p (a) and S-wave model preconditioning matrix \mathbf{M}_s (b) at iteration $n = 0$. Model preconditioning matrices are calculated from semblance values (Figure 4.7).

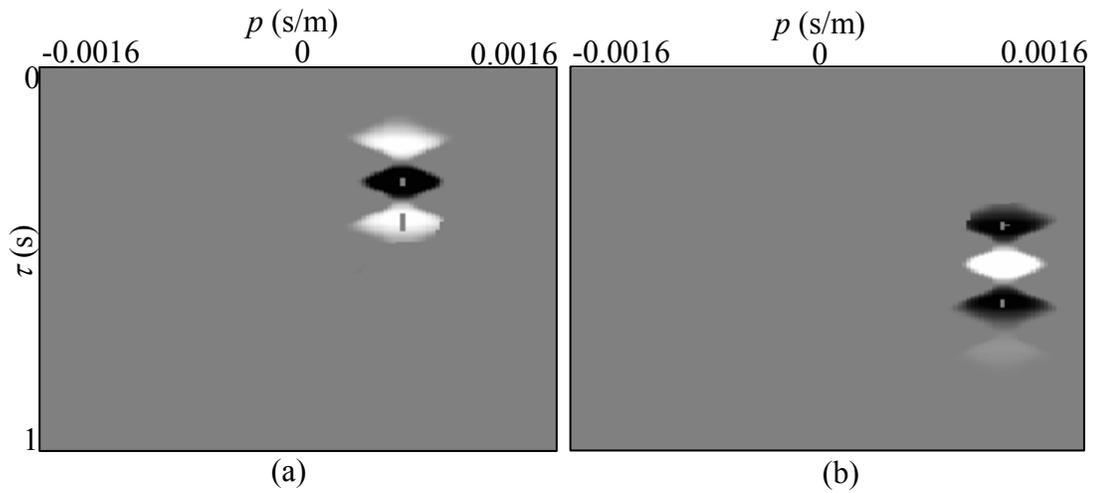


FIG. 4.9. A simple, one-layer model synthetic data P-wave conjugate gradient \mathbf{h}_p (a) and S-wave conjugate gradient \mathbf{h}_s (b) at iteration $n = 0$. Conjugate gradient is the zero iteration gradient (Figure 4.6) with the model preconditioning matrix (Figure 4.7) applied.

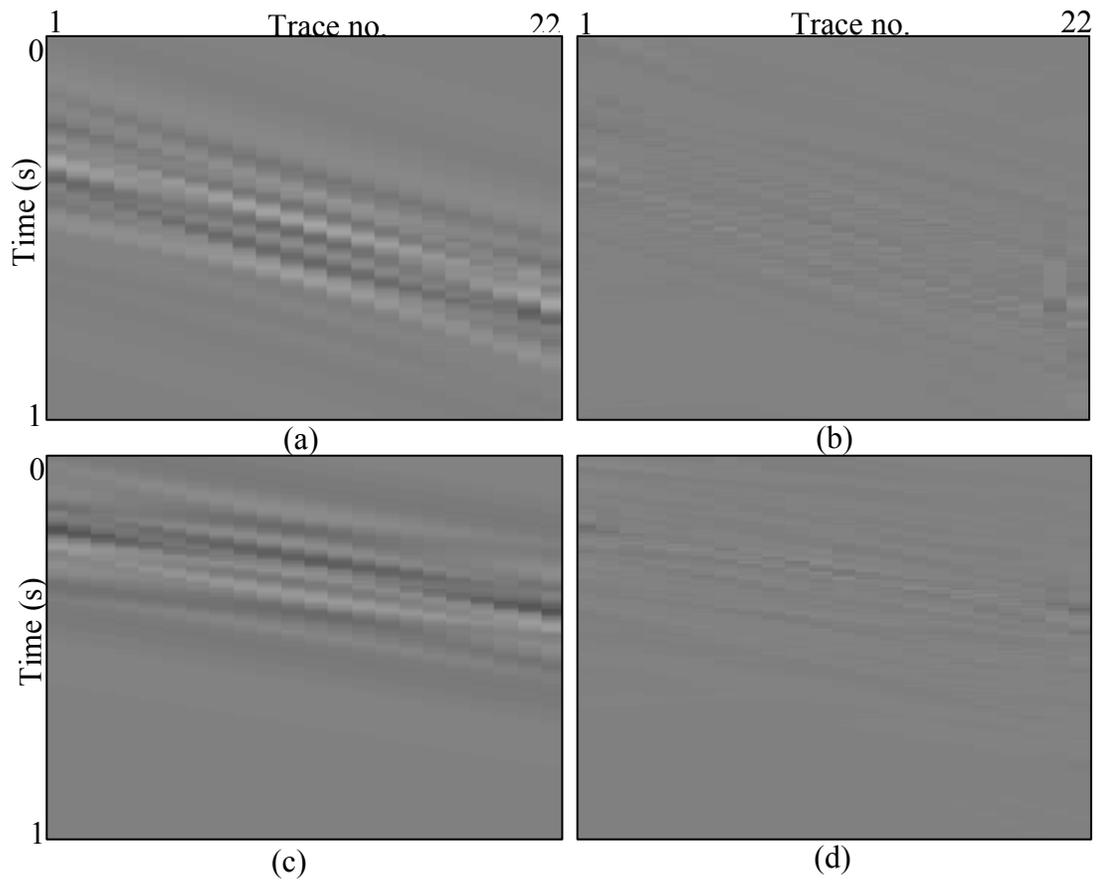


FIG. 4.10. A simple, one-layer model synthetic data x-component residual vector \mathbf{r}_x at iteration $n = 1$ (a) and iteration $n = 10$ (b); z-component residual vector \mathbf{r}_z at iteration $n = 1$ (c) and $n = 10$ (d).

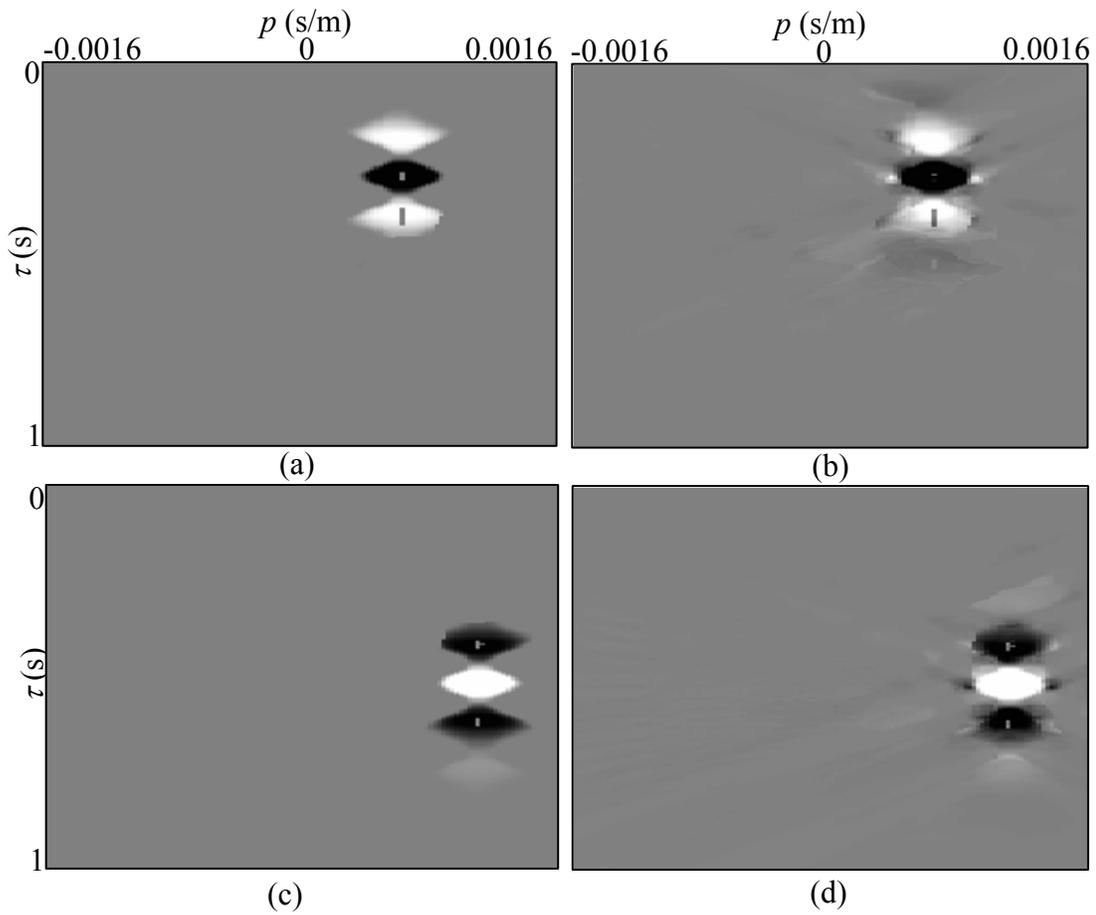


FIG. 4.11. A simple, one-layer model synthetic data P-wave model vector \mathbf{m}_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S-wave model vector \mathbf{m}_s at iteration $n = 1$ (c) and at iteration $n = 10$ (d). Model vectors at the last iteration $n = 10$ are separated P-waves and S-waves in τ - p space.

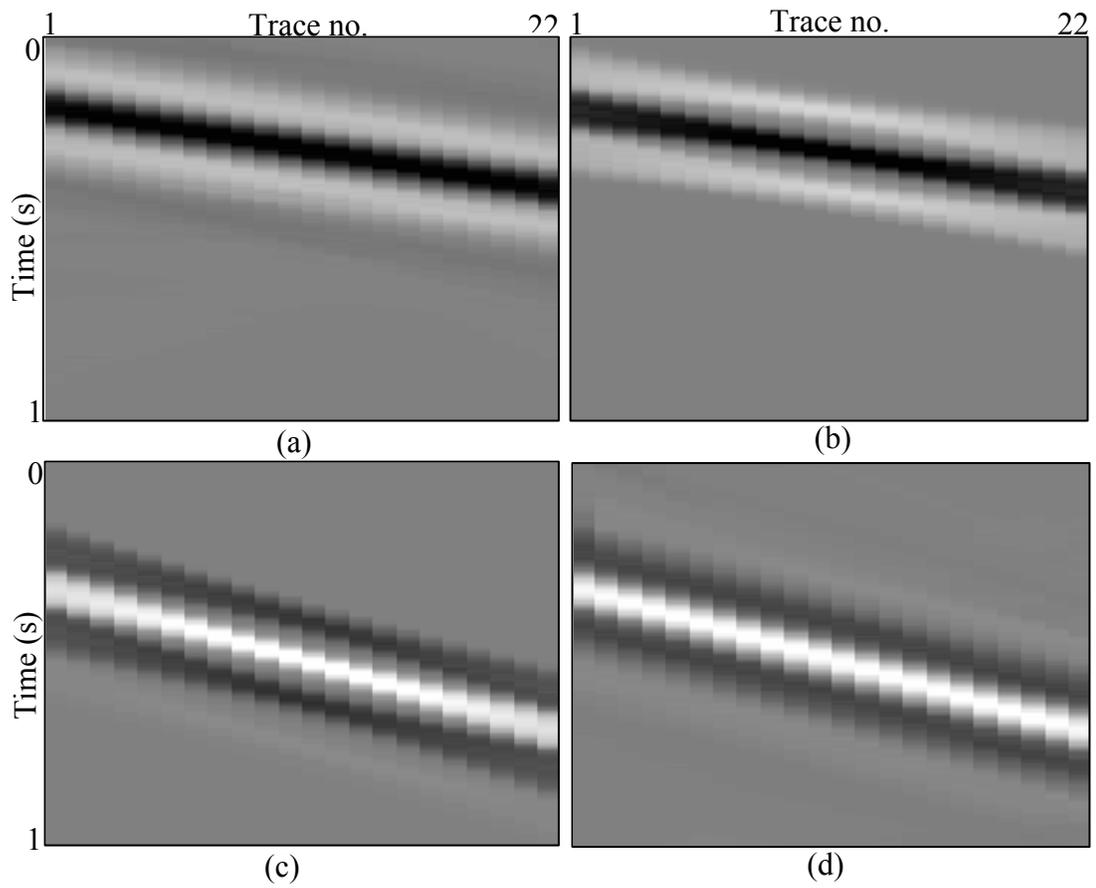


FIG. 4.12. A simple, one-layer model synthetic data x-component reconstruction vector $\mathbf{d}_{x,n}$ at iteration $n = 1$ (a) and iteration $n = 10$ (b); z-component reconstruction vector \mathbf{d}_z at iteration $n = 1$ (c) and iteration $n = 10$ (d).

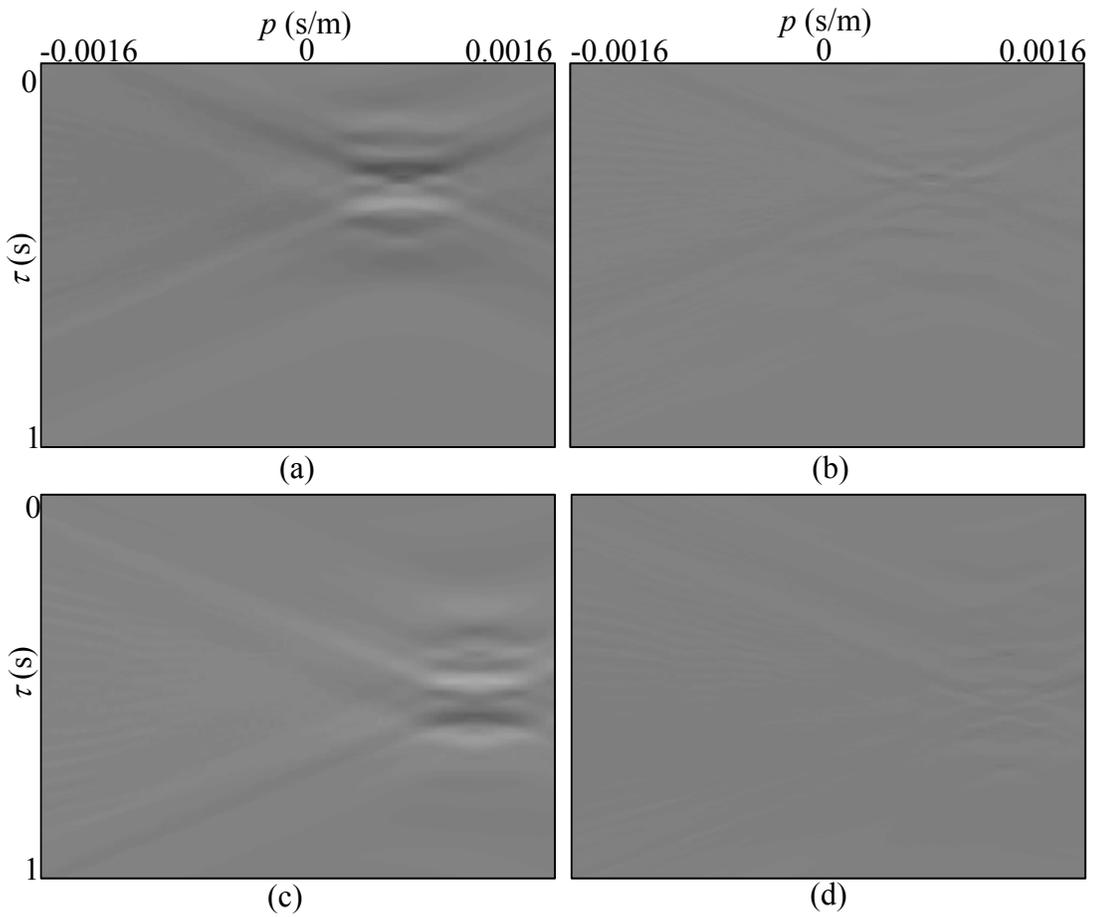


FIG. 4.13. A simple, one-layer model synthetic data P-wave gradient vector \mathbf{g}_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S-wave gradient vector \mathbf{g}_s at iteration $n = 1$ (c) and at iteration $n = 10$ (d). Comparing to Figure 4.6, gradient becomes significantly smaller after only one iteration.

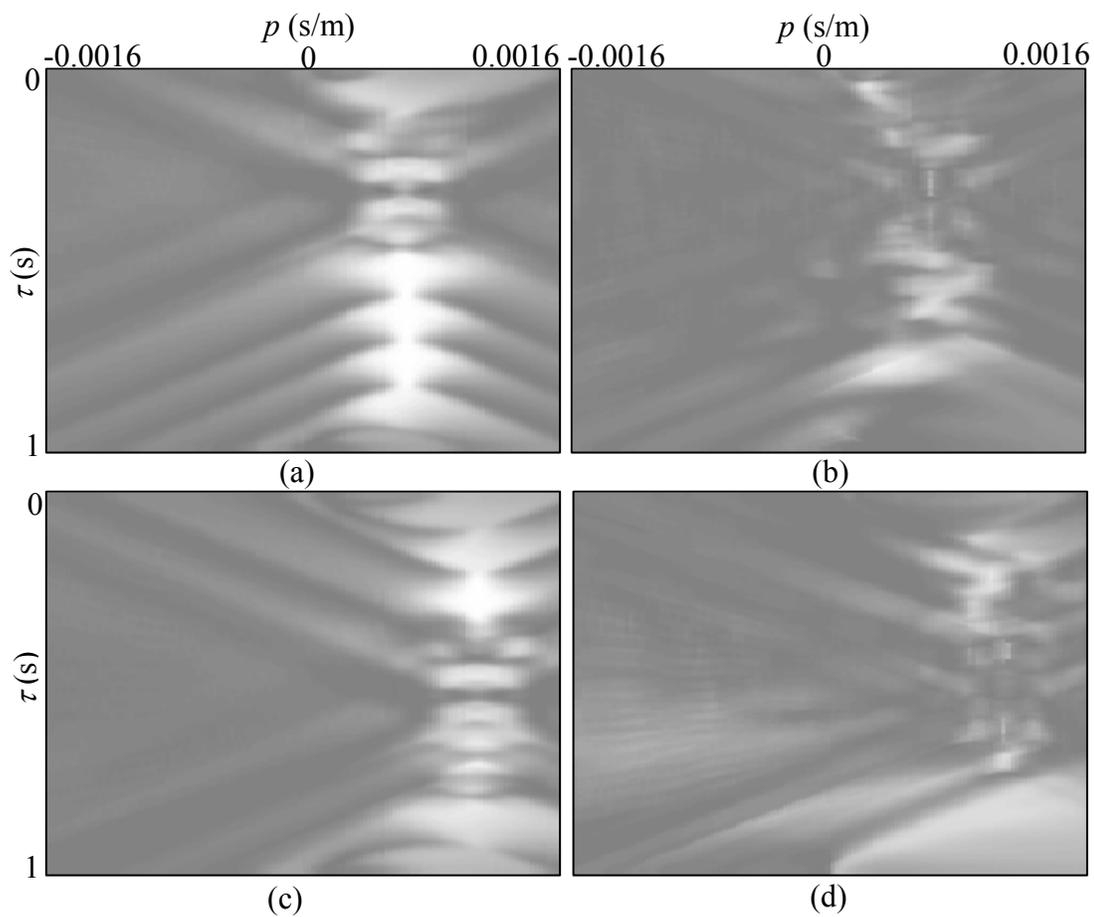


FIG. 4.14. A simple, one-layer model synthetic data P-wave semblance s_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S-wave semblance s_s at iteration $n = 1$ (c) and at iteration $n = 10$ (d).

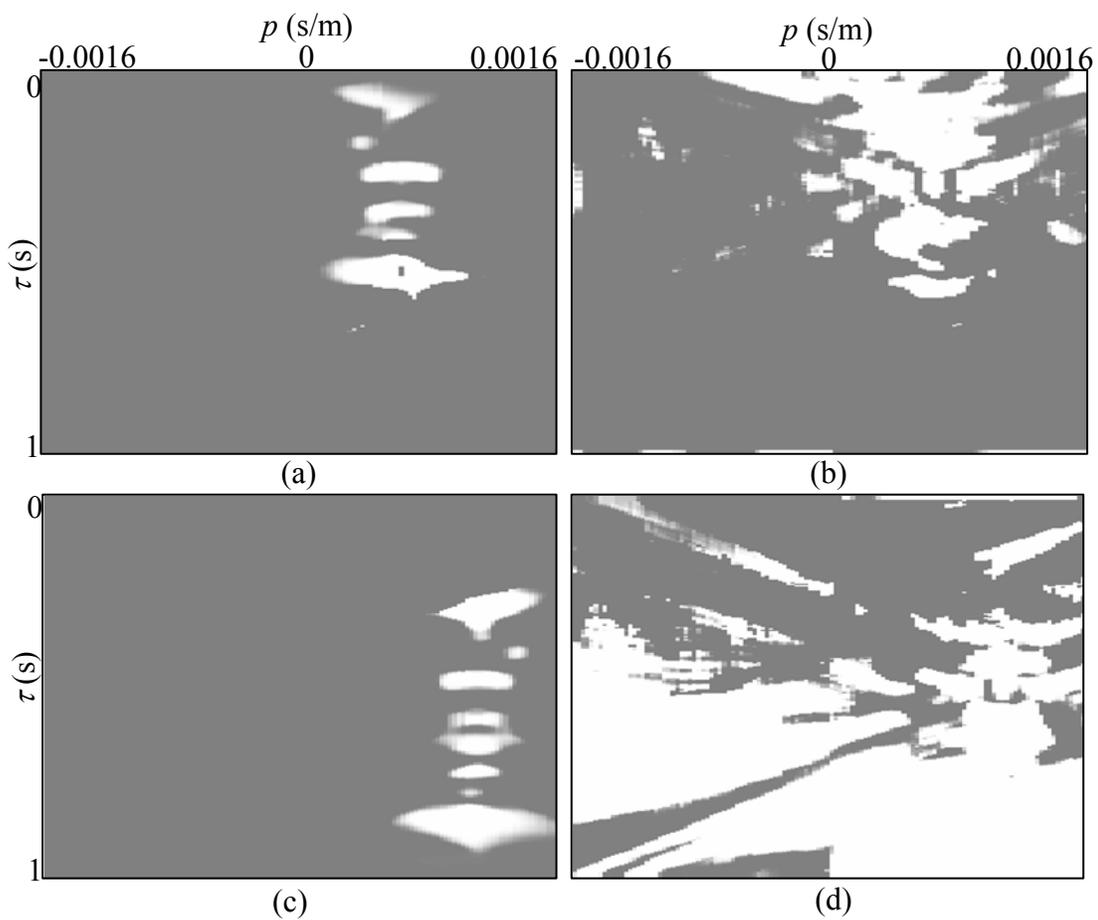


FIG. 4.15. A simple, one-layer model synthetic data P-wave model preconditioning matrix \mathbf{M}_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S wave model preconditioning matrix \mathbf{M}_s at iteration $n = 1$ (b) and at iteration $n = 10$.

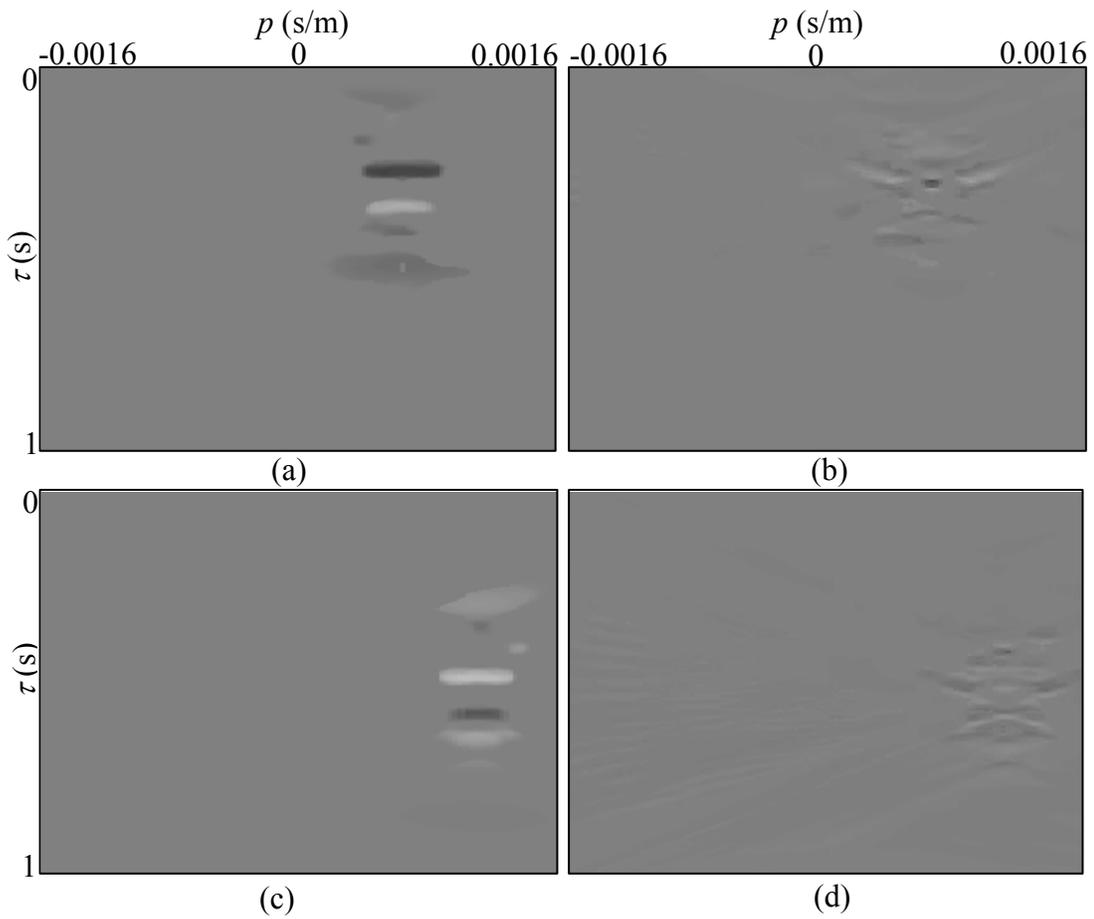


FIG. 4.16. A simple, one-layer model synthetic data P-wave conjugate gradient \mathbf{h}_p at iterations $n = 1$ (a) and $n = 10$ (b); an S-wave conjugate gradient \mathbf{h}_s at iteration $n = 1$ (c) and $n = 10$ (d).

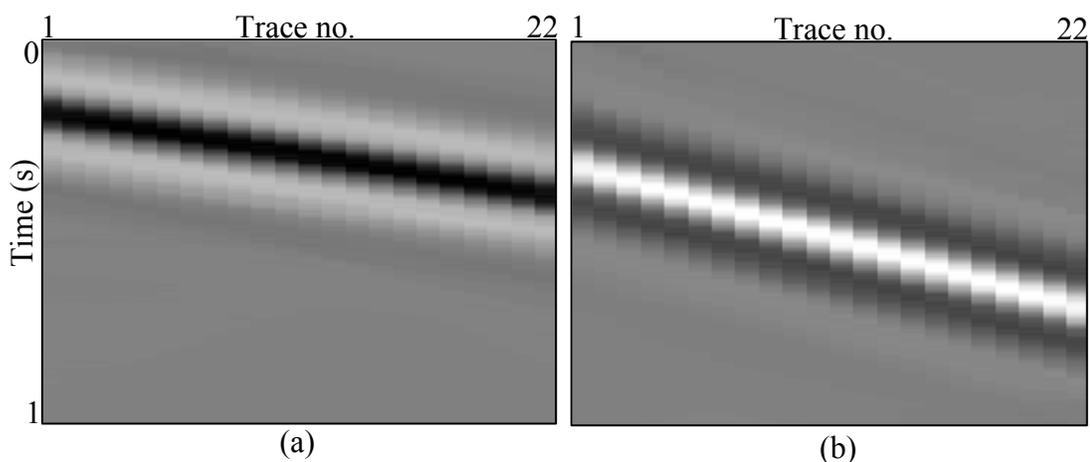


FIG. 4.17. A simple, one-layer model separated P-waves \mathbf{d}_p (a) and S-waves \mathbf{d}_s (b).

4.2.1.1. A simple test model with added random noise

In order to test our algorithm on data more similar to real data, I added some random noise to our simple model synthetic data from Figure 4.4. Synthetic data with added noise (Figure 4.18) will have more values different from zero. This will help in the semblance calculation. Comparing to Figure 4.7, semblance is now more focused (Figure 4.19), resulting in a better model preconditioning matrix \mathbf{M} (Figure 4.20) and better conjugate gradient \mathbf{h} (Figure 4.21).

In the following iteration step, semblances \mathbf{s}_p (Figure 4.22a) and \mathbf{s}_s (Figure 4.22c) are much smaller than the corresponding semblances in Figure 4.14 because we have accurately modeled the data within the first iteration. Semblances in the last iteration step, \mathbf{s}_p (Figure 4.22b) and \mathbf{s}_s (Figure 4.22d), correspond to the noise which cannot be modeled. The model preconditioning matrix \mathbf{M} (Figure 4.23) corresponds to the

calculated semblances. Separated P-waves and S-waves (Figure 4.25) are obtained from the last updated model at iteration $n=10$ (Figure 4.24).

Separation results in this simple example with the noise are not much different from the results obtained in the example without the noise (Figure 4.17), but the importance of adding the noise to the synthetic data is obvious in the case of more complicated models.

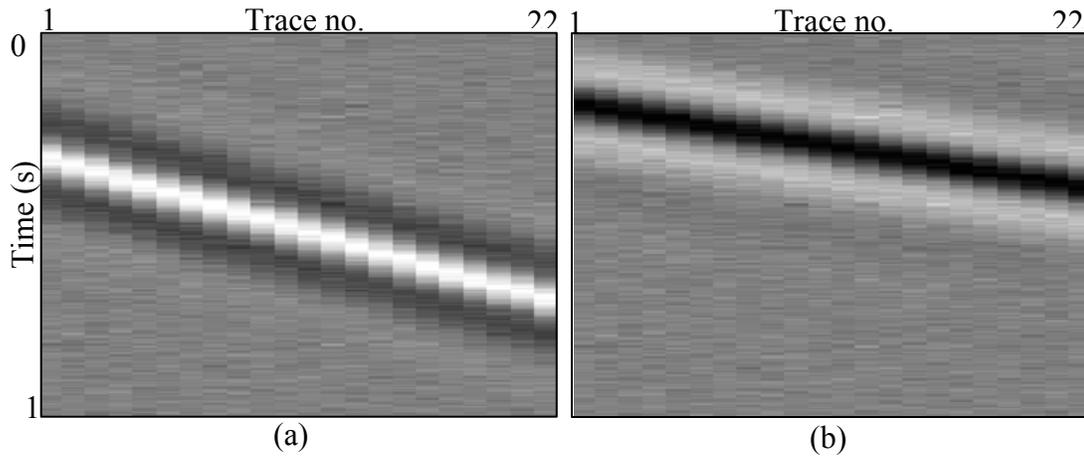


FIG. 4.18. Horizontal component $\mathbf{d}_{x_{obs}}$ (a) and vertical component $\mathbf{d}_{z_{obs}}$ (b) from Figure 4.4 with added random noise.

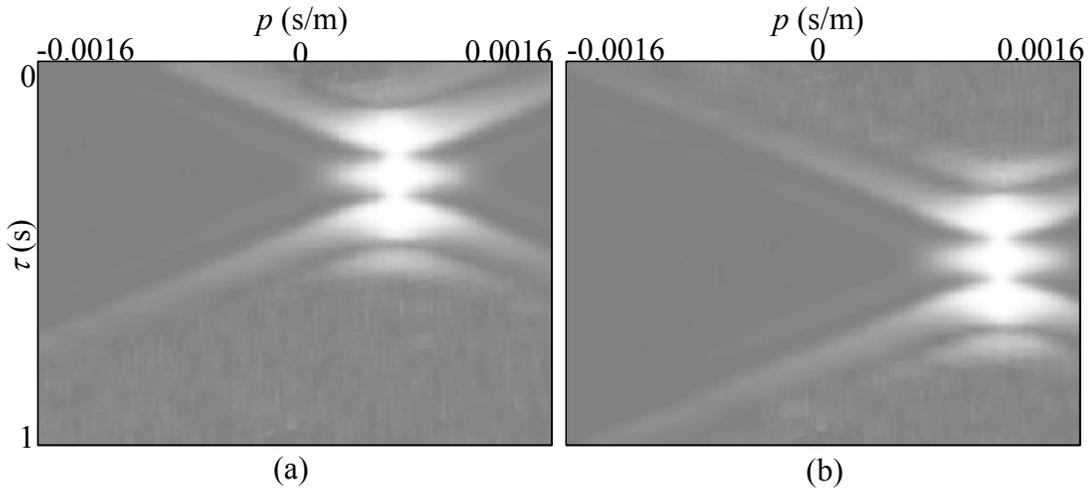


FIG. 4.19. With some random noise added, P-wave semblances s_p (a) and S-wave semblances s_s (b) are more focused than the semblances from Figure 4.7.

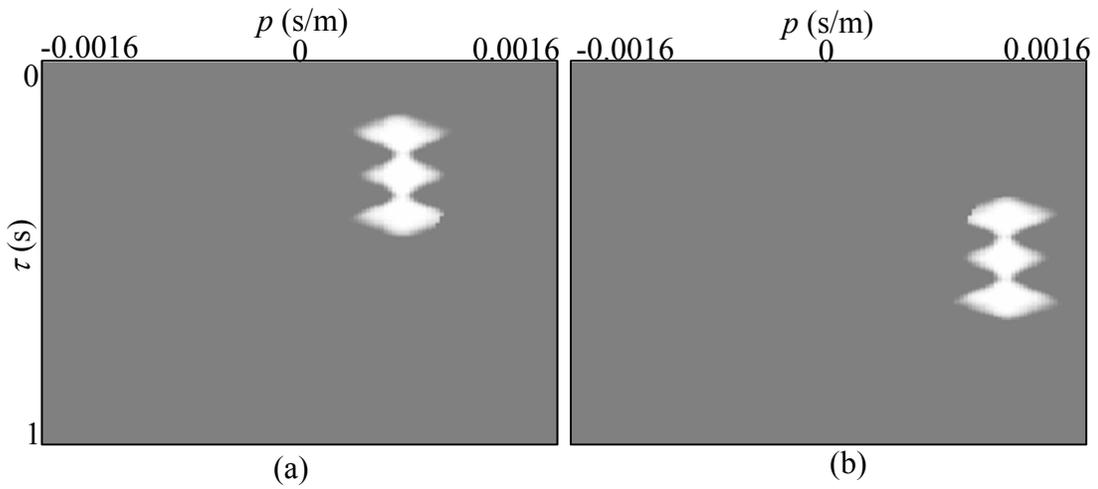


FIG. 4.20. Better semblances (Figure 4.19) give better P-wave model preconditioning matrix \mathbf{M}_p (a) and S-wave model preconditioning matrix \mathbf{M}_s (b) at iteration $n = 0$.

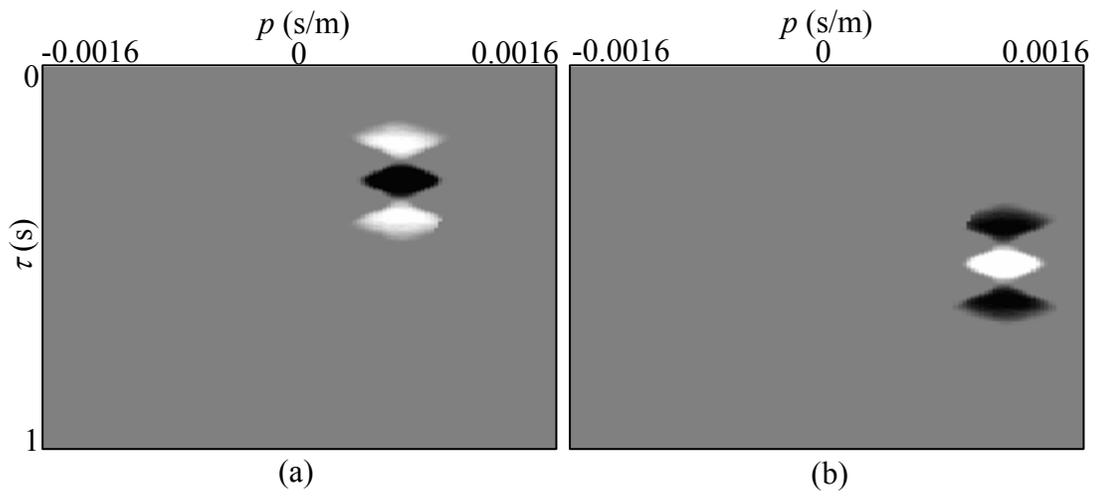


FIG. 4.21. Synthetic data with the noise P-wave conjugate gradient \mathbf{h}_p (a) and S-wave conjugate gradient \mathbf{h}_s (b) at iteration $n = 0$.

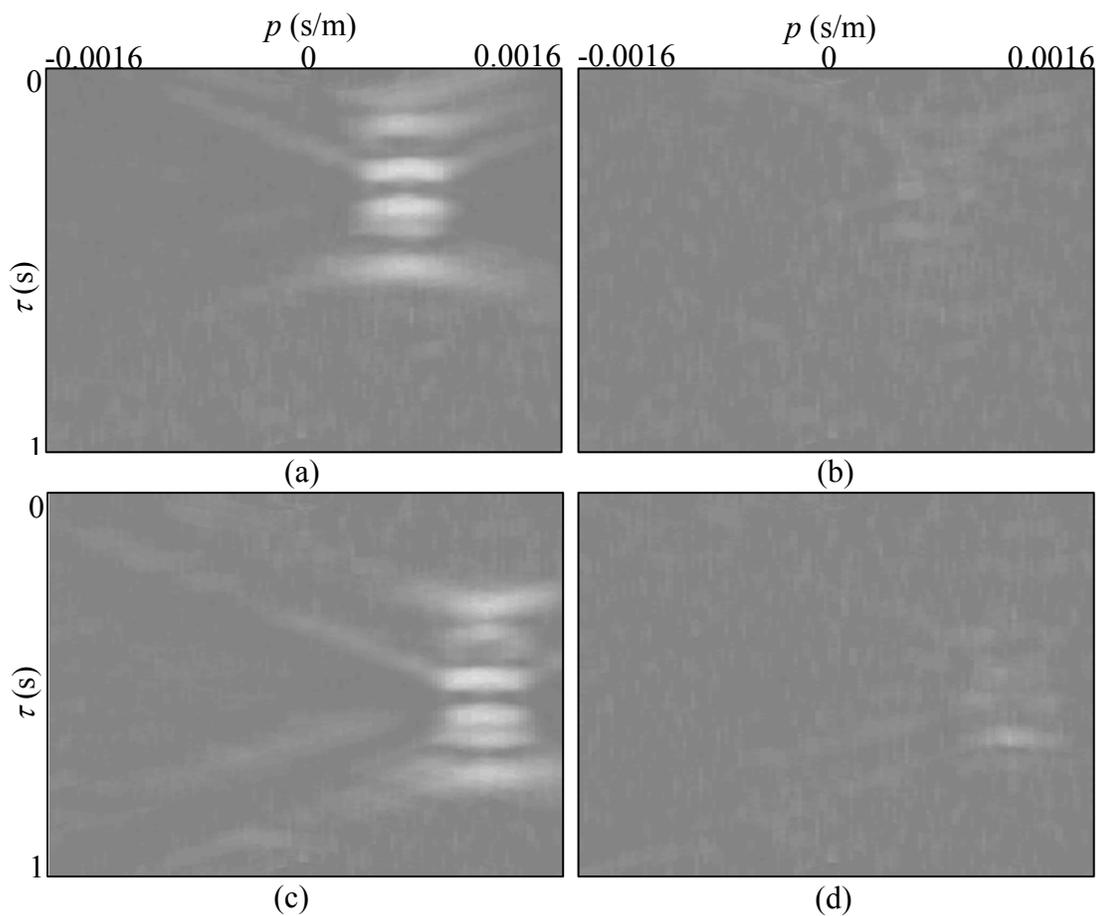


FIG. 4.22. Synthetic data with noise P-wave semblance s_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S-wave semblance s_s at iteration $n = 1$ (c) and at iteration $n = 10$ (d).

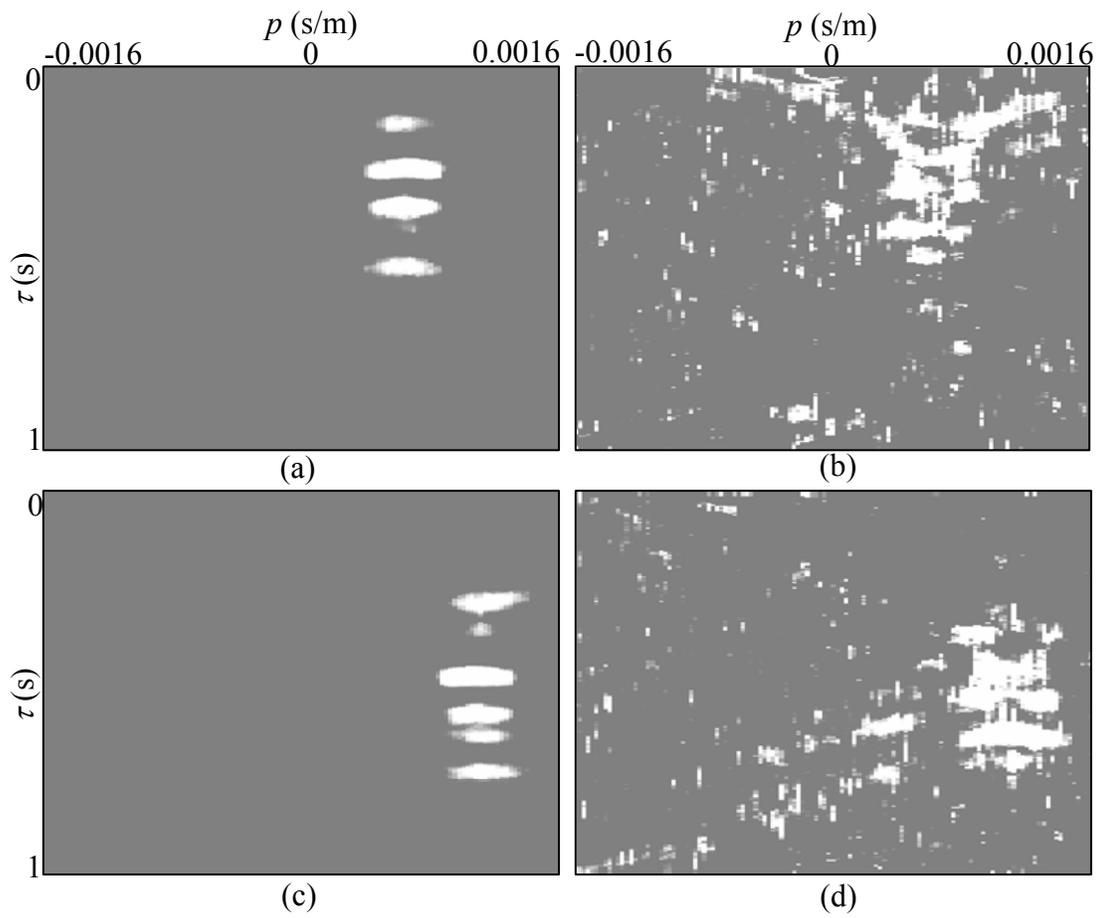


FIG. 4.23. Synthetic data with noise P-wave model preconditioning matrix \mathbf{M}_p at iteration $n = 1$ (a) and at iteration $n = 10$ (b); an S-wave model preconditioning matrix \mathbf{M}_s at iteration $n = 1$ (c) and at iteration $n = 10$ (d).

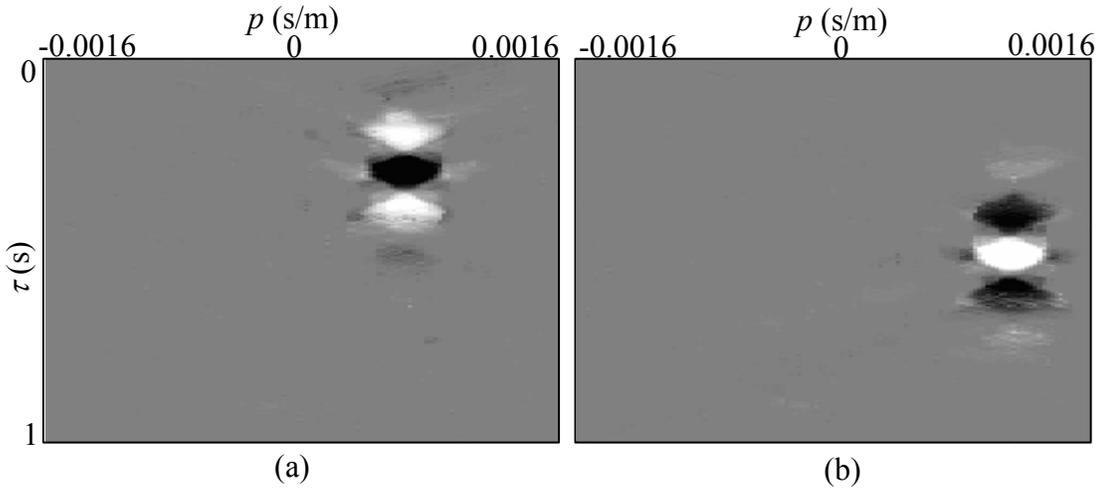


FIG. 4.24. Synthetic data with noise P-wave model \mathbf{m}_p (a) and S-wave model \mathbf{m}_s (b) at iteration $n = 10$.

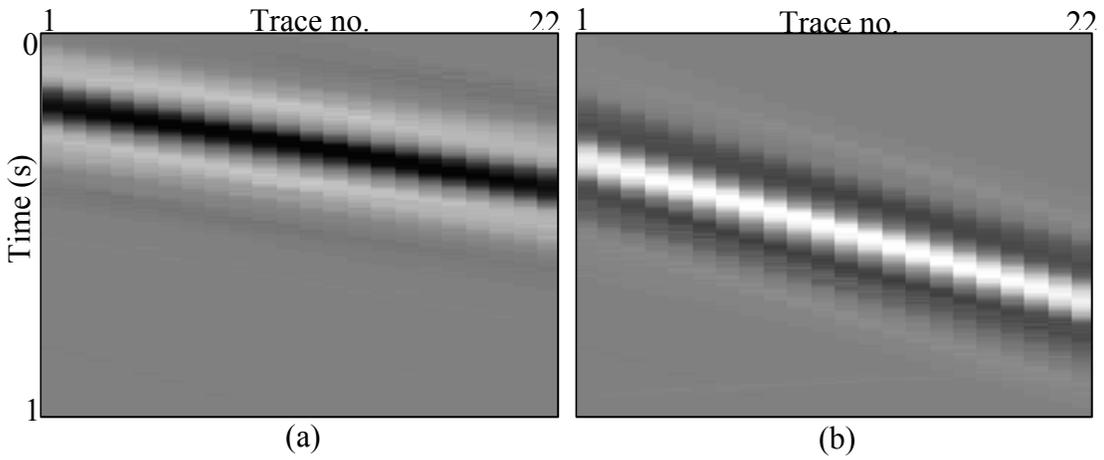


FIG. 4.25. Synthetic data with noise separated P-waves \mathbf{d}_p (a) and S-waves \mathbf{d}_s (b).

4.2.2. Vinton Dome test model

I described separation of the Vinton Dome synthetic data by classical separation procedures, the f - k filter and median filter in Chapter 2. I will use the same data (Figure 2.14) to test the new algorithm.

To assure better semblance calculation, I added random noise to the input data (Figure 4.26). Amplitude spectra of the synthetic data are shown in Figure 4.27. Required input parameters (Table 4.1) depend on the frequency range, offsets and velocities. I performed separation by the three different approaches described in section 4.1.1. of this chapter.

4.2.2.1. A fixed, large spatial window

In the first approach, I used the complete range of 61 traces as the input to the algorithm. This range of 915 m spans four different layers in the Vinton Dome model (Figure 2.8). Velocities from these layers give an average P-wave velocity of 2232 m/s and an average S-wave velocity of 1289 m/s. In this approach, the same average velocity is used in every angle of the incidence calculation.

I show two seconds of separated VSP data in Figure 4.27. Both strong downgoing P and downgoing S arrivals are successfully separated. However, there is a leakage of the transmitted downgoing P-waves into the S wavefield and vice versa. While I favor the slant stack with the higher energy, in the case when the energy is about the same, some leakage is almost inevitable. The large spatial window approach fits each of the arrivals with one ray parameter across the whole receiver array such that upgoing waves do not stop at the boundary, i.e. the intersection with the transmitted downgoing wave. The most obvious problem with the large spatial window approach is in separating the salt

reflection. Beside the fact that we try to approximate hyperbola with one linear ray, a problem that demands many more ray parameters is the change in the salt reflection polarity.

In order to estimate quality of this separation technique, I compared wavefields with the $f-k$ separation results. Since I separated x-component P-waves and S-waves separately from the z-component wave modes (Figures 2.18 and 2.19), I summed horizontal and vertical component P-waves (Figures 2.18b and 2.19b) and horizontal and vertical component S-waves (Figures 2.18c and 2.19c) to obtain comparable wavefields (Figure 4.29). Both techniques have leakage of one wave mode to the other and they both fail to separate P from the S salt face reflection. However, the $f-k$ separation results are highly contaminated with the incorrect mute selection artifacts which do not exist on the real data. The Radon transformation technique does not produce those artifacts.

4.2.2.2. A small, running spatial window

In this type of filtering, we form a spatial window using a fixed number of traces and apply the algorithm. We then shift the window position by one trace and apply the algorithm again. Here I used a 21 trace running window. The first window includes traces 1-21, the second 2-22, the third 3-23 and so on (Figure 4.30). For the array of 61 input traces (Figure 4.26), the last window is 41-61. From every window application run, I used the middle output trace only, i.e. traces 11, 12, 13....51, and obtained separate P (Figure 4.31a) and S wavefields (Figure 4.31b). Though not separated by a running window technique, one can use separated traces 1-11 from the first window and 51-61 from the last window to fill up 'missing' parts and avoid lost data (Figure 4.31).

Comparing the results obtained by one large spatial window filtering approach (Figure 4.28) and $f-k$ separation (Figure 4.29), I note that the running window technique has less leakage of one mode to the other and better modeling of the salt reflections (Figure 4.31). Unfortunately, inaccurate modeling still occurs at discontinuities.

4.2.2.3. Fixed, sequential windows

In the third filtering approach, I exploit the fact that we can estimate major impedance discontinuities within the well and reliably determine our spatial window size. As was shown in the Vinton Dome model (Figure 2.8), the VSP well spans four different layers. Traces 1-28 are within the first layer with $v_{p1}=1948$ m/s and $v_{s1}=1125$ m/s, traces 29-41 are in the second layer with velocities $v_{p2}=2250$ m/s and $v_{s2}=1298$ m/s, traces 42-53 span third layer with velocities $v_{p3}=2307$ m/s and $v_{s3}=1332$ m/s, and finally, the last traces 54-61 are in the fourth zone where we have $v_{p4}=2426$ m/s and $v_{s4}=1402$ m/s. Based on this, I divided the input data into four zones and then applied my separation algorithm. Obtained results are shown in Figure 4.32. Linear events are properly modeled such that we do not see upgoing waves beyond the transmitted downgoing P-wave. As in the previous approaches, the linear Radon transform does not accurately fit the hyperbolic moveout of the salt face reflections. In addition, the blocky (fixed window) Radon transform tries to fit each part of the salt reflection with the most appropriate ray parameter for that zone. Unfortunately, the fixed window Radon transform introduces discontinuities in the continuous P and SV salt reflections. For these events, it appears that the running window approach better approximates continuous, hyperbolic events.

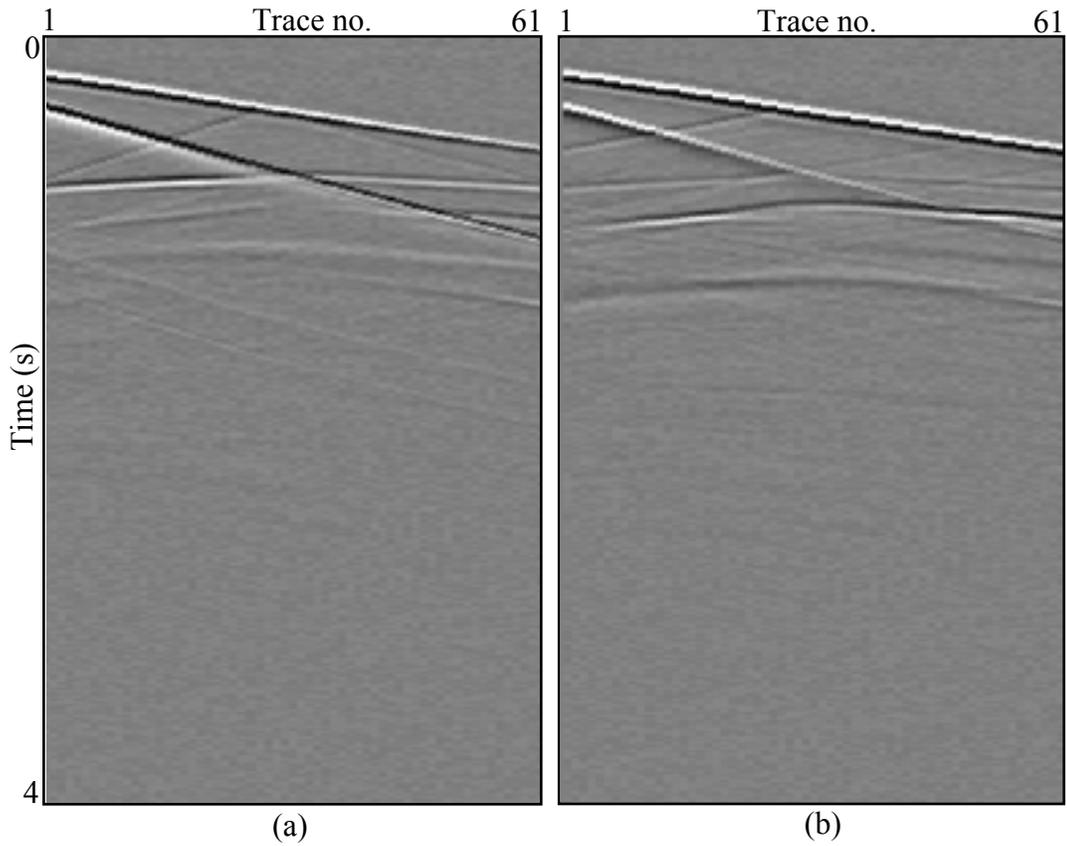


FIG. 4.26. Vinton Dome model synthetic data (Figure 2.12) recorded on the horizontal (a) and vertical component (b). Random noise is added to the data to assure a better semblance calculation.

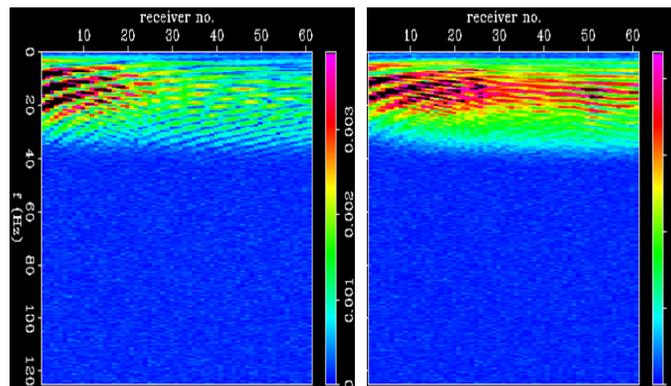


FIG. 4.27. Vinton Dome synthetic data from Figure 4.26 x-component (a) amplitude spectrum and z-component amplitude spectrum (b).

p_{\min}	-0.0016 s/m
p_{\max}	0.0016 s/m
x_r	1200 m
N	10
ϕ	0.7
ξ_{\min}	0.1
f_1	0.8 Hz
f_2	2 Hz
f_3	50 Hz
f_4	70 Hz

Table 4.1. Required input parameters for the algorithm application on the synthetic Vinton Dome data from Figure 4.26.

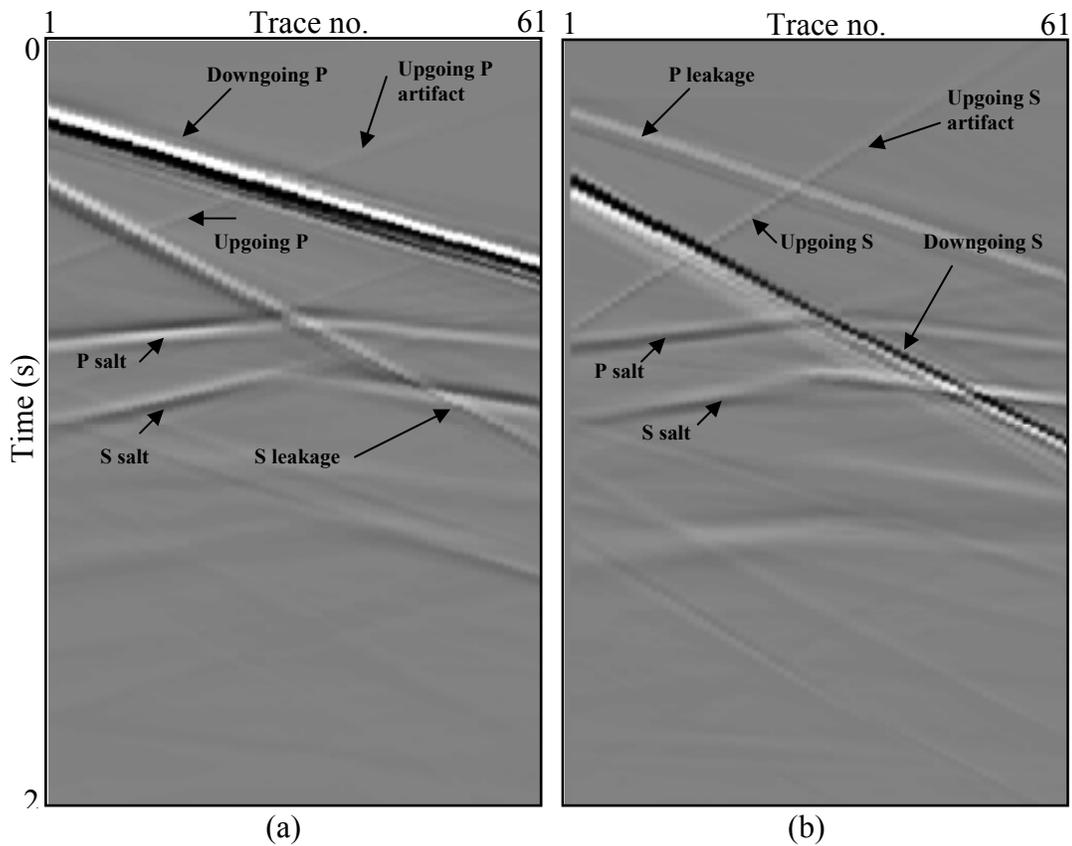


FIG. 4.28. Separated synthetic P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the fixed, large spatial window approach.

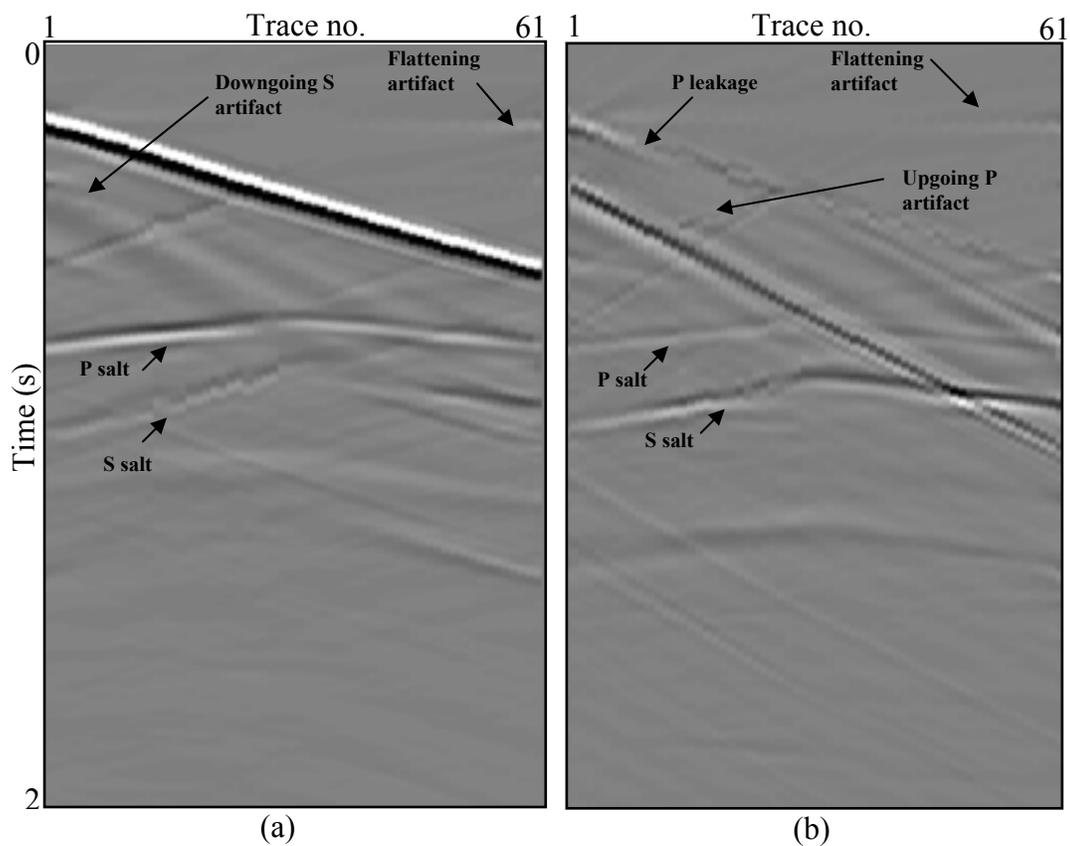


FIG. 4.29. The f - k separated downgoing and upgoing P-waves from Figures 2.23b and 2.24b are summed in order to obtain a P wavefield comparable to the P wavefield from Figure 4.28a. Downgoing and upgoing S-waves from Figures 2.23c and 2.24c give an S wavefield comparable to the S wavefield from Figure 4.28b.

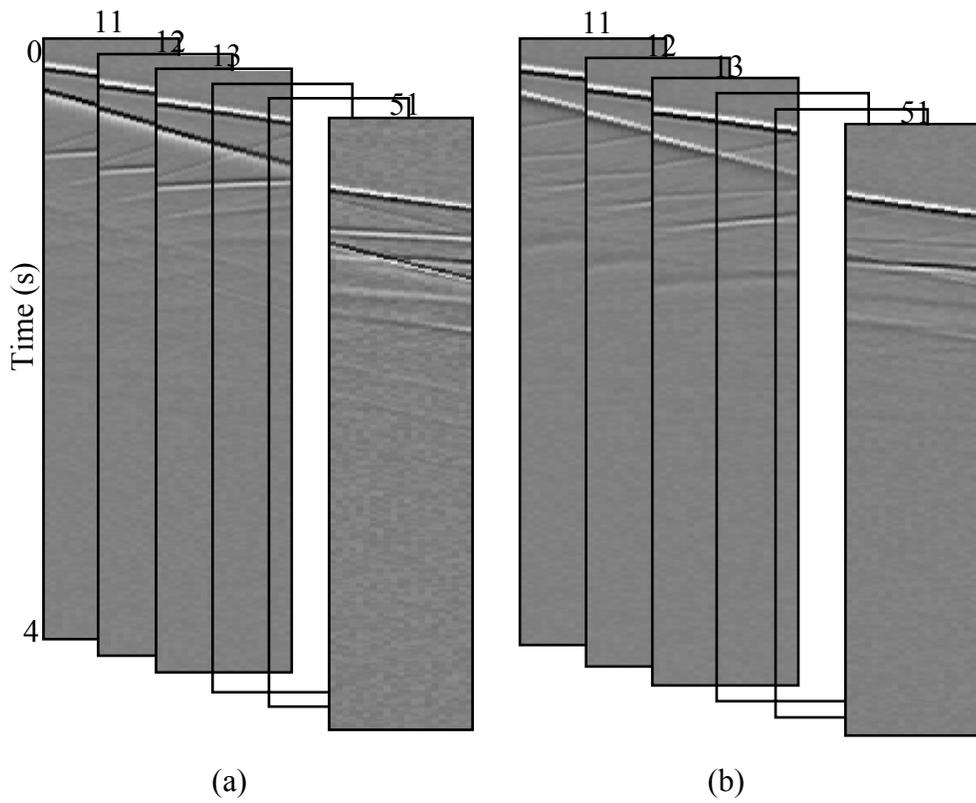


FIG. 4.30. Algorithm application by a small, 21 trace long running spatial window. Input x-component (a) and input z-component (b) are processed as 40 separate algorithm runs.

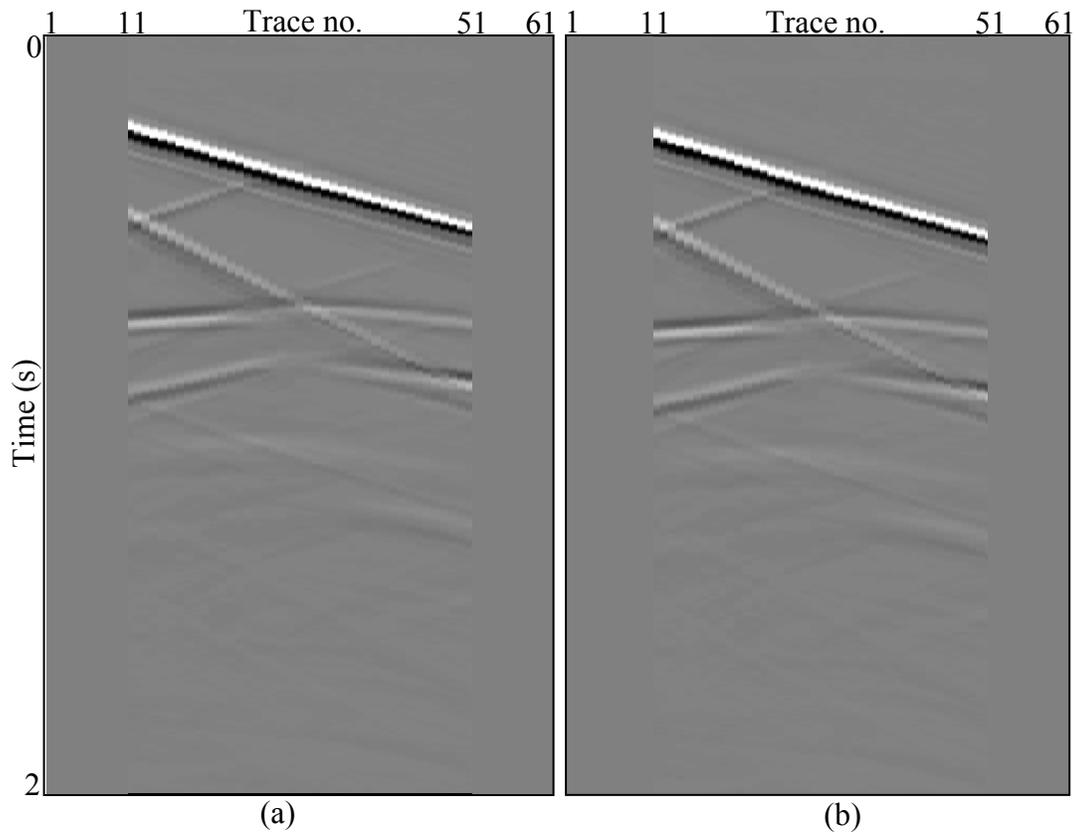


FIG. 4.31. Separated synthetic P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the small, running spatial window approach. Each window includes 21 traces, such that the first obtained trace is 11th and the last 51st trace.

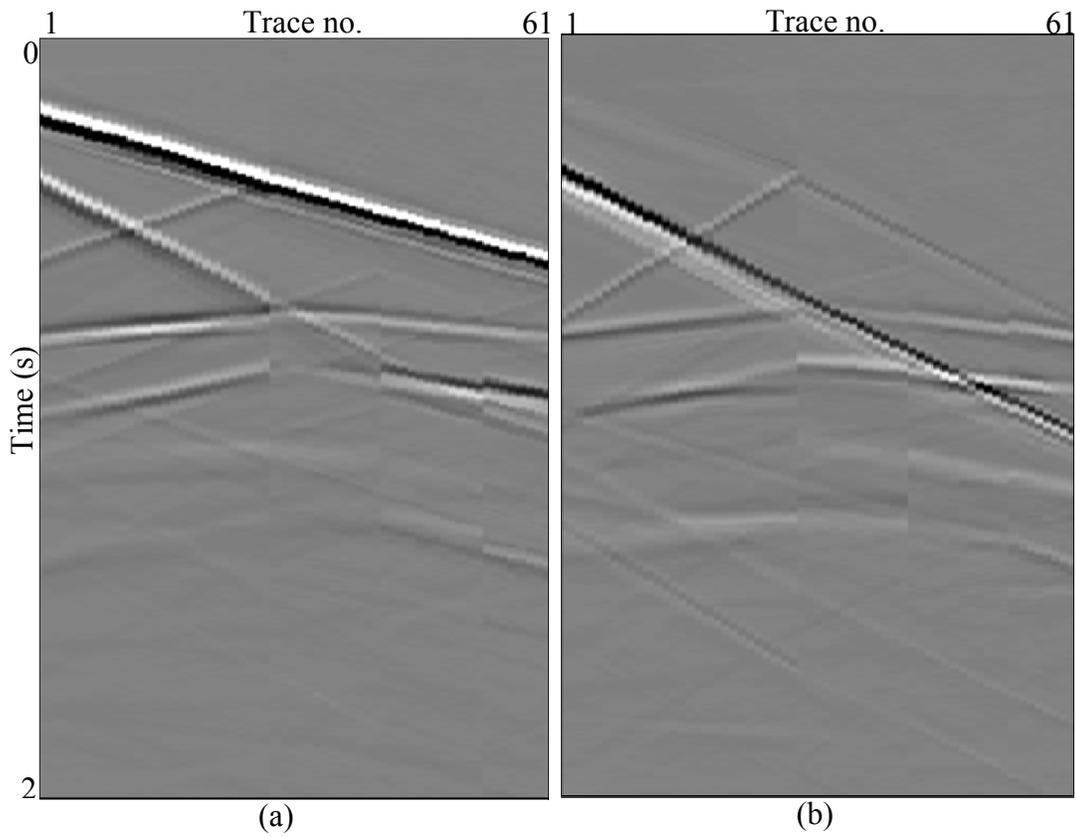


FIG. 4.32. Separated synthetic P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the fixed, zonal spatial window approach. Window sizes are determined based on the information about the major discontinuities within the well.

4.2. Algorithm testing on the Vinton Dome VSP data

I am now ready to test my high resolution discrete Radon transform separation algorithm on the one second of the recorded radial and vertical component from the shot 3068 (Figure 2.5). The input parameters for algorithm application (Table 4.2) are calculated to prevent edge effects (high ray parameter value). Also, setting the intercept value, $\tau_{max}=1.5s$, to be higher than the length of the input data, $t_{max} = 1.0 s$, allows proper modeling of the steep events within the data. The input velocity file is the same as for the synthetic data. As we can see from the obtained results (Figure 4.33), only the strongest events, i.e. downgoing events are partially modeled.

To allow modeling of all events, I have applied amplitude balancing and muting to the input shot gathers (Figure 4.34). To balance the data, I multiplied each sample by t^4 event such that early and late events of a seismogram have similar amplitudes. Such a deterministic gain can easily be removed by multiplying the result of filtering by t^{-4} . I muted all events to be zero before the first arrival. The separated P and S wavefields (Figure 4.35) now contain all arrivals. In order to check quality of separation, I reconstructed radial and vertical components (Figure 4.36), and calculated residuals between the original data and the reconstructed components (Figure 4.37).

4.2.1. Tests of constraints

In order to understand the importance of applied constraints for wave mode separation, I will compare three examples with different constraints parameter.

To evaluate an unconstrained vector Radon transform, I set a value my model weights to one, $\mathbf{M}=1$. Since all events are equally favored, the separation is based on the

ray parameter and polarization only. In Figures 4.38a and 4.38b, I note that \mathbf{m}_p and \mathbf{m}_s have strong dispersion due to data aliasing.

In the next example, I apply the constraints described in the previous chapter, but allow all τ - p events to be modeled at the last iteration (Figures 4.38c and 4.38d).

In the final example, I maintain constraints for every iteration (Figures 4.38e and 4.38f).

The reconstructed P and S wavefields are shown in Figure 4.39. Without constraints, P wavefield (Figure 4.39a) and S wavefield (Figure 4.39b) are very poor. There are also strong edge effects. The sparsely constrained results (Figures 4.39c and 4.39d) have better P and S separation and there are minimal edge effects. However, the most constrained separated data (Figures 4.39e and 4.39f) show the most coherent arrivals.

4.2.2. Spatial window test

As it was explained in section 4.1.1. of this chapter and later applied to the synthetic Vinton Dome data, there are three possibilities for spatial τ - p filtering.

Results from application of one large spatial window are shown in Figure 4.40. Most of the S wavefield events stretch across whole shot and that is the major problem with this window type. However, with a running window application (Figure 4.41) events are better positioned, more coherent and better separated. I have divided data into four zones for the fixed, zonal window application: traces 1-17, traces 18-33, traces 34-45 and 46-61. While some deeper events become most coherent with this type of spatial filtering (Figure 4.42), there are still artificial brakes and artifacts.

4.2.3. Offset test

I tested the algorithm on a near (shot 3060, offset 300 m), mid (shot 3068) and far (shot 3046, offset 1000m) offset data. Inputs to the algorithm are two seconds of the recorded radial and vertical component. Separated P and S wavefields for shot 3060 (Figure 4.43) and shot 3068 (Figure 4.44) have better resolution than the results obtained for the long offset data (Figure 4.45). However, the quality of the separation highly depends on the rotation, number of iterations, relaxation and stopping criteria.

4.2.4. Data separation by filtering in τ - p space

There are two ways for data filtering in τ - p space. One way is to select ray parameter values to be muted. The other way is to select a range of angles to be muted.

With a proper mute selection in the model space, one can filter the data. As it was described in section 4.1.2, this is a simple way for separation of the downgoing from the upgoing VSP waves. Muting the waves with negative ray parameters, I allowed modeling of downgoing P-waves and downgoing S-waves only (Figures 4.46a and 4.46b). Reversing the mute separates upgoing P and upgoing S wavefields from the downgoing waves (Figures 4.46c and 4.46d).

Beside downgoing-upgoing separation, we can select a certain range in the τ - p space we want to filter. In Figure 4.47 I passed arrivals that are more horizontal and muted steeper events. Though it is hard to select a proper mute range, this can be one of the possible methods to filter salt reflections from the rest of the data. The large edge effects are due to an abrupt mute selection in the τ - p space; a taper would avoid this problem.

To better image horizontal salt reflections, I have filtered data based on the incidence angles ψ_p and ψ_s . Since we know from my modeling that the salt reflection travel horizontally to the VSP well, I have selected angles -30° to $+30^\circ$ from the positive x-axis for the P-waves. This range is the same for the S-waves ($+30^\circ / -30^\circ$), but the filtered zone in τ - p space is larger due to the smaller velocity values. With this filtering approach we are able to extract any particular wave and to avoid strong τ - p filtering artifacts.

p_{\min}	-0.0023 s/m
p_{\max}	0.0023 s/m
τ_{\max}	1.5 s
x_r	1200 m
N	20
ϕ	0.7
ξ_{\min}	0.1
f_1	0.8 Hz
f_2	2 Hz
f_3	90 Hz
f_4	120 Hz

Table 4.2. Required input parameters for the algorithm application on the real Vinton Dome data from Figure 4.35.

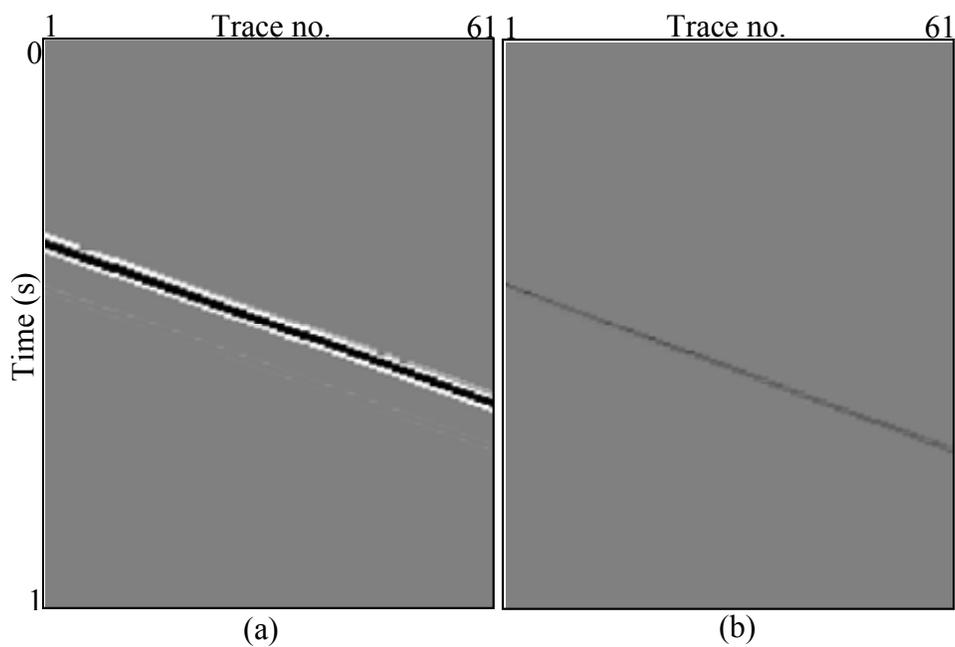


FIG. 4.33. Separated P-waves (a) and S-waves (b) for shot 3068. Due to the strength of the first arrivals, we can not model other events.

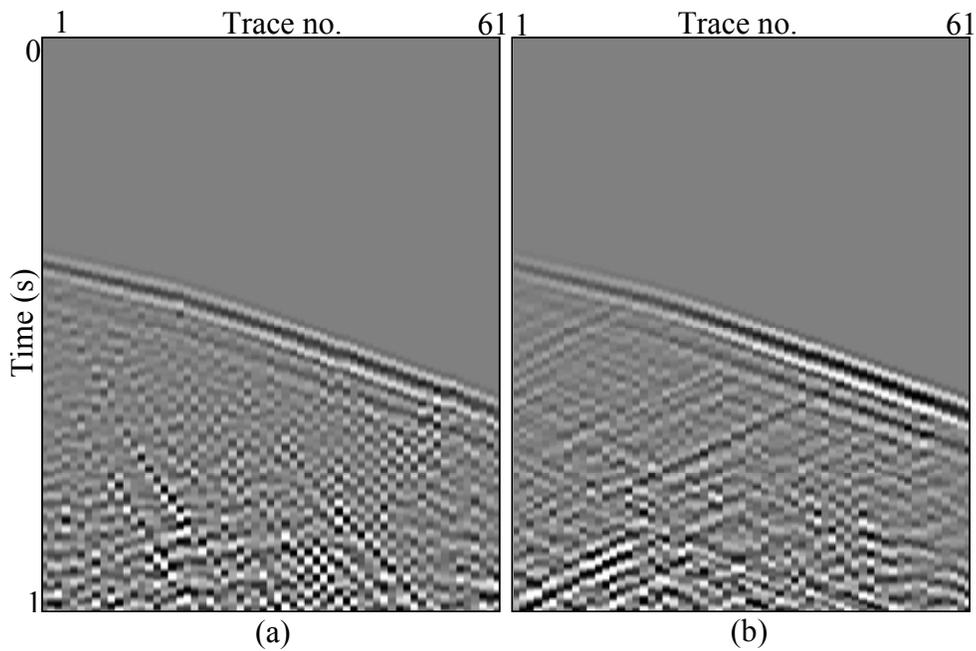


FIG. 4.34. Scaled and muted radial (a) and vertical (b) component, shot 3068, offset 600 m.

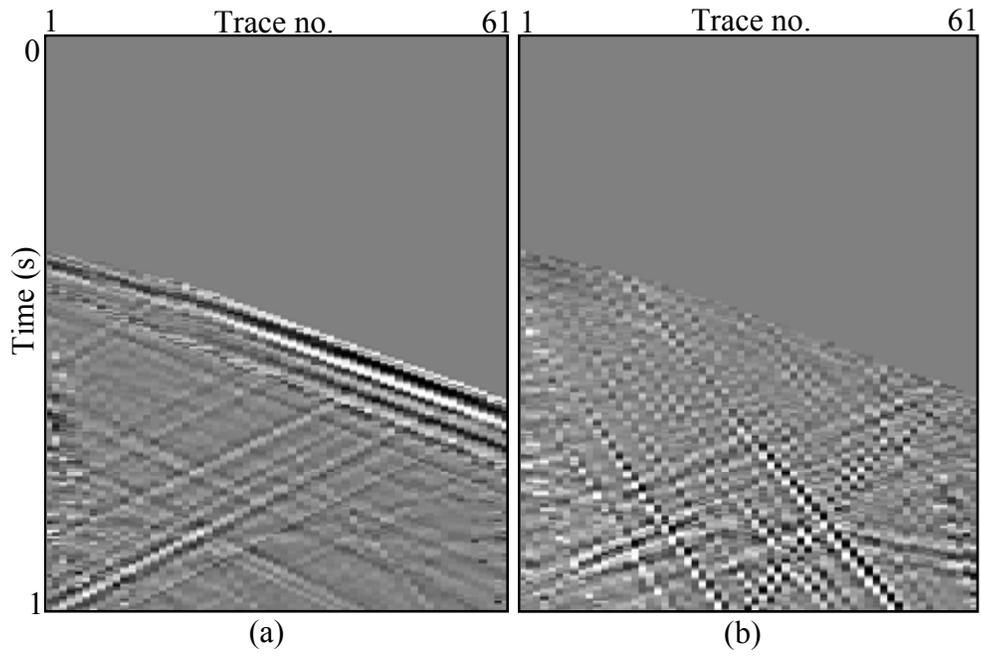


FIG. 4.35. Separated P (a) and S (b) wavefields for shot 3068.

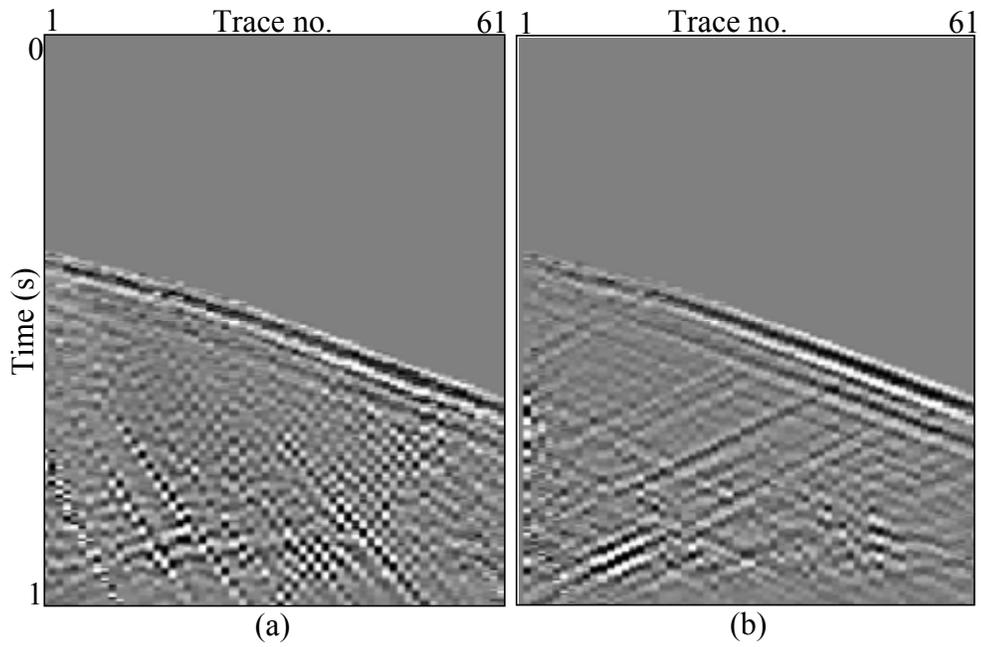


FIG. 4.36. Reconstructed radial (a) and vertical component (b) for shot 3068.

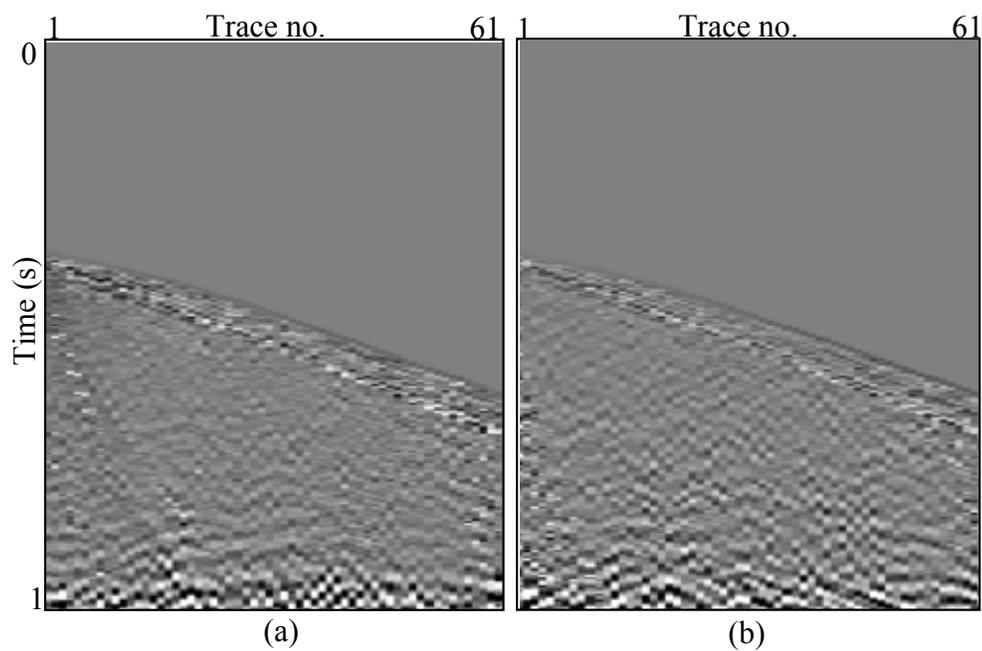


FIG. 4.37. (a) Residual between the original (Figure 4.37a) and reconstructed radial component (Figure 4.39b) and (b) residual between the original (Figure 4.37b) and reconstructed vertical component (Figure 4.39b).

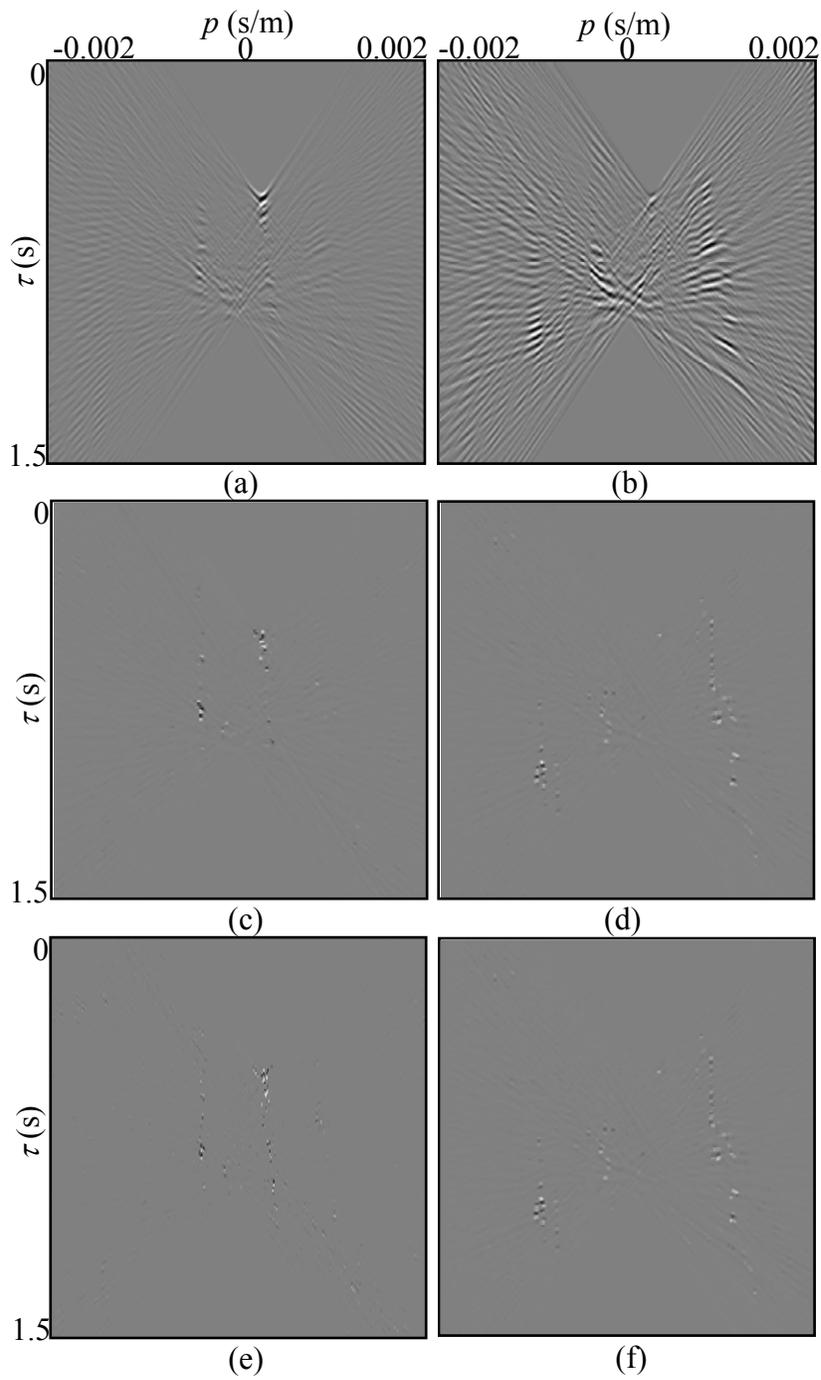


FIG. 4.38. The last iteration models \mathbf{m}_p and \mathbf{m}_s depend on the constraints applied: no constraints at all P-wave model (a) and S-wave model (b), sparse constraints P-wave model (c) and S-wave model (d) and constrained P-wave model (e) and S-wave model (f).

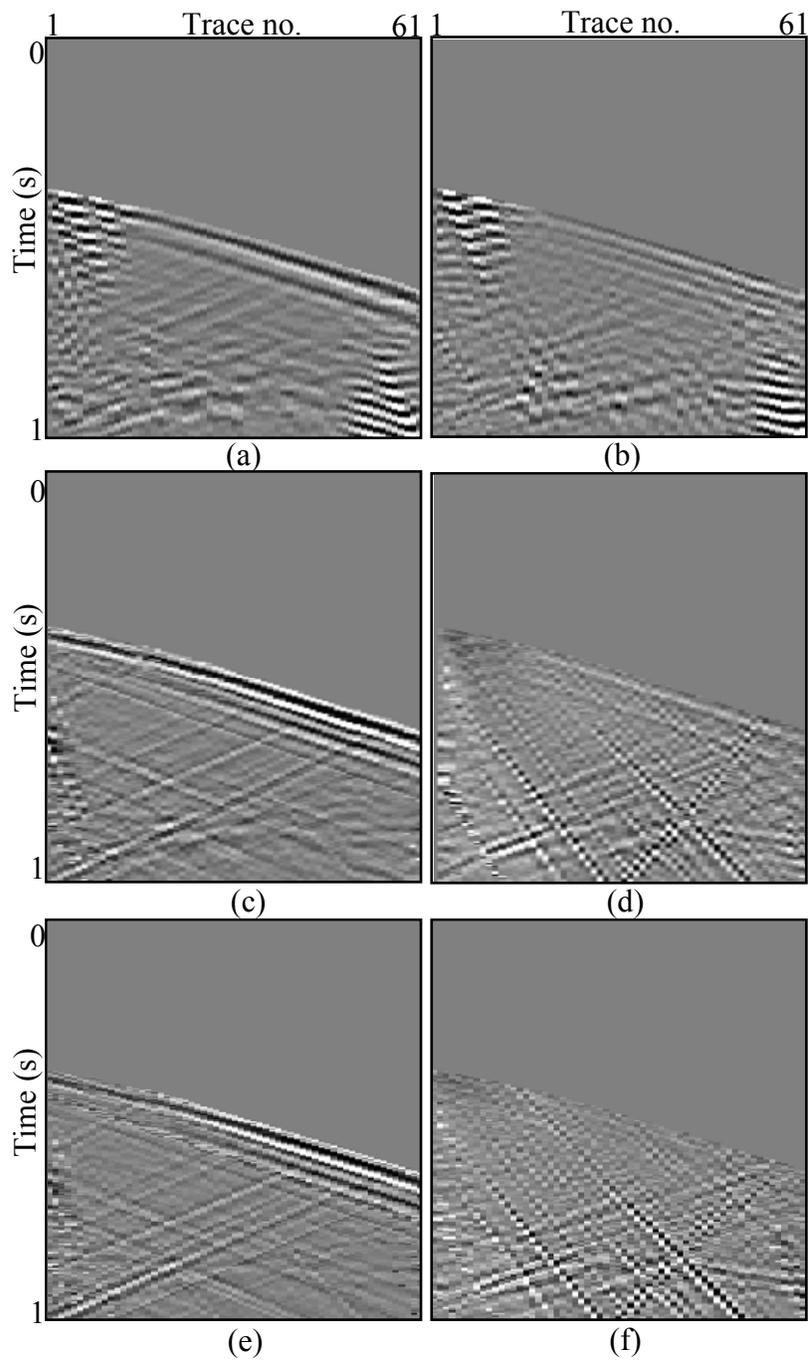


FIG. 4.39. P wavefield (a) and S wavefield (b) with no constraints applied, sparse constraints P wavefield (c) and S wavefield (d), and constrained P wavefield (e) and S wavefield (f).

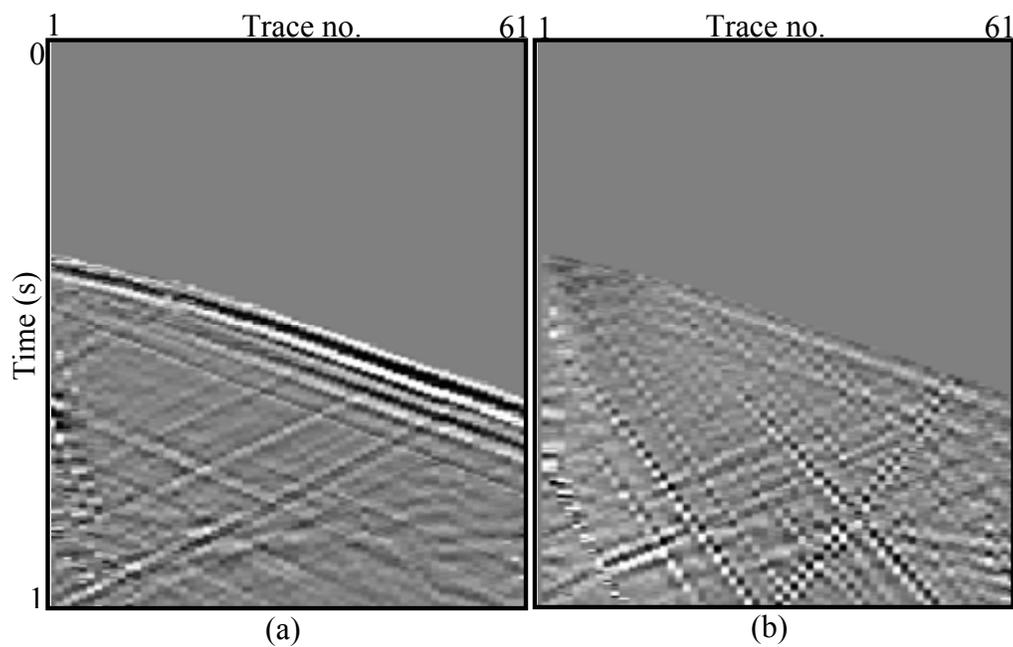


FIG. 4.40. Separated Vinton Dome shot 3068 P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the fixed, large spatial window approach.

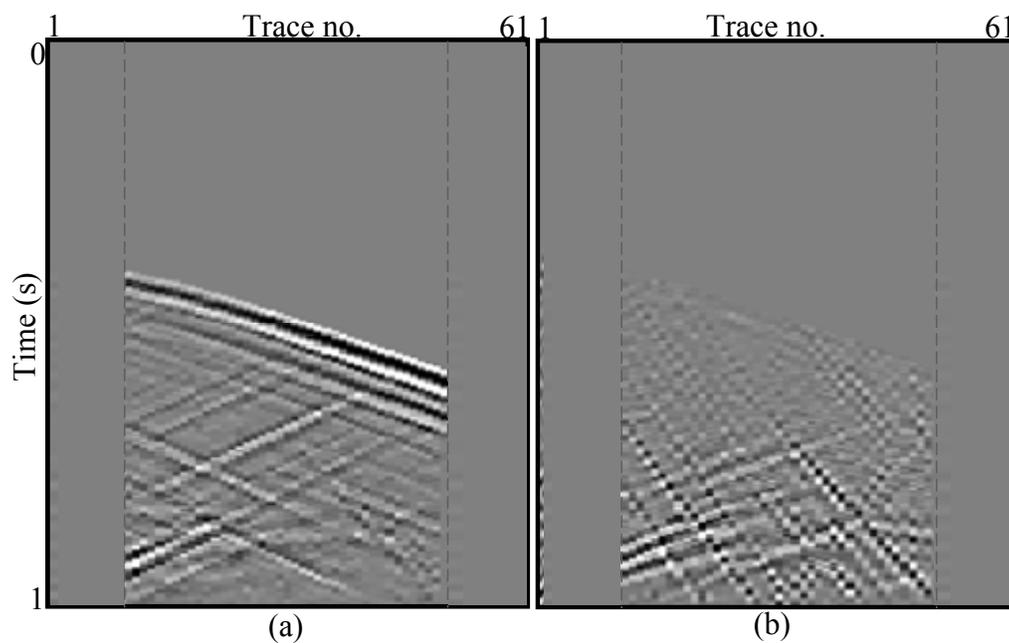


FIG. 4.41. Separated Vinton Dome shot 3068 P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the small, running spatial window approach. Each window includes 21 traces, such that the first obtained trace is 11th and the last 51st trace.

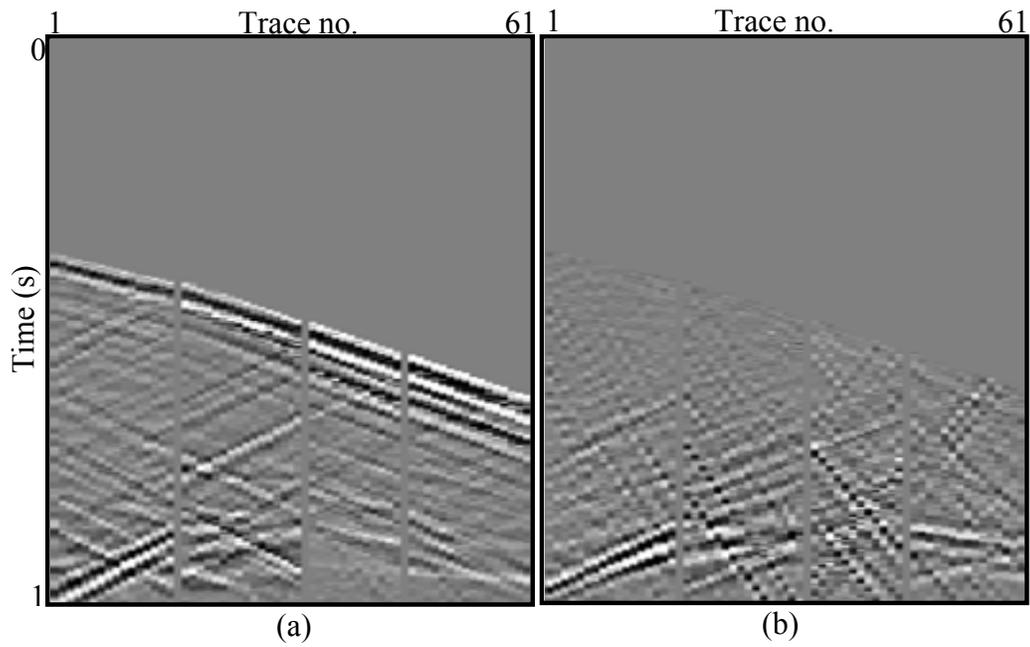


FIG. 4.42. Separated Vinton Dome shot 3068 P-waves (a) and S-waves (b). High resolution discrete Radon transform algorithm is applied using the fixed, zonal spatial window approach. Window sizes are determined based on the information about the major discontinuities within the well. I have added a gap to show the position of the windows.

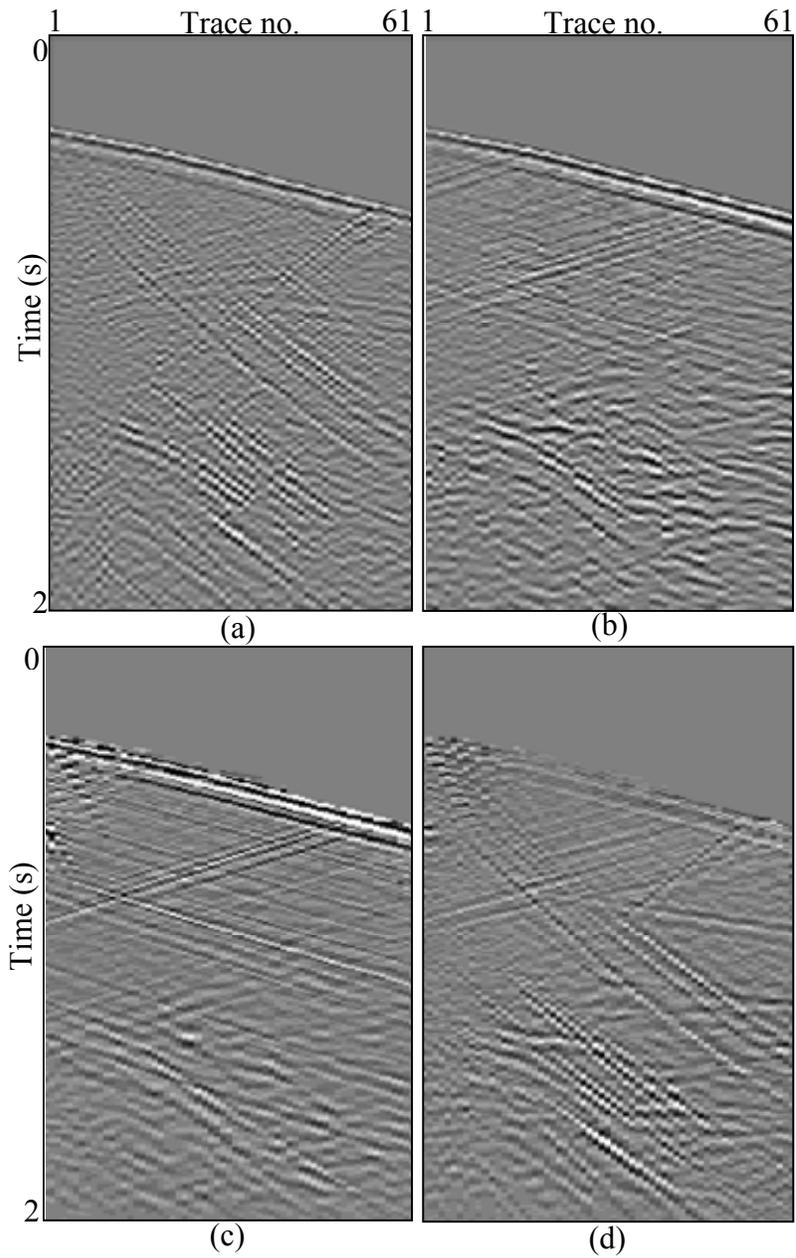


FIG. 4.43. Shot 3060, near offset 300 m, radial (a) and vertical (b) components. High resolution discrete Radon transform algorithm separates VSP data into P wavefield (c) and S wavefield (d).

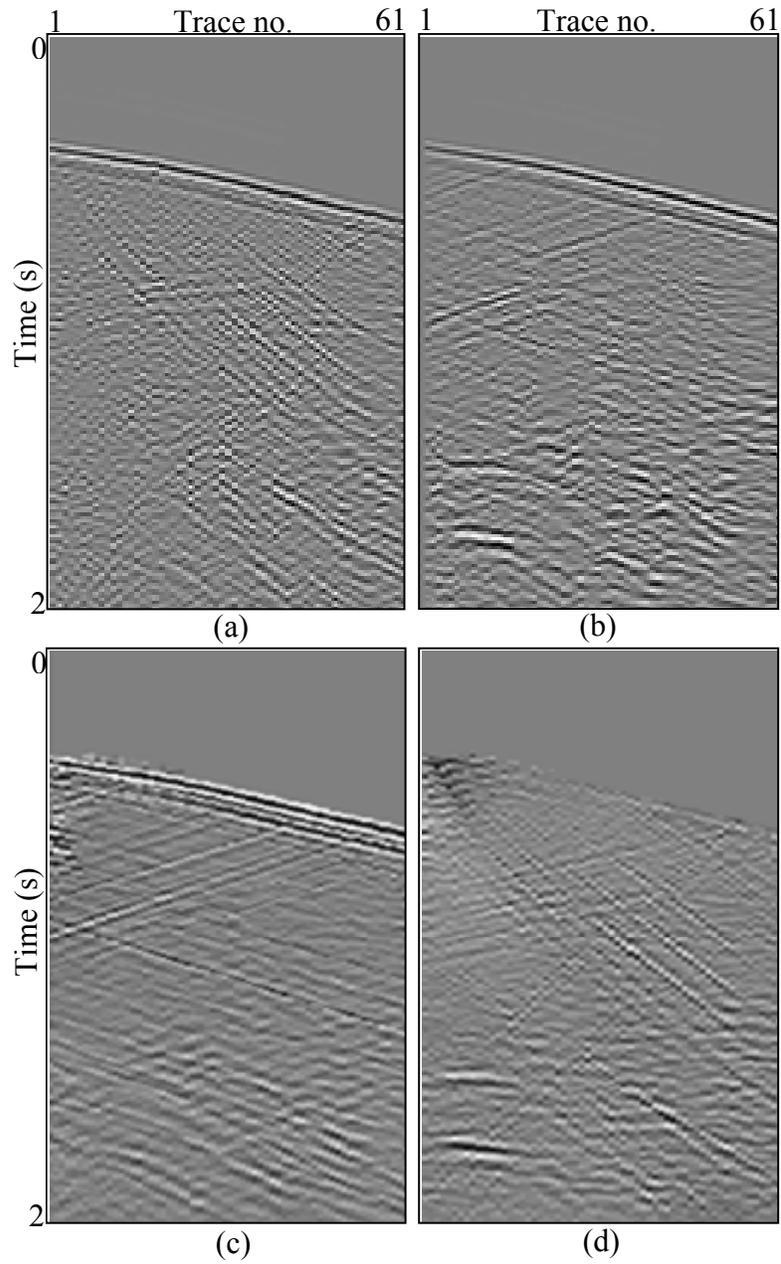


FIG. 4.44. Shot 3068, mid offset 600 m, radial (a) and vertical (b) components. High resolution discrete Radon transform algorithm separates VSP data into P wavefield (c) and S wavefield.

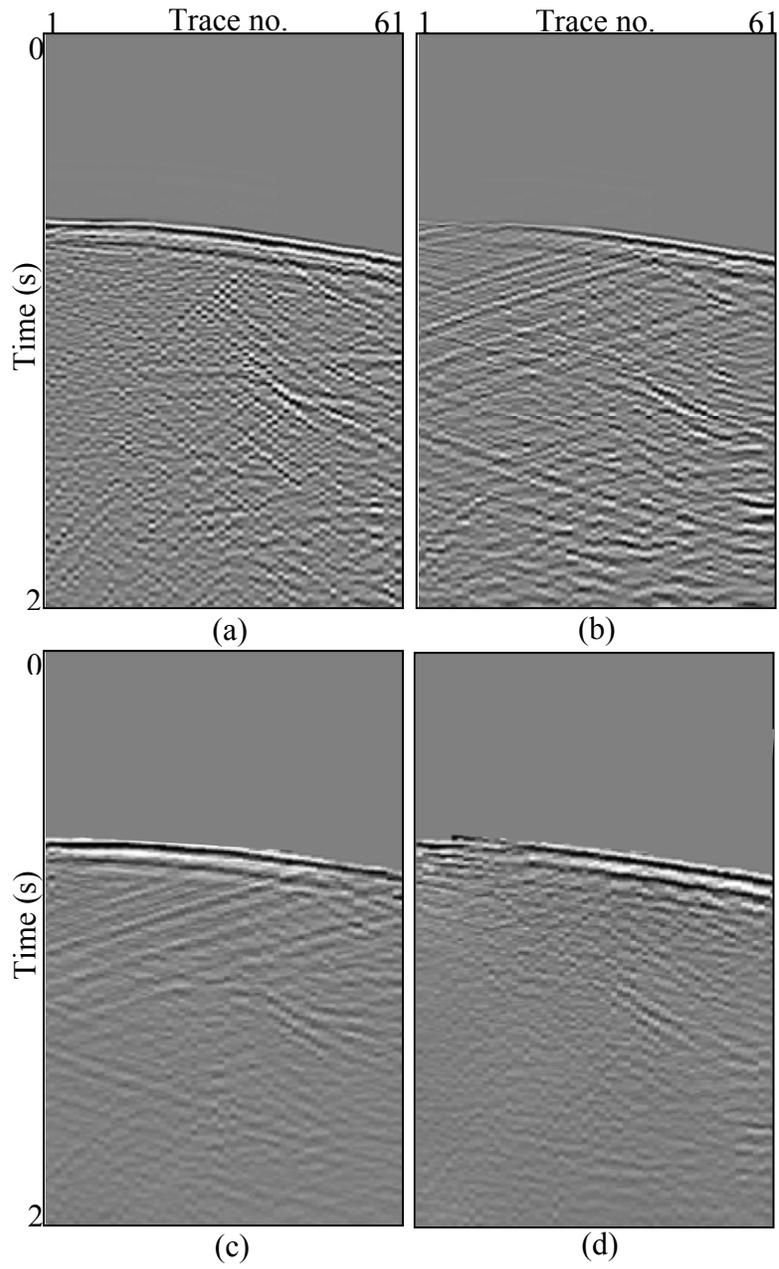


FIG. 4.45. Shot 3046, far offset 1000 m, radial (a) and vertical (b) components. High resolution discrete Radon transform algorithm separates VSP data into P wavefield (c) and S wavefield.

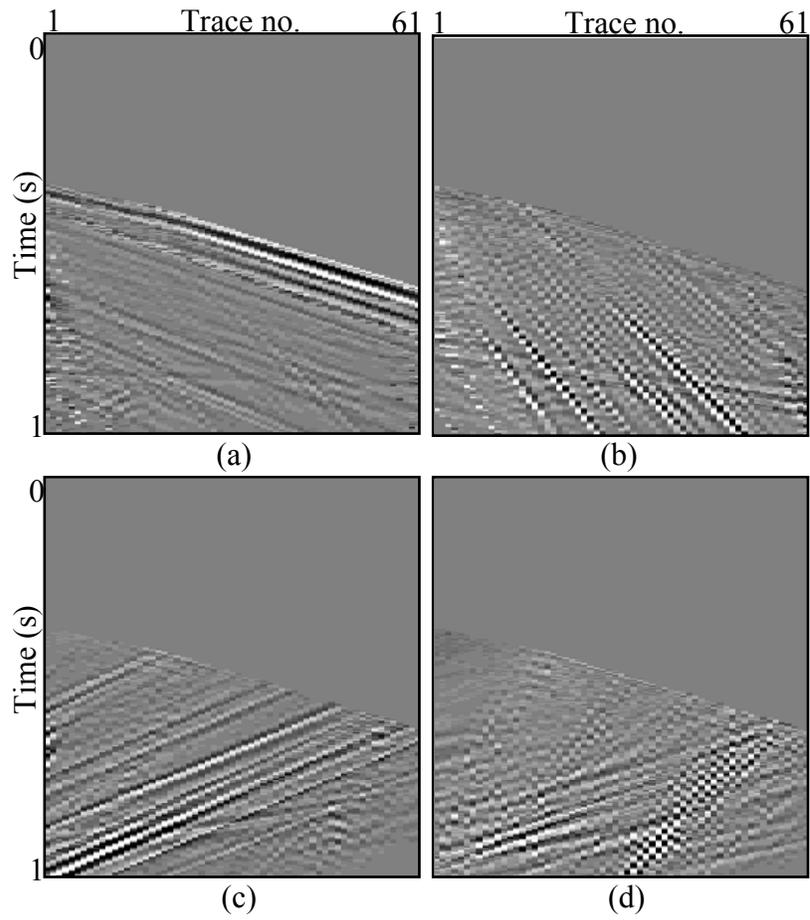


FIG. 4.46. Negative ray parameters in the P and S model space are muted to separate downgoing P-waves (a) and downgoing S-waves (b). Selection of the opposite mute, filters upgoing P-waves (c) and upgoing S-waves (d).

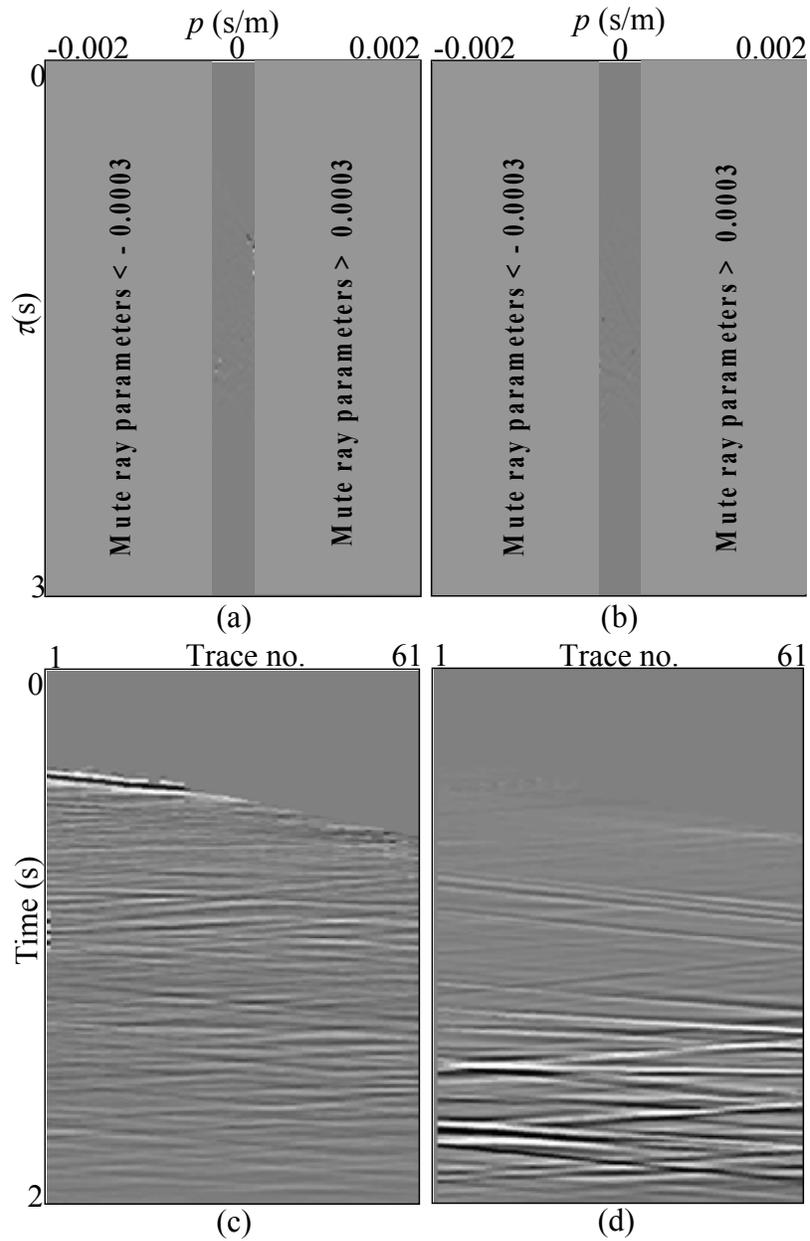


FIG. 4.47. Mute can be selected for any range of P-waves ray parameter values (a) and S waves ray parameter values (b). Passing the values with the ray parameters close to zero, one can filter out steeper events.

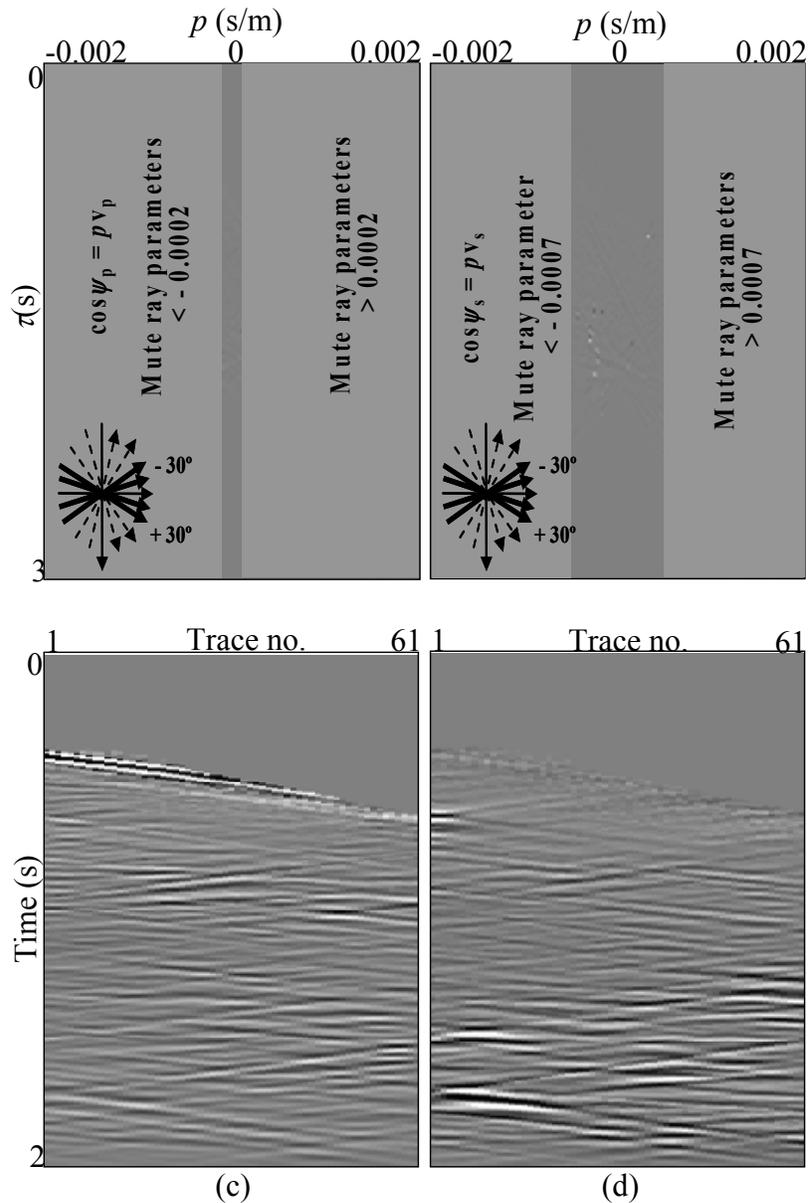


FIG. 4.48. Data are filtered based on the incidence angle ψ . For the extraction of the sideways salt reflections, selection of more horizontal events filters the possible salt reflections from the rest of the wavefield (a) and (b). For the extraction of P-wave salt reflections (c), I have selected waves with the incidence angles between $+30^\circ$ and -30° . The same range is selected for the S-wave salt reflections (d).

5. CONCLUSIONS

Despite a good acquisition pattern, separation of P and SV waves from a multicomponent 3-D VSP is hard due to hyperbolic events from complicated salt dome geology. Additionally, long offsets and aliased data make the classical f - k or median filter separation less effective.

I present a new separation algorithm in this thesis that uses a high resolution discrete Radon transform approach. Performed in the time, rather than the frequency domain, my Radon transform is both an antialiasing and polarization filter. The major advantages of the developed multicomponent conjugate gradient solution are: the iteration process can be stopped at any time, providing a partial solution, and the user can specify quality of the separation giving the relaxation criteria and number of iterations.

I have shown on both synthetic and real data corresponding to Vinton Dome, that my algorithm successfully separates major linear events, though with some minor leakage of one wave mode into the other. Unfortunately, this method requires the model to be a good fit to the data, and because it is a linear Radon transform, it fails to fit my hyperbolic salt reflections.

Additional problems related to a linear model approximation are 'kinks' in the VSP arrivals and projection of some upgoing events beyond transmitted downgoing wave. These problems are successfully overcome using the different spatial windows. Good results are obtained both with a running window and with a velocity based zonal window. Unfortunately, this window approach significantly increases the already high costs of the high resolution Radon transform.

Tests performed on the several different Vinton Dome shots show that quality of the obtained results depends on the preprocessing of the data. This algorithm is strongly affected by the quality of the rotation to the radial and transversal component and strength of the first downgoing arrival, such that some scaling is necessary before the algorithm application. In general, this code requires good balancing of all input parameters.

As it was expected, challenges due to the long offsets and lack of resolution remain. Separation results are better on the near and mid offsets than on the far offset data.

Using the semblance and energy driven constraints, one of the major strengths of this algorithm can be exploited for the purposes of data filtering. Once separated to the P and S wavefields, data can be filtered and separated to the downgoing, upgoing, or any wave type using the mute in the τ - p space. Unfortunately, this type of filtering may produce severe edge effects due to the mute selection. I solved this problem using an incidence angle filter, where any particular wave can be filtered based on the direction of arrival to the well in the vertical plane between a source and receiver.

Extension of this approach to include azimuthal direction for the τ - p - φ space VSP separation is going to be the next step in the high resolution discrete Radon transform algorithm development for P and S waves separation of coarsely sampled 3-D 3-C VSP data.

REFERENCES

- Aki, K. and Richards, P., 1980, Quantitative seismology: Theory and methods, W. H. Freeman and Company.
- Boelle, J.-L., Canadas, G. and Hugonnet, P., 1998a, Wavefield separation in borehole seismic by Linear Radon Decomposition, 68th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 272-275.
- Boelle, J.-L., Kaiser, P. and Maocec, E., 1998b, Difficulties and clues in 3-D VSP processing, 68th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 373-376.
- Chopra, S., Alexeev, V., Manerikar, A., Kryzan, A., 2004, Processing/integration of simultaneously acquisition of 3D surface seismic and 3D VSP data, The Leading Edge, **23**, 422-430.
- Claerbout, J. F., 1985, Imaging the Earth's Interior, Blackwell Scientific Publications.
- Constance, P., 2003, Vinton Salt Dome, current issues in structural imaging, 73rd Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2357-2361.
- Cornish, B., Minear, J., Kempner, W. and Nemeth, T., 2000, Processing of 3-D VSP data from a permanent borehole array, 70th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1755-1758.
- Dankbaar, J. W. M., 1987, Vertical seismic profiling - Separation of P-waves and S-waves, Geophys. Prosp., **35**, 803-814.
- Devaney, A. J. and Oristaglio, M. L., 1986, A plane-wave decomposition for elastic wave fields applied to the separation of P-waves and S-waves in vector seismic data (short note), Geophysics, **51**, 419-423.

- Diegel, F. A., Karlo, J. F., Schuster, J. C., Shoup, R. C., Tauvers, P. R., 1996, Cenozoic structural evolution and tectono-stratigraphic framework of the northern Gulf Coast continental margin, AAPG Memoir, **63**, 109-151.
- Esmersoy, C., 1990, Inversion of P and SV waves from multicomponent offset vertical seismic profiles, Geophysics, **55**, 39-50.
- Foster, D. J. and Gaiser, J. E., 1986, Elastic wave-field decomposition of offset VSP data, 56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 563-565.
- Hardage, B., 1983, Vertical Seismic Profiling, part A: Principles, Geophysical Press.
- Hestenes, M. R., 1973, Iterative Methods for Solving Linear Equations, Journal of Optimization Theory and Applications, **11**, 323-334.
- Kostov, C., 1990, Toeplitz structure in slant-stack inversion, 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1618-1621.
- Labonte, S., 1990, Modal separation, mapping, and inverting three-component VSP data, M.Sc. Thesis, The University of Calgary.
- Leaney, W., 2002, Anisotropic vector wavefield decomposition for 3D VSP data, 72th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2369-2372.
- Luo, M. and Evans, B., 2002, Imaging fractures and faults from 3D VSP(s), 72nd Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2389-2392.
- Marfurt, K. J., Schneider, R. V. and Mueller, M. C., 1996, Pitfalls of using conventional and discrete Radon transforms on poorly sampled data, Geophysics, **61**, 1467-1482.
- Menke, W., 1984, Geophysical data analysis: discrete inverse theory, Academic Press Inc., Orlando, Florida.

- Moon, W., Carswell, A., Tang, R. and Dilliston, C., 1986, Radon transform wave-field separation for vertical seismic profiling data, *Geophysics*, **51**, 940-947.
- Morozov, I. B., Bradley, C. J., Smithson, S. B., 1997, P and SV separation by polarization-dependant velocity filtering: application to vertical seismic profiles from Kola superdeep borehole, Russia, *Computers and Geosciences*, **23**, 1051-1061.
- Roche, S., Constance, P., Bloor, R., Bicquart, P., Bryans, B., Gelinsky, S., Holland, M. and Ralph, J. G., 1999, Simultaneous acquisition of 3-D surface seismic and 3-C, 3-D VSP data, 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 104-107
- Scales, J. A., 1987, Tomographic inversion via the conjugate gradient method, *Geophysics*, **52**, 179-185.
- Shewchuk, J. R., 1994, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, Technical report, School of Computer Science, Carnegie Mellon University, Vol. CMU-CS-94-125.
- Scott, W., Leaney, P. and Esmersoy, C., 1989, Parametric decomposition of offset VSP wave fields, 59th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 26-29.
- Seeman, B. and Horowicz, L., 1983, Vertical-seismic-profiling - Separation of upgoing and downgoing acoustic waves in a stratified medium, *Geophysics*, **48**, 555-568.
- Sheriff, R. E., 2002, *Encyclopedic Dictionary of Applied Geophysics*, Society of Exploration Geophysics.

- Stewart, R. and Disiena, J., 1989, The values of VSP in interpretation, *The Leading Edge*, **8**, 16-23.
- Stewart, R. R., Gulati, J. S., Zhang, Q., Parkin, J. M. and Peron, J., 1998, Analysing 3C-3-D VSP data: The Blackfoot, Alberta survey, 68th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 381-384.
- Stoer, J. and Bulirsch, R., 2002, *Intorduction to Numerical Analysis*, Springer-Verlag New York, Inc.
- Stoffa, P. L., Buhl, P., Diebold, J. B., Wanzel, F., 1981, Direct mapping of seismic data to the domain of intercept time and ray parameter - A plane wave decomposition, *Geophysics*, **46**, 255-267.
- Trad, D., Ulrych, T. J. and Sacchi, M. D., 2003, Latest views of the sparse Radon transform, *Geophysics*, **68**, 386-399.
- Turner, G., 1990, Aliasing in the tau- p transform and the removal of spatially aliased coherent noise, *Geophysics*, **55**, 1496-1503.
- Warren, A. D., 1957, The Anahuac and Frio sediments in Louisiana, *Gulf Coast Association of Geological Societies Transactions*, **7**, 221-237.
- Yilmaz, O. and Taner, M. T., 1994, Discrete plane-wave decomposition by least-mean-square-error method, *Geophysics*, **59**, 973-982.
- Yilmaz, O., 2001, *Seismic Data Analysis: Processing, Inversion and Interpretation of seismic Data*, Society of Exploration Geophysicist.
- Zhang, Q., Stewart, R. R. and Sun, Z., 1996, 3-D - VSP survey design and analysis, 66th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 190-193.

APPENDIX A: CONJUGATE GRADIENT METHOD

Many geophysical applications require the solution of large systems of linear equations, $\mathbf{Lm}=\mathbf{d}$, where \mathbf{m} is an unknown vector, \mathbf{d} is a known vector and \mathbf{L} is a known matrix. The simplest way to solve this system of equations is by solving the normal equations

$$\mathbf{m}=(\mathbf{L}^T\mathbf{L} + \varepsilon\mathbf{I})^{-1}\mathbf{L}^T\mathbf{d}, \quad (\text{A-1})$$

where \mathbf{L}^T is a transpose of known matrix \mathbf{L} , \mathbf{I} is an identity matrix and ε is a scalar introduced to ensure minimum energy solution (Menke, 1984). If the matrix \mathbf{L} is sparse, i.e. has only a relatively few nonzero elements, the classical elimination methods given by equation A-1 may not be efficient for solving the system because of formation of dense intermediate matrices $\mathbf{L}^T\mathbf{L}$, with number of operations too large even for present days computers. The memory for making these intermediate matrices can also be prohibitively large (Stoer and Bulirsch, 2002). *Iterative methods* are suitable for solving such system of equations. In the iterative method, we start with an initial estimate of the vector $\mathbf{m}^{(0)}$ and subsequently generate a sequence of vectors, $\mathbf{m}^{(0)} \rightarrow \mathbf{m}^{(1)} \rightarrow \mathbf{m}^{(2)} \rightarrow \dots$, which converge toward the desired solution, \mathbf{m} . In each iteration step we basically perform only matrix-vector multiplication which is a very modest computation when matrix \mathbf{L} is sparse. Therefore, we can do a large number of iterations with a relatively small effort. Iterative methods provide an acceptable approximate solution after a certain number of iterations without reaching the exact solution. Statements for the iterative methods also apply to Krylov space methods (Stoer and Bulirsch, 2002), except that these methods, in exact arithmetic, terminate with the exact solution \mathbf{x} after a finite number of

steps. A method of this type, which is the most prominent method for solving sparse systems of linear equations, is the conjugate gradient method (Hestenes, 1973).

The conjugate gradient method is an optimizer and simultaneous equation solver. Its iterations can be interrupted at any stage. The obtained partial result is an approximation that is often useful, while the exact solution is attained in a finite number of steps. In order to understand the conjugate gradient method, I will first review the method of steepest descent.

A.1. The method of steepest descent

For a simple, two-dimensional linear system (Figure A.1a), the solution lies at the intersection of two lines. If \mathbf{L} is a symmetric and positive-definite, the quadratic form

$$f(\mathbf{m}) = \frac{1}{2} \mathbf{m}^T \mathbf{L} \mathbf{m} - \mathbf{d}^T \mathbf{m} + c \quad (\text{A-2})$$

(Figure A.1b) is minimized by the solution to $\mathbf{Lm}=\mathbf{d}$ (Shewchuk, 1994). The minimum point of this surface is the solution to $\mathbf{Lm}=\mathbf{d}$ (Figure A.1c). The *gradient* of a quadratic form is a vector field that, for a given point \mathbf{m} , points in the direction of greatest increase of $f(\mathbf{m})$, and is orthogonal to the contour lines (Figure A.1d). At the bottom of the paraboloidal bowl, the gradient is zero.

In the method of steepest descent, we start at an arbitrary point $\mathbf{m}_{(0)}$. With a series of steps, we slide to the bottom of the paraboloid, i.e. until we are satisfied that we are close to the solution \mathbf{m} . We take steps in the direction in which $f(\mathbf{m})$ decreases most quickly. This direction is the opposite of the gradient direction. The *error* $\mathbf{e}_{(i)} = \mathbf{m}_{(i)} - \mathbf{m}$ is a vector that indicates how far we are from the solution. The *residual* $\mathbf{r}_{(i)} = \mathbf{d} - \mathbf{Lm}_{(i)}$ shows how far we are from the correct value of \mathbf{d} . The residual is the direction of steepest

descent. We begin along the direction of steepest descent at point $\mathbf{m}_{(0)}$. When we reach the point where we should 'climb' again, we should turn 90° , and continue on the new direction of steepest descent. But, the question is, where is that point? The point of our 'turn' is the minimum value of an intersecting parabola made between the vertical plane of the direction of steepest descent and paraboloid (Shewchuk, 1994). We define this point by the step length α :

$$\mathbf{m}_{(1)} = \mathbf{m}_{(0)} + \alpha \mathbf{r}_{(0)}. \quad (\text{A-3})$$

We call the procedure to find α at the base of parabola a *line search*. α should be chosen such that the residual and new gradient are orthogonal

$$\mathbf{r}_{(1)}^T \mathbf{r}_{(0)} = 0 \quad (\text{A-4})$$

$$(\mathbf{d} - \mathbf{L} \mathbf{m}_{(1)})^T \mathbf{r}_{(0)} = 0 \quad (\text{A-5})$$

$$(\mathbf{d} - \mathbf{L} (\mathbf{m}_{(0)} + \alpha \mathbf{r}_{(0)}))^T \mathbf{r}_{(0)} = 0 \quad (\text{A-6})$$

$$(\mathbf{d} - \mathbf{L} \mathbf{m}_{(0)})^T \mathbf{r}_{(0)} + \alpha (\mathbf{L} \mathbf{r}_{(0)})^T \mathbf{r}_{(0)} = 0 \quad (\text{A-7})$$

$$(\mathbf{d} - \mathbf{L} \mathbf{m}_{(0)})^T \mathbf{r}_{(0)} = \alpha (\mathbf{L} \mathbf{r}_{(0)})^T \mathbf{r}_{(0)} \quad (\text{A-8})$$

$$\mathbf{r}_{(0)}^T \mathbf{r}_{(0)} = \alpha \mathbf{r}_{(0)}^T (\mathbf{L} \mathbf{r}_{(0)}) \quad (\text{A-9})$$

$$\alpha = \frac{\mathbf{r}_{(0)}^T \mathbf{r}_{(0)}}{\mathbf{r}_{(0)}^T \mathbf{L} \mathbf{r}_{(0)}}. \quad (\text{A-10})$$

The algorithm above requires two matrix-vector multiplications per iteration. In our Radon transform this will be two subroutine calls - one to the forward transform and one to the (non-orthogonal) reverse transform. One of these multiplications can be eliminated by premultiplying both sides of equation (A-3) by \mathbf{L} and adding \mathbf{d}

$$\mathbf{r}_{(i+1)} = \mathbf{r}_{(i)} - \alpha_{(i)} \mathbf{L} \mathbf{r}_{(i)}. \quad (\text{A-11})$$

After calculating the first residual $\mathbf{r}_{(1)}$, and product \mathbf{Lr} , the equation (A-11) can be used for every iteration thereafter. The result is shown in Figure A.2.

A.2. The Method of Conjugate Gradients

Steepest descent often takes the new direction to be the same as the previous one. A better idea is to search for the right direction after one step. Thus we will be done after exactly n steps. The solution is to make search directions $\mathbf{h}_{(i)}$ \mathbf{L} -orthogonal, or *conjugate* $\mathbf{h}_{(i)}^T \mathbf{Lh}_{(j)} = 0$ (Figure A.3). When the search directions are conjugate, α is

$$\alpha_{(i)} = \frac{\mathbf{h}_{(i)}^T \mathbf{r}_{(i)}}{\mathbf{h}_{(i)}^T \mathbf{Lh}_{(i)}}. \quad (\text{A-12})$$

In order to find a set of \mathbf{L} -orthogonal search directions $\mathbf{h}_{(i)}$, we will introduce a set of n linearly independent vectors $\mathbf{g}_{(i)}$, such that $\mathbf{h}_{(0)} = \mathbf{g}_{(0)}$ and $\mathbf{h}_{(i)} = \mathbf{g}_i + \sum_{k=0}^{i-1} \beta_{ik} \mathbf{h}_{(k)}$. The values for β are given by

$$\beta_{ij} = -\frac{\mathbf{g}_i^T \mathbf{Lh}_{(j)}}{\mathbf{h}_{(j)}^T \mathbf{Lh}_{(j)}}. \quad (\text{A-13})$$

In the conjugate gradient method, the search directions are constructed by conjugation of the residuals. The residual is orthogonal to the previous search directions, so they will always produce a new, linearly independent search direction. The classic conjugate gradient algorithm for solving $\mathbf{Lm}=\mathbf{d}$, where \mathbf{L} is an $n \times n$, symmetric, positive definite matrix is as follows (Figure A.4):

$$\mathbf{h}_{(0)} = \mathbf{r}_{(0)} = \mathbf{d} - \mathbf{L} \mathbf{m}_{(0)}, \quad (\text{A-14})$$

$$\alpha_{(i+1)} = \frac{\mathbf{r}_{(i)}^T \mathbf{r}_{(i)}}{\mathbf{h}_{(i)}^T \mathbf{Lh}_{(i)}}, \quad (\text{A-15})$$

$$\mathbf{m}_{(i+1)} = \mathbf{m}_{(i)} + \alpha_{(i+1)} \mathbf{h}_{(i)} , \quad (\text{A-16})$$

$$\mathbf{r}_{(i+1)} = \mathbf{r}_{(i)} - \alpha_{(i+1)} \mathbf{Lh}_{(i)} , \quad (\text{A-17})$$

$$\beta_{i+1} = \frac{\mathbf{r}_{(i+1)}^T \mathbf{r}_{(i+1)}}{\mathbf{r}_{(i)}^T \mathbf{r}_{(i)}} , \text{ and} \quad (\text{A-18})$$

$$\mathbf{h}_{(i+1)} = \mathbf{r}_{(i+1)} + \beta_{(i+1)} \mathbf{h}_{(i)} . \quad (\text{A-19})$$

If the matrix \mathbf{L} is an $m \times n$ matrix (Hestenes, 1973), the conjugate gradient method gives the following generalization for the normal equation $\mathbf{L}^T \mathbf{Lm} = \mathbf{L}^T \mathbf{d}$:

$$\mathbf{r}_{(0)} = \mathbf{d} - \mathbf{L} \mathbf{m}_{(0)} , \quad (\text{A-20})$$

$$\mathbf{g}_{(0)} = \mathbf{h}_{(0)} = \mathbf{L}^T (\mathbf{d} - \mathbf{L} \mathbf{m}_{(0)}) , \quad (\text{A-21})$$

$$\Delta \mathbf{d} = \mathbf{Lh}_{(0)} \quad (\text{A-22})$$

$$\alpha_{(i+1)} = \frac{\mathbf{g}_{(i)}^T \mathbf{g}_{(i)}}{\Delta \mathbf{d}_{(i)}^T \Delta \mathbf{d}_{(i)}} , \quad (\text{A-23})$$

$$\mathbf{m}_{(i+1)} = \mathbf{m}_{(i)} + \alpha_{(i+1)} \mathbf{h}_{(i)} , \quad (\text{A-24})$$

$$\mathbf{r}_{(i+1)} = \mathbf{r}_{(i)} - \alpha_{(i+1)} \Delta \mathbf{d}_{(i)} , \quad (\text{A-25})$$

$$\beta_{i+1} = \frac{\mathbf{g}_{(i+1)}^T \mathbf{g}_{(i+1)}}{\mathbf{g}_{(i)}^T \mathbf{g}_{(i)}} , \quad (\text{A-26})$$

$$\mathbf{h}_{(i+1)} = \mathbf{g}_{(i+1)} + \beta_{(i+1)} \mathbf{h}_{(i)} , \text{ and} \quad (\text{A-27})$$

$$\Delta \mathbf{d}_{(i+1)} = \mathbf{Lh}_{(i+1)} . \quad (\text{A-28})$$

The idea is to calculate the residual $\mathbf{r}_{(0)}$ (A-20) and then multiply by \mathbf{L}^T rather than subtracting $\mathbf{L}^T \mathbf{Lm}$ from $\mathbf{L}^T \mathbf{d}$ (Scales, 1987).

To summarize, the conjugate gradient method is effective because:

a) The matrix appears only through the matrix-vector product, and this can be implemented as a subroutine, and

b) The search vector \mathbf{h} is calculated recursively and not stored (Scales, 1987).

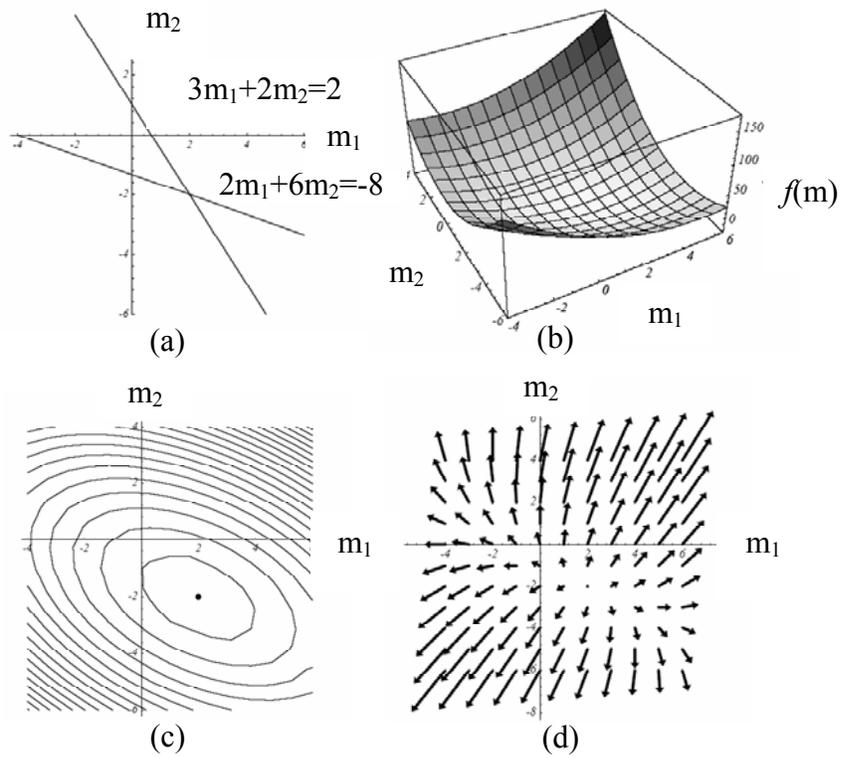


FIG. A.1. (a) Two-dimensional linear system, (b) Graph, (c) Contours, and (d) Gradient of the quadratic form given by equation A-2 (after Shewchuk, 1994).

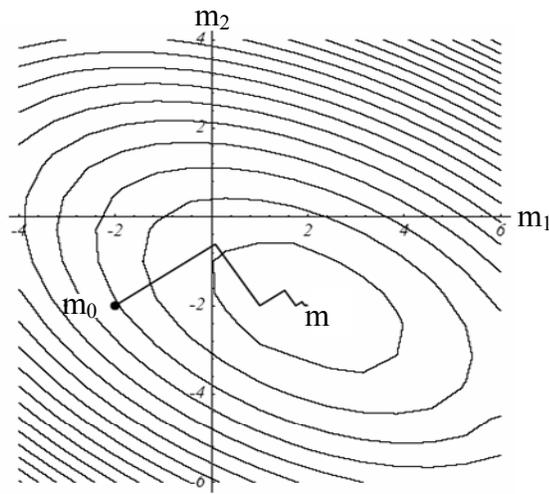


FIG. A.2. The method of Steepest Descent (after Shewchuk, 1994).

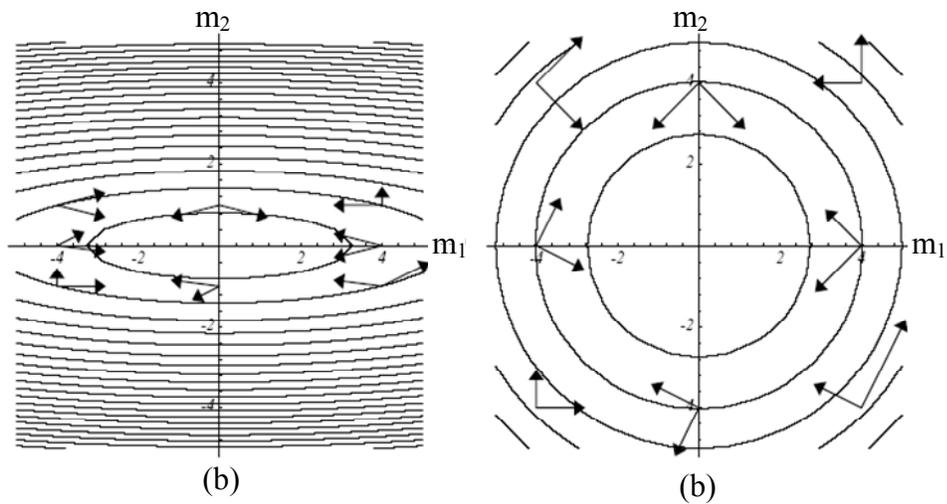


FIG. A.3. (a) Pairs of orthogonal vectors. (b) Pairs of L-orthogonal or conjugate vectors (after Shewchuk, 1994).

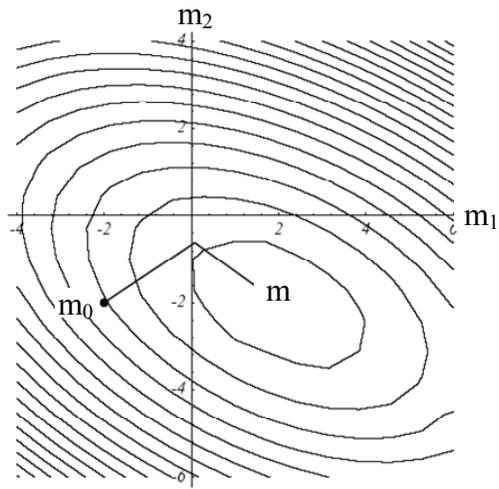


FIG. A.4. The method of conjugate gradients converges in two steps for a 2-D problem (after Shewchuk, 1994).