Important Notice

UNIVERSITY OF CALGARY

Seismic parameter estimation from AVO inversion

by

Jonathan E. Downton

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOLOGY AND GEOPHYSICS

CALGARY, ALBERTA
JANUARY, 2005

UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Seismic parameter estimation from AVO inversion" submitted by Jonathan E. Downton in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Supervisor, Dr. Laurence R. Lines
Department of Geology and Geophysics

Dr. John Bancroft
Department of Geology and Geophysics

Dr. Edward S. Krebes
Department of Geology and Geophysics

Dr. Michael Lamoureux
Department of Mathematics and Statistics

Dr. Robert R. Stewart
Department of Geology and Geophysics

External Examiner Dr. Mauricio D. Sacchi
University of Alberta, Department of Physics

Date

Abstract

AVO inversion is a well established seismic exploration methodology to predict the earth's elastic parameters and thus rock and fluid properties. This thesis introduces a series of theoretical improvements to the AVO inverse problem. These include, a practical methodology to perform three-term AVO inversion instead of the industry standard of performing two-term AVO inversion. Two-term AVO inversion constrains the inversion either implicitly by truncating a term, or explicitly through some empirical relationship, thus introducing bias into the estimate. Using probabilistic constraints based on local geologic information, the three-term inversion can be constrained so as to give stable estimates, but with minimal bias. To help the explorationist understand the reliability of these estimates, a series of quality controls are developed and verified.

In order to estimate reliable density reflectivity, long offsets are required. NMO stretch and offset-dependent tuning introduce problematic distortions at these offsets. Two methods are investigated to address these issues. First, stretch-free NMO is investigated as a way to precondition the data. Secondly, an AVO waveform inversion which incorporates into the forward model NMO stretch and offset-dependent tuning is considered.

Both the AVO inversion and AVO waveform inversion are developed using a Bayesian framework. Gaussian and long-tailed distributions are explored for the likelihood and *a priori* probability distributions. The optimization problem results in non-linear solutions. The AVO inversion is solved using Newton-Raphson while the AVO waveform inversion is solved using conjugate gradient. In each case, the methodology is demonstrated on synthetic and real data examples.

The synthetic example shows that the AVO waveform inversion provides the most accurate estimates in the presence of NMO stretch and tuning. Preconditioning the data with stretch-free NMO improves the results of the AVO inversion, but not as much as the AVO waveform inversion results. For real seismic data the AVO waveform inversion also provided the best results of the methods tested. An example is shown where both the AVO inversion and AVO waveform inversion are able to estimate density reflectivity that differentiate commercial from noncommercial gas.

Acknowledgements

I would like to thank Larry Lines, my dissertation supervisor, for his guidance and many useful suggestions. Larry's vast knowledge of the geophysical literature, specifically about inverse theory and keen intuitive understanding of geophysics, helped me throughout my graduate studies. Larry became the nucleus of a group interested in AVO, inversion and the relationship between rock and seismic properties. He attracted a talented group of individuals including fellow graduate students, Andrew Royle, Brian Russell and Ian Watson. I learned much through our discussions and interactions.

The CREWES environment provided a friendly atmosphere to learn about a wide range of topics in geophysics and present ideas. From this group, I made a number of good friendships including Chuck Ursenbach, Richard Bale and Hugh Geiger. I would also like to thank a number of Professors that I learned so much from, Ed Krebes from whom I learned a great deal about seismology, Gary Margrave from whom I learned about mathematical geophysics and migration, and Doug Schmitt from whom I learned about rock physics. In addition I would like to thank Rob Stewart, John Bancroft, and Michael Slawinski for teaching interesting, accessible courses that got me interested in taking further graduate courses in geophysics after leaving university for ten years.

Before returning to university, to do my dissertation, I had been in the oil industry for fourteen years. A number of key individuals shaped my ideas and interests. Foremost among them were Shlomo Levy and Kerry Stinson. While working for ITA, I was given the opportunity to investigate poststack impedance inversion, AVO, and the relationship between rock and seismic properties. This is where my interest in stratigraphic inversion was allowed to blossom. Later, upon forming our own company, Integra, they provided excellent council and guidance.

Within Integra and later CoreLab RTD, I was privileged to work with a number of talented individuals. We learned much from each other and grew together. Among these were Yongyi Li, Jan Dewar, Bob Somerville, Jim Howell, Don Daub, Yong Xu and Glen Brewer. In addition, I would like to acknowledge Huimin Guan for coding up the Stretch-free NMO algorithm in Fortran and J.R. Hostetler for creating the real Stretch-free NMO data example. I learned much from our clients and grew with our collaborations. Foremost among these were Bill Goodway, and Dave Mackidd. Others who have helped my understanding of stratigraphic inversion include Lee

Hunt, Rainer Tonn, Holger Mandler and Weimin Zhang.

I would also like to thank Mark Klingbeil who was the President of Integra. I learned much about running a company from him and personally grew through my relationship with him. He also was responsible for setting up the flexible employment agreement that allowed me the financial resources to pursue my graduate studies.

I am most indebted to Larry Lines, Pat Daley, Jan Dewar, and Chuck Ursenbach for proof reading this thesis. I thank them for their suggestions and all the work this entailed.

Finally, I am thankful to my family for persevering with me and supporting me as I went down this path for the last five years. I thank my wife, Beatrix, who has put up with the time pressures of me working and going to university for so long. My daughters, Annika and Zoe, have grown up with me being in university for most of their lives. I look forward to spending more time with my family upon the completion of this thesis.

This work is dedicated to the memory of my father, William H. Downton.

Contents

List of Tables

List of Figures

List of Symbols

Throughout this thesis, I use the convention that upper case bold face lettering designates a matrix, lower case bold face lettering designates a vector and a scalar is italicized. Estimates are denoted with the symbol ˆ over the variable and averages with a bar over them. Summarized below, for the readers convenience are the symbols used in this thesis.

A, B, C reflectivity attributes generated by Shuey equation

b intercept used in mudrock relationship $\alpha = m\beta + b$

b_I intercept used in modified mudrock relationship $I_p = m_I I_s + b_I$

$\mathbf{C_d}$ parameter covariance matrix for the data $\mathbf{d}$

$\mathbf{C_x}$ parameter covariance matrix for variable $\mathbf{x}$

$\tilde{\mathbf{C}}_\mathbf{x}$ normalized parameter covariance matrix for variable $\mathbf{x}$

$\hat{\mathbf{C}}_{\hat{\mathbf{x}}}$ estimate of parameter covariance matrix for variable $\hat{\mathbf{x}}$

$\mathbf{d}$ data

$\mathbf{D}$ data vector with replicates

f Gardner-like coefficient used in density, S-wave velocity relationship

$FE\,(R)$ Fractional error of reflectivity attribute R

g Gardner coefficient used in density, P-wave velocity relationship

$\mathbf{E}$ Diagonal matrix containing weights used in AVO equation

$\mathbf{F}$ Diagonal matrix containing weights used in AVO equation

$\mathbf{G}$ Linear operator used to model data

$\mathbf{H}$ Diagonal matrix containing weights used in AVO equation

h offset

$\mathbf{I}$ identity matrix

I_p P-wave impedance

I_s S-wave impedance

J number of point impedance constraints

K number of offsets

ℓ^1 L1 norm

ℓ^2 L2 norm

ℓ^p LP norm

L number of time samples

m slope used in mudrock relationship $\alpha = m\beta + b$

m_I slope used in modified mudrock relationship $I_p = m_I I_s + b_I$

M number of data points

$\mathbf{M}$ mask operator

N number of unknown parameters

$\mathbf{n}$ random noise vector

$\mathbf{N}$ NMO operator

p ray parameter (horizontal slowness)

$P(\mathbf{x}|I)$ probability of $\mathbf{x}$ given some information I

$\mathbf{Q}$ diagonal weighting matrix used in regularization

$\mathbf{r}_d$ density reflectivity vector (element is $R_d = \frac{1}{2}\frac{\Delta\rho}{\rho}$)

$\mathbf{r}_f$ fluid stack based on velocity reflectivity

$\mathbf{r}_g$ fluid stack based on impedance reflectivity

$\mathbf{r}_K$ Bulk modulus reflectivity vector (element is $R_K = \frac{1}{2}\frac{\Delta K}{\bar{K}}$)

$\mathbf{r}_p$ P-wave impedance reflectivity vector (element is $R_p = \frac{1}{2}\frac{\Delta I_P}{\bar{I}_P}$)

$\mathbf{r}_s$ S-wave impedance reflectivity vector (element is $R_s = \frac{1}{2}\frac{\Delta I_S}{\bar{I}_S}$)

$\tilde{\mathbf{r}}_s$ scale S-wave impedance reflectivity vector (element is $\tilde{R}_s = \bar{\gamma}\frac{1}{2}\frac{\Delta I_S}{\bar{I}_S}$)

$\mathbf{r}_\alpha$ P-wave velocity reflectivity vector (element is $R_a = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}}$)

$\mathbf{r}_\beta$ S-wave velocity reflectivity vector (element is $R_\beta = \frac{1}{2}\frac{\Delta\beta}{\bar{\beta}}$)

$\mathbf{r}'_\beta$ scaled S-wave velocity reflectivity vector (element is $R'_\beta = \bar{\gamma}\frac{1}{2}\frac{\Delta\beta}{\bar{\beta}}$)

$\mathbf{r}_\lambda$ λ reflectivity vector (element is $R_\lambda = \frac{1}{2}\frac{\Delta\lambda}{\bar{\lambda}}$)

$\mathbf{r}_\mu$ Rigidity reflectivity vector (element is $R_\mu = \frac{1}{2}\frac{\Delta\mu}{\bar{\mu}}$)

$\mathbf{r}_{\lambda\rho}$ $\lambda\rho$ reflectivity vector (element is $R_\lambda = \frac{1}{2}\frac{\Delta(\lambda\rho)}{\lambda\bar{\rho}}$)

$\mathbf{r}_{\mu\rho}$ $\mu\rho$ reflectivity vector (element is $R_\mu = \frac{1}{2}\frac{\Delta(\mu\rho)}{\mu\bar{\rho}}$)

$\mathbf{r}_\gamma$ β/α ratio reflectivity vector (element is $R_\gamma = \frac{1}{2}\frac{\Delta\gamma}{\bar{\gamma}}$)

$\mathbf{r}_\sigma$ pseudo Poisson ratio reflectivity vector (element is $R_\sigma = \frac{\Delta\nu}{2(1-\bar{\nu})^2}$)

$\mathbf{r}_\nu$ Poisson ratio reflectivity vector (element is $R_\nu = \frac{1}{2}\frac{\Delta\nu}{\nu}$)

$r_{R_\alpha R_\beta}$ correlation coefficient of $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$

$R(\theta)$ Angle-dependent reflectivity

$\mathbf{S}$ spray operator

$\mathbf{T}$ transform matrix

t_0 zero offset two-way travel time

V_{stak} stacking velocity

w weights used to construct weighting matrix

$\mathbf{W}$ convolution matrix containing wavelet

$\mathbf{W}_e$ weighting matrix that weights misfit function

$\mathbf{W_x}$ parameter weighting matrix

$\mathbf{V}$ eigenvectors of $\mathbf{C}_x$

$\mathbf{x}$ parameter vector

$\mathbf{x}'$ transformed parameter vector

$\mathbf{X}$ parameter vector with replicates

z depth

$\mathbf{0}$ Block matrix filled with zeros (size dependent on context)

α P-wave velocity

α_{back} smooth background P-wave velocity used for ray tracing

β S-wave velocity

ϵ threshold for ℓ^1 and Huber norm

$\boldsymbol{\varepsilon}$ misfit vector (whose elements are ε_k)

$\mathbf{\Sigma}$ eigenvalues of $\mathbf{C}_x$ (diagonal of $\mathbf{\Sigma}$ are eigenvalues σ_i^2)

λ Lamé parameter λ

μ rigidity

Θ influence function

γ S-wave to P-wave velocity ratio

ρ density (d used as subscript for reflectivity to indicate density)

θ average angle of incidence

ν Poisson ratio

τ_k offset-dependent two-way travel time for the k^{th} offset

σ_{R_α} variance of $\mathbf{r}_\alpha$

$\sigma_{R_\alpha R_d}$ covariance of $\mathbf{r}_\alpha$ and $\mathbf{r}_d$

σ_N^2 variance of the noise

Chapter 1

Introduction

1.1 Amplitude variations with offset (AVO)

Compressional wave amplitude variations with offset (AVO) provides the explorationist with additional information over conventional stacked seismic data. In the 1960's, geophysicists discovered that gas deposits were often associated with amplitude anomalies on stacked sections known as "bright spots". However, these bright spots proved to be less than a perfect gas indicator, since other large impedance contrasts due to rapid changes in lithology can also cause them. Examples of this are igneous intrusions, carbonates and coals in clastic sequences. AVO was introduced as a methodology to provide additional information so as to more accurately predict the rock and fluid properties of the earth. For example, Ostrander (1984), Fatti et al. (1994), and Allen and Peddy (1993) published case histories showing AVO's usefulness in predicting and mapping hydrocarbons. Verm and Hilterman (1995), Goodway et al. (1997), and Avseth (2000) used it to distinguish lithology and map porosity in clastics. Further, Li et al. (2003) demonstrated its usefulness in exploring for carbonates. AVO has also proven useful in planning drilling programs by performing pore pressure prediction (Carcione and Gangi, 2000) and characterizing the reservoir in time-lapse seismic monitoring studies (Tura and Lumley, 1999; and Landro, 2001).

Compressional wave AVO provides additional information without extra acquisition costs unlike converted or shear wave acquisition. The extra cost associated with AVO are (1) the cost of processing and inverting the prestack seismic data properly, and (2) incorporating this extra information into the interpretation. In practice both these considerations limit the use of AVO. The seismic data are noisy and the AVO

problem ill-conditioned resulting in elastic parameter estimates with potentially large and unknown errors. Further, rock and fluid properties estimated from the AVO estimates are ambiguous due to the non-uniqueness of the elastic to rock property transform. Thus, interpreting the significance of the AVO inversion results can be daunting. For these reasons AVO tends to get used on a subset of plays where it has worked in the past or has the potential to clear up ambiguous poststack interpretations. This means that for many plays, we are throwing away information that might be useful in describing the reservoir and finding hydrocarbons. This seems a waste.

To address the first of these issues, namely the issue of reliably estimating AVO attributes, this thesis proposes a series of improvements. Robust statistics and probabilistic constraints based on geologic control are incorporated into the inversion using Bayes theorem resulting in more reliable estimates in the presence of noise and sub-optimal acquisition geometries. For isotropic elastic media the methodology estimates three independent attributes. Various quality controls are developed and illustrated on both synthetic and real data. These quality controls help the explorationist evaluate the reliability and the information content actually in the AVO attributes. For noisy data, the constraints dominate the solution, particularly the density estimate, providing little independent information.

In order to obtain reliable density results, large offsets are required. Incorporating these large offsets introduces its own problems, including NMO stretch and offset-dependent tuning. These distortions and their effects are described and then several methods for dealing with them are developed and demonstrated, including a method to do stretch-free NMO and AVO waveform inversion. The results of the AVO waveform inversion may be output in terms of elastic parameters or their perturbations.

Throughout this work it is assumed that the earth is composed of a series of flat, homogenous, elastic, isotropic layers. The assumption of horizontal layers, and laterally invariant velocities may be relaxed somewhat if the data are preconditioned by prestack migration. Only compressional wave reflectivity is considered. Transmission losses, converted waves, and multiples are not incorporated in this model and so must be addressed through prior processing.

Both AVO and AVO waveform inversion are a subset and simple forms of elastic inversion. In its most general form elastic inversion incorporates all multiples, modes

of propagation including shear and converted waves, the reflectivity model based on a propagator matrix methods (Kennett, 1984) or equivalently the Zoeppritz equations, and makes use of both the dynamic and kinematic information in the seismic data. The model is nonlinear and the solution space non-unique. This non-uniqueness is a result of the band-limited nature of the seismic data and the ill-conditioned nature of AVO problem (Debski and Tarantola, 1995). Further complicating the analysis is the fact that both the reflectivity and traveltime inversion problems are largely decoupled (Claerbout, 1986) and non-unique (Spagnolini, 1994). Elastic inversion is typically solved using nonlinear inversion techniques. Examples of elastic inversion include Tarantola (1986), Crase et al. (1990), Sen and Stoffa (1995), and Gouveia and Scales (1998). The latter two techniques solve the problem recursively in depth and are based on propagator matrix theory that assumes the velocity structure of the earth varies only in the vertical direction.

Despite the maturity of these papers (Tarantola, 1986; Stoffa, 1995; Gouveia and Scales, 1998), and advances in computing power over recent years, elastic inversion is still not commercialized on a wide basis. This is partly due to the wide scope of problems being addressed by these algorithms and the complications this entails. Because of this, various authors have tried to simplify and solve the inversion as a series of smaller, separate, and easier problems with more approximations. Often this requires that the seismic data be preprocessed to satisfy the assumptions of the algorithm. The results are easier to quality control, with the processor having a chance to adjust parameters before going on to the next step. This approach is less of a black box. For example, Carazzone and Srnka (1993) simplify the inversion problem by ignoring converted waves, allowing them to ignore registration issues between the P-wave and S-wave reflectors. Grion et al. (1998) ignore multiples and converted waves, but still perform a joint traveltime and reflectivity inversion. Simmons and Backus (1996) perform AVO waveform inversion using only the reflectivity information, having explicitly defined the traveltime relationships. Similarly, Buland and Omre (2003) perform AVO waveform inversion on NMO corrected gathers using a linearized approximation. AVO inversion (Wiggins et al., 1983; Smith and Gidlow, 1987; and Lortzer and Berkhout, 1993) ignores the band-limited nature of the seismic data and is performed on NMO corrected or prestack migrated gathers. The resulting band-limited reflectivity estimates may be inverted to obtain their related elastic parameters using poststack impedance inversion (Lindseth, 1979; Oldenburg et al.,

1983; and van Riel and Berkhout, 1985). AVO and poststack impedance inversion are the most common methods for performing elastic inversion in the geophysical industry today.

The ultimate objective of the explorationist is to determine rock and fluid properties. Elastic parameter and density estimates are an intermediate step towards obtaining rock and fluid properties; a secondary inversion is required to get from elastic and density to rock and fluid properties. Discussing this secondary inversion in detail is beyond the scope of this thesis. Interested readers are referred to work by Avseth (2000), Takahashi (2000), and Loures and Moraes (2002) to name but a few. The major issues with this estimation problem are quite different than those of the elastic inversion problem, hence it is best to treat this second inversion as a separate problem. The models used to map the elastic parameters to rock and fluid properties are based on a wide range of theoretical and empirical models (Mavko et al., 1998) that must be used in their proper geological context. Since these relationships are empirical, or in the case of the theoretical models, highly idealized, the role of theoretical error needs to considered in the accuracy of the estimates. There is also a major issue with nonuniqueness since the number of rock and fluid properties that control a reservoir's elastic parameters vastly exceed the number of elastic parameters estimated by the AVO inversion. In order to make the problem better posed, certain simplifying assumptions based on the geologic context must be made. Further, complicating the analysis is the fact there are scale issues due to the different frequency ranges used to measure the core, well, and seismic data and effective media concepts must be used. Despite the fact this thesis does not examine the rock property inversion in any detail, certain basic relationships and interpretation strategies are briefly reviewed in this chapter, in order to give the reader the basic tools to be able to judge how well the elastic estimates describe the reservoir and what additional information they provide over a conventional interpretation. Further, geologic information, such as well logs, may be used to help constrain the elastic inversion. With this in mind, this chapter reviews a number of basic relationships linking the rock and seismic properties.

This chapter reviews and lays the theoretical foundation for the rest of the thesis. The chapter starts by reviewing the Zoeppritz equations and showing how these are the basis of AVO. Further, it is shown that AVO inversion is a simplified form of elastic inversion. A linearized approximation of the Zoeppritz equations is then introduced

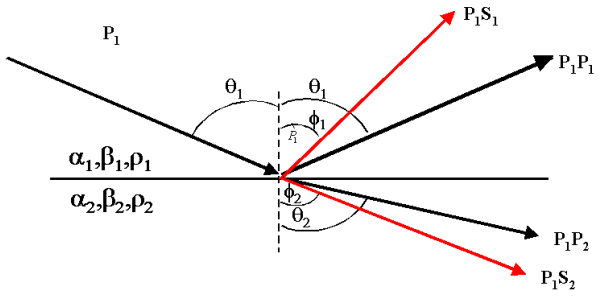

Figure 1.1: Reflected and transmitted rays for an incident P-wave on an interface between two elastic isotropic halfspaces.

and justified. The assumptions behind the Zoeppritz equations are idealizations and as such preprocessing must be done to the seismic data to attempt to satisfy these assumptions. These assumptions and processes are described in the next section. Then several empirical rock property relationships and interpretation methodologies are described, illustrated by several examples. The chapter closes by summarizing the contributions of this thesis to AVO inversion.

1.2 AVO theory

1.2.1 Zoeppritz equations

For an isolated interface the Zoeppritz equations (Aki and Richards, 1980) predict that the reflected amplitude changes as a function of angle of incidence. This is the basis of AVO (Amplitude Variations with Offset) or more properly AVA (Amplitude Variations with Angle). Based on how the amplitude changes as a function of angle, various inferences about the elastic parameters may be made. The Zoeppritz equations are derived for the idealized situation of two half-spaces in welded contact. In this thesis it is assumed that these half-spaces are elastic and in particular isotropic but this is not a necessary assumption. See for example, Daley and Hron

(1977), Thomsen (1988), Ruger (2002) and Tsvankin (2001) for AVO relationships for polar anisotropy and beyond. Isotropy is assumed in this thesis since it leads to a linearized AVO inversion which is overdetermined, which is not the case for VTI anisotropy. Section 1.3 discusses in greater detail some of the other assumptions that must be made, and the preconditioning that must be done to the seismic data, in order to apply the Zoeppritz equations for AVO analysis.

The upper and lower half-space are parameterized by their density, P-wave and S-wave velocities ρ, α, and β respectively, with the subscript 1 referring to the upper half-space and subscript 2 the lower half-space. For an incident downgoing plane P- or SV-wave upon an interface, the Zoeppritz equations predict that there will be two reflected upgoing plane waves and two downgoing plane waves (Figure 1.1). For P-SV waves, Aki and Richards (1980) write the Zoeppritz equations in matrix form as

$$\begin{bmatrix} \grave{P}_1\acute{P}_1 & \grave{S}_1\acute{P}_1 & \acute{P}_2\acute{P}_1 & \acute{S}_2\acute{P}_1 \\ \grave{P}_1\acute{S}_1 & \grave{S}_1\acute{S}_1 & \acute{P}_2\acute{S}_1 & \acute{S}_2\acute{S}_1 \\ \grave{P}_1\grave{P}_2 & \grave{S}_1\grave{P}_2 & \acute{P}_2\grave{P}_2 & \acute{S}_2\grave{P}_2 \\ \grave{P}_1\grave{S}_2 & \grave{S}_1\grave{S}_2 & \acute{P}_2\grave{S}_2 & \acute{S}_2\grave{S}_2 \end{bmatrix} = \mathbf{M}^{-1}\mathbf{N}, \tag{1.1}$$

where

$$\mathbf{M} = \begin{bmatrix} -\alpha_1 p & -\sqrt{1-p^2\beta_1^2} & \alpha_2 p & \sqrt{1-p^2\beta_2^2} \\ \sqrt{1-p^2\alpha_1^2} & -\beta_1 p & \sqrt{1-p^2\alpha_2^2} & -\beta_2 p \\ 2\rho_1\beta_1^2 p\sqrt{1-p^2\alpha_1^2} & \rho_1\beta_1\left(1-2p^2\beta_1^2\right) & 2\rho_2\beta_2^2 p\sqrt{1-p^2\alpha_2^2} & \rho_2\beta_2\left(1-2p^2\beta_2^2\right) \\ -\rho_1\alpha_1\left(1-2\beta_1^2p^2\right) & 2\rho_1\beta_1^2 p\sqrt{1-p^2\beta_1^2} & \rho_2\alpha_2\left(1-2\beta_2^2p^2\right) & -2\rho_2\beta_2^2 p\sqrt{1-p^2\beta_2^2} \end{bmatrix}, \tag{1.2}$$

and

$$\mathbf{N} = \begin{bmatrix} \alpha_1 p & \sqrt{1-p^2\beta_1^2} & -\alpha_2 p & -\sqrt{1-p^2\beta_2^2} \\ \sqrt{1-p^2\alpha_1^2} & -\beta_1 p & \sqrt{1-p^2\alpha_2^2} & -\beta_2 p \\ 2\rho_1\beta_1^2 p\sqrt{1-p^2\alpha_1^2} & \rho_1\beta_1\left(1-2\beta_1^2p^2\right) & 2\rho_2\beta_2^2 p\sqrt{1-p^2\alpha_2^2} & \rho_2\beta_2\left(1-2\beta_2^2p^2\right) \\ \rho_1\alpha_1\left(1-2\beta_1^2p^2\right) & -2\rho_1\beta_1^2 p\sqrt{1-p^2\beta_1^2} & -\rho_2\alpha_2\left(1-2\beta_2^2p^2\right) & 2\rho_2\beta_2^2 p\sqrt{1-p^2\beta_2^2} \end{bmatrix}. \tag{1.3}$$

The notation for the matrix in equation (1.1) is quite general allowing for both upgoing and downgoing incident P- and SV- waves. The letter indicates whether it is a P-wave or SV-wave, with the accent mark indicating whether it is an upgoing wave (e.g. $\acute{P}_1$) or a downgoing wave (e.g. $\grave{P}_1$), while the subscript indicates which

layer the wave is traveling in. So, for example, $\grave{P}_1\acute{P}_1$ is the reflected relative amplitude of the upgoing P-wave from an incident downgoing P-wave in layer 1 while $\grave{S}_1\grave{S}_2$ is the transmitted downgoing SV-wave traveling in layer 2 from an incident downgoing SV-wave in layer 1. The reflection and transmission coefficients are functions of the horizontal slowness p, and the layer properties.

As indicated in the introduction, this thesis only considers $\grave{P}\acute{P}$ reflections. Both $\grave{P}\acute{S}$ and $\grave{S}\acute{S}$ AVO inversion are conceptually similar to $\grave{P}\acute{P}$ AVO inversion. The interested reader is referred to Lortzer and Berkhout (1993), and Gray (2003) for case histories demonstrating converted wave $\grave{P}\acute{S}$ AVO. Larsen et al. (1999) demonstrate a joint inversion for $\grave{P}\acute{P}$ and $\grave{P}\acute{S}$ data. Shear sources and $\grave{S}\acute{S}$ AVO are generally not considered due to the poor signal-to-noise ratio of the recorded shear wave data. The AVO waveform inversion demonstrated in this thesis may be extended to perform $\grave{P}\acute{P}$ and $\grave{P}\acute{S}$ joint inversion. In practice difficulties arise due to the problem of estimating unique waveforms for both the $\grave{P}\acute{P}$ and $\grave{P}\acute{S}$ data, and event registration between the $\grave{P}\acute{P}$ and $\grave{P}\acute{S}$ data.

If horizontal layering and lateral velocity invariance are assumed, it is possible to generalize this analysis to many layers and interfaces using propagator matrices (Claerbout, 1973; Ursin, 1983; and Kennett, 1984). The parameterization of these layers is somewhat arbitrary in that the velocity is scale dependent. Due to the limited resolution of the seismic measurement, effective media concepts (Backus, 1962) must be used. These layers may be parameterized in terms of zero offset travel time if it is assumed that each layer's thickness has the constant two way travel time (Goupillaud, 1961)

$$\Delta t_0 = \frac{\Delta z_n}{\alpha_{eff_n}}, \tag{1.4}$$

where the thickness of the n^{th} layer Δz_n is chosen appropriately to compensate for changes in the effective velocity α_{eff_n} of the n^{th} layer. This may be done since α_{eff_n} changes slowly on a sample by sample basis.

Both Sen and Stoffa (1995), and Gouveia and Scales (1998) use propagator matrices as the starting point for their elastic inversions. This form includes the total wavefield response, including conversions, and all surface and intrabed multiples. In addition, modifications are introduced to incorporate surface boundary conditions, source, and receiver responses. The inversion problem is multi-modal and highly nonlinear as a result. Kennett (1984) writes the propagator matrices in an interesting form in which multiples be included or excluded. Using his formulation and

neglecting multiples, conversions, and transmission losses the convolution model of the earth may be derived as

$$d\left(h\right) = W * R\left(h\right), \tag{1.5}$$

where W is the wavelet, h is offset, $R\left(h\right)$ is the offset-dependent $\grave{P}\acute{P}$ reflectivity and $d\left(h\right)$ is the band-limited offset-dependent reflectivity. Using the convolution model and some other simplifying assumptions, the velocity and density may be solved using linear inverse techniques (Buland and Omre, 2003).

It is important to note that throughout this discussion horizontal layering and lateral velocity invariance has been assumed. Kennett actually performs the modeling in the spatial frequency domain k_x. To allow for lateral velocity variations, spatial finite differences may be used. Crase et al. (1990) use this as the basis for their elastic inversion. Further, Wapenaar and Berkhout (1989) use this as the basis for their elastic migration discussed in Section 1.3.2. If multiples are ignored this leads to the Kirchhoff modeling integral (Bleistein et al., 2001, Appendix E).

1.2.2 Linear Approximation

Returning now to the discussion of a single interface, note that the solution of the Zoeppritz equations (1.1) involves the inversion of a 4×4 matrix and a matrix multiplication. The set of equations are nonlinear, and nonlinear inversion techniques must be used to solve it. More importantly a null space exists because of the band-limited nature of the seismic data making the estimate of the elastic parameters underdetermined. This null space may be avoided by inverting for band-limited elastic parameters or some derivative of them. A number of papers have simplified the Zoeppritz equations by linearizing them namely Bortfeld (1962), Richards and Frasier (1976), Aki and Richards (1980). The linearization approximation requires that the fractional perturbations in elastic parameters is small (for example $\frac{\Delta\alpha}{\bar{\alpha}} << 1$). Aki and Richards published the linearized approximation for all the reflection and transmission coefficients shown in matrix (1.1). The $\grave{P}\acute{P}$ angle-dependent reflectivity $R\left(\theta\right)$

$$R(\theta) = \frac{1}{2}\left(1 - 4\bar{\gamma}^2 \sin^2\bar{\theta}\right)\frac{\Delta\rho}{\bar{\rho}} + \frac{1}{2}\sec^2\bar{\theta}\frac{\Delta\alpha}{\bar{\alpha}} - 4\bar{\gamma}^2\sin^2\bar{\theta}\frac{\Delta\beta}{\bar{\beta}}, \tag{1.6}$$

is parameterized in terms of $\bar{\alpha}$, $\bar{\beta}$,$\bar{\rho}$,$\bar{\gamma}$, $\bar{\theta}$, the average P-wave velocity, S-wave velocity, density, $\bar{\beta}/\bar{\alpha}$ ratio, and angle of incidence across the interface respectively. Similarly, $\Delta\alpha$, $\Delta\beta$, $\Delta\rho$ are the change in P-wave velocity, S-wave velocity and density

across the interface. This thesis refers to these fractional changes in parameters as reflectivity attributes or sometimes reflectivity for conciseness. These have the same mathematical form as P-wave and S-wave impedance reflectivity but not the same physical significance. For example, the P-wave velocity, S-wave velocity and density reflectivity are defined as

$$R_\alpha = \frac{\alpha_2 - \alpha_1}{\alpha_2 + \alpha_1} = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}}, \tag{1.7}$$

$$R_\beta = \frac{\beta_2 - \beta_1}{\beta_2 + \beta_1} = \frac{1}{2}\frac{\Delta\beta}{\bar{\beta}}, \tag{1.8}$$

and

$$R_d = \frac{\rho_2 - \rho_1}{\rho_2 + \rho_1} = \frac{1}{2}\frac{\Delta\rho}{\bar{\rho}}, \tag{1.9}$$

respectively. In the density reflectivity symbol the subscript d is used rather than ρ since it is easy to confuse ρ with p which will later be used to define P-wave impedance reflectivity. Ruger (2002) shows a nice derivation for equation (1.6) for the more complex case of VTI media where the isotropic case is a simplification.

In order to linearize the problem it is assumed that $\bar{\gamma}$ and the average angle-of-incidence to offset mapping is known *a priori*. This information is not always known with great certainty. This next two sections describe how this information may be supplied to the inversion. Later, Chapter 4 analyzes the influence of error in these parameters on the AVO inversion.

Angle-to-offset mapping

The linearized AVO approximation (1.6) is expressed in terms of average angle of incidence $\bar{\theta}$ rather than offset. The data are recorded as a function of offset so a mapping must be found to relate offset to average angle of incidence. As will be shown in Section 1.3.2, certain prestack migrations implicitly perform this mapping. This is the preferred methodology for structurally complex geology. However, if the data are not prestack migrated or if the output of migration is in the offset domain then this mapping must be still be performed.

If it is assumed that the earth is composed of a series of flat, homogenous, isotropic layers, then there are a number of numerically efficient methods to calculate the relationship between offset and the angle of incidence. Two approaches are discussed in this section, one is based on ray tracing while the other is a data driven approach following Walden (1991). Both these approaches require, at a minimum, that the

P-wave velocity α_n is known *a priori*. This velocity field may be constructed from local well control or travel time inversions from the seismic data. Because of the limited resolution of travel time inversion and the sparse nature of well control, a smooth velocity background model is typically used for this mapping. It is shown in Chapter 4 that it is sufficient to provide a smooth background velocity α_{back} for the ray tracing as long as it follows the running average of α_n. Using a smooth background has a certain theoretical appeal in that ray tracing is based upon a long wave-length approximation (Bleistein et al. 2001). From a practical standpoint, rapid variations in the velocity field α_i can cause rays to be refracted at or beyond the critical angle of refraction, making the ray tracing algorithm unstable.

Classical Ray Theory approach The first method used in this thesis follows the classical ray theory approach (Aki and Richards, 2002). The approach makes use of two relationships, the first of which calculates the two-way travel time τ_N, the time taken for a primary P-wave to travel from the surface to the N^{th} layer and back again along the ray path parameterized by the horizontal slowness, p, as

$$\tau_N(p) = 2\sum_{n=1}^{N} \frac{\Delta z_n}{\alpha_n} \frac{1}{\sqrt{1-\alpha_n^2 p^2}}, \tag{1.10}$$

where Δz_n is the layer thickness, and α_n is the P-wave velocity of the n^{th} layer. The second relationship calculates the horizontal distance traveled (offset) h_N for this ray parameter p as

$$h_N(p) = 2\sum_{n=1}^{N} \frac{p\alpha_n \Delta z_n}{\sqrt{1-\alpha_n^2 p^2}}. \tag{1.11}$$

Note, the first relationship is not required to map the offsets to angle of incidence, but is useful in generating the travel times for forward modeling. In deriving these relationships CDP and flat layer geometry is assumed so that the offset is twice the mid point and the two-way travel time is twice the traveltime to layer n.

If layers are parameterized so that they result in a constant time sample $\Delta t = \Delta z_n/\alpha_n$ then the algorithm may be parameterized either with respect to layer number n or two way zero offset travel time $t_{0_n} = n\Delta t$. The output of the algorithm is parameterized in terms of K regularly sampled offsets, with the user defining the minimum offset h_1 and maximum offset h_K.

The algorithm then consists of calculating the offset and traveltime for a series of

regularly sampled p values, so that if J is the number of offsets, then $p_j\,(1 \leq j \leq J)$,sampled from $p_1 = 0$ to $p = p_J$ where $p_J = \frac{1}{\alpha_{\min}}$, and $\alpha_{\min}$ is the minimum layer velocity. Choosing $p_{\max} = \frac{1}{\alpha_{\min}}$ implies that some rays are refracted at or beyond the critical angle of refraction. To address this, the ray is defined to be horizontal at the interface for which p becomes imaginary. Equations (1.10) and (1.11) are used to construct the two two-dimension arrays $h\,(t_{0_n}, p_j)$ and $\tau\,(t_{0_n}, p_j)$. Then, these arrays are then interpolated and reordered so that functional dependence is changed to $p\,(t_{0_n}, h_k)$ and $\tau\,(t_{0_n}, h_k)$. The angle of incidence may then be calculated using Snell's law

$$p = \frac{\sin\theta}{\alpha_n}. \tag{1.12}$$

This approach was used to generate the forward models in this thesis as well as calculate the angle of incidence for the majority of AVO inversions shown. The major weakness of this approach is that the results are dependent on overburden velocities. If there are errors in the overburden velocity model or unaccounted static shifts exist, perhaps due to elevations statics not being properly accounted for, significant errors may be introduced into the mapping.

Walden approach to ray tracing The second approach used in this thesis is data driven. The horizontal slowness may be calculated from the stacking velocity and then subsequently to angle of incidence using Snell's law (Walden, 1991).

Two-term NMO parameterization The two-term NMO equation (Taner and Koehler, 1969) calculates the two-way travel time τ

$$\tau^2 = t_0^2 + \frac{h^2}{V_{stak}^2}, \tag{1.13}$$

at offset h and is parameterized in terms of the stacking velocity V_{stak} and zero offset two-way travel time t_0. The horizontal slowness is the spatial derivative of equation (1.13)

$$p = \frac{d\tau}{dh} = \frac{h}{V_s^2\sqrt{\left(\frac{h^2}{V_{stak}^2} + t_o^2\right)}} = \frac{h}{V_{stak}^2\tau}. \tag{1.14}$$

Once again Snell's law (equation 1.12) may be used to convert to angle of incidence

$$\sin\theta = \frac{\alpha h}{V_{stak}^2 \tau}. \tag{1.15}$$

Three-term NMO parameterization In areas with significant vertical velocity heterogeneity higher order terms are significant and the two term formulation is not as accurate as the classical ray tracing approach. This is most problematic for large offsets and angles. For large offsets, calculating the horizontal slowness based on a higher order approximation (Taner and Koehler, 1969; de Bazelaire, 1990; Thore et al. 1994; Castle, 1994; Alkhalifah and Tsvankin, 1995) is more appropriate. This thesis uses the Castle (1994) shifted hyperbola approach

$$\tau \approx \tau_s + \sqrt{\tau_0^2 + \frac{h^2}{v^2}}, \tag{1.16}$$

because of its accuracy and the fact it only assumes isotropic layering. This simplifies future theoretical developments. The approach is parameterized in terms of the time variant static shift

$$\tau_s = t_0 \left(1 - \frac{1}{S}\right), \tag{1.17}$$

and the dimensionless shifting parameter

$$S \equiv \frac{\mu_4}{\mu_2^2}, \tag{1.18}$$

where time weighted moments of velocity distribution are given by

$$\mu_j = \frac{\sum_{k=1}^{N} \Delta t_{0_k} \alpha_k^j}{\sum_{k=1}^{N} \Delta t_{0_k}}. \tag{1.19}$$

Note that $\mu_2 = V_{rms}^2$. in addition, the parameters τ_0 and v are defined as

$$\tau_0 = \left(\frac{t_0}{S}\right), \tag{1.20}$$

and

$$v^2 = S\mu_2 = SV_{rms}^2. \tag{1.21}$$

The horizontal slowness calculated from the spatial derivative is

$$p = \frac{d\tau}{dh} = \frac{h}{(\tau - \tau_s)\, v^2}. \tag{1.22}$$

Once again Snell's law (equation 1.12) may be used to convert to angle of incidence

$$\sin\theta = \frac{\alpha h}{(\tau - \tau_s)\, v^2}. \tag{1.23}$$

Note that if there is no vertical velocity heterogeneity ($S = 1$), then the two-term and three-term results are equivalent.

The advantage of this approach is that it is data driven. Stacking velocities may be picked with a high degree of accuracy. The background velocity is still required to calculate the angle of incidence, but only errors associated with the analysis time influence the offset-to-angle mapping error unlike the classical ray tracing approach.

Background β/α ratio

In addition to the offset-to-angle mapping, the ratio $\bar{\gamma}$ must be defined in the linear operator. This may be constructed from local well control or from some empirical relationship such as the mudrock relationship. In order to map the offset to angle a background P-wave velocity model was constructed. The mudrock relationship that is introduced later in this chapter (equation 1.34) may be modified to estimate γ

$$\gamma \approx \frac{1}{m}\left(1 - \frac{b}{\alpha}\right), \tag{1.24}$$

from the P-wave velocity α and the slope m and intercept terms b. These last two parameters are described in Section 1.4.1. Chapter 4 explores the impact of $\bar{\gamma}$ errors on the AVO inversion.

1.2.3 AVO Inversion

Given $\bar{\gamma}$, the relationship between offset and angle-of-incidence, equation (1.6) may be written in matrix notation for M offsets as

$$\begin{bmatrix} R(\bar{\theta}\,(h_1)) \\ \vdots \\ R(\bar{\theta}\,(h_M)) \end{bmatrix} = \begin{bmatrix} \sec^2\bar{\theta}_1 & -8\gamma^2\sin_1^2\bar{\theta} & \left(1-4\bar{\gamma}^2\sin^2\bar{\theta}_1\right) \\ \vdots & \vdots & \vdots \\ \sec^2\bar{\theta}_M & -8\gamma^2\sin^2\bar{\theta}_M & \left(1-4\bar{\gamma}^2\sin^2\bar{\theta}_M\right) \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \quad (1.25)$$

Including random noise $\mathbf{n}$ in the linear model, equation (1.25) may be written succinctly as

$$\mathbf{G}\mathbf{x} = \mathbf{d} + \mathbf{n}, \quad (1.26)$$

where $\mathbf{G}$ is the linear operator, $\mathbf{x}$ the unknown parameter vector and $\mathbf{d}$ the input data vector (offset-dependent reflectivity). The parameter vector is composed of the P-wave and S-wave velocity and density reflectivity attributes. The matrix $\mathbf{G}$ is completely determined by the acquisition geometry and the background velocity fields.

If there are more than three unique offsets and the noise is assumed to be uniform, Gaussian, and uncorrelated, then equation (1.26) may be solved by least squares (Menke, 1984)

$$\mathbf{x} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}. \quad (1.27)$$

Smith and Gidlow (1987) made the observation that this operation is equivalent to applying a series of offset-dependent weights to the data and then summing the weighted data. These weights are completely defined by the acquisition geometry and the background velocity fields so they may be precalculated. Since this weighted summation is a linear process, the least squares inversion is also linear.

The reflectivity solution provided by equation (1.27) is only for one interface. For many interfaces, the inversion must be performed for each interface, thus for N interfaces the inversion is performed N times. If the seismic data are full-spectrum then the number of layers can be taken to be the number of time samples. Up to now, it has been ignored that the seismic data are the result of a convolution with some sort of source wavelet. Provided that the wavelet does not change as a function of offset, the order of the operators does not matter, since both the convolution and the inversion are themselves linear processes. The inversion may be thought of as

being performed on full spectrum reflectivity and then convolving the result with the source wavelet. This results in three band-limited reflectivity attribute time series.

1.2.4 Impedance inversion

The output of the AVO inversion is in terms of velocity and density reflectivity. These may be written in terms of P-wave and S-wave impedance reflectivities (Appendix B)

$$R_p = R_\alpha + R_d, \tag{1.28}$$

and

$$R_s = R_\beta + R_d, \tag{1.29}$$

respectively. Peterson et al. (1955) noted that for small reflection coefficients integrating the P-wave impedance reflectivity

$$I_{p_N} \approx I_{p_0} \exp\left\{2\sum_{n=1}^{N} R_{p_n}\right\}, \tag{1.30}$$

it is possible to calculate the P-wave impedance I_p. The calculation is actually more complex since the seismic data and impedance reflectivity estimate are both band-limited. Typically, both the low and high frequencies are missing. The lack of high frequencies limit the resolution of the inversion. The low frequencies are actually more problematic since they define the background trend of the inversion. Lindseth (1979), Oldenburg et al. (1983), van Riel and Berkhout (1985), Debeye and van Riel (1990), and Sacchi (1997) developed inversion algorithms that invert band-limited reflectivity data to impedance. The major difference between the methods listed is how they deal with the non-uniqueness in the problem either through the parameterization of the problem, their constraints, or solution method. Note, other reflectivity attributes may calculate their elastic parameter counterparts in a similar fashion.

Having estimated the density, P-wave and S-wave impedances, other elastic parameters may be calculated using various transforms found in the literature. One popular transform (Goodway et al. 1997) transforms the estimates to $\lambda\rho$ and $\mu\rho$ using

$$\lambda\rho = I_p^2 - 2I_s^2, \tag{1.31}$$

and

$$\mu\rho = I_s^2. \tag{1.32}$$

Upon inverting the reflectivity to elastic parameters, the seismically derived results may be compared to the well control for analysis and comparison.

Before leaving this section it is important to point out that both the AVO and poststack impedance inversion problems may be posed as linear problems. The two inversions have different issues, the AVO problem is overdetermined, though ill-conditioned and its major issue is how to deal with noise. The poststack impedance inversion problem is underdetermined with its major issue of how to deal with the missing data. The two inversion problems are linear and may be combined; this is done in Chapter 6. This means the combined problem is also underdetermined. Due to the band-limited nature of the seismic data, all elastic inversions are underdetermined and therefore their solutions will be non-unique.

1.2.5 Nonlinear inversion

The Zoeppritz equations may also be approximated by a nonlinear equation. Ursenbach (2002) shows an approximation including all second order terms. In this paper he also includes an error analysis related to ignoring these higher order terms. It turns out that most of the error is related to ignoring the 2^{nd} order S-wave velocity reflectivity term. In a subsequent paper, Ursenbach (2004) shows an efficient method to include this term while leaving the other terms linear.

In this thesis I choose not to use a higher order approximation. I felt that if higher order terms were to be included, then first order multiples and converted waves should also be included. They both introduce error of the same order. Simmons and Backus (1994) show a modeling study where they compare how well various approximations model the actual data. The modeling includes both the converted waves and intrabed multiples. For the interface they analyzed they found that the Aki and Richards approximation equation (1.6) better matched the observed data than the Zoeppritz equations due to interference from an intrabed multiple and a converted wave.

1.3 Seismic data preconditioning

1.3.1 Factors that affect AVO

Sheriff (1975) lists several broad categories of factors which affect amplitude variations with offset including data collection, wave propagation effects, data processing and geologic factors. The geologic factors and their corresponding elastic and density parameters are the ultimate objective of AVO inversion. In order to infer these parameters through AVO, the seismic data must be properly preconditioned so that the assumptions made in the previous section are met. Failure to do so introduces both systematic and random error into the data. Systematic error is particularly problematic since it is difficult to detect and leads to biased estimates. In this section, the various factors affecting AVO unrelated to the geology are first summarized and then a processing sequence is outlined to address these. This type of processing is generally called controlled amplitude processing (Soroka et al., 2002). Finally, an example is shown illustrating the importance of processing the data following a controlled amplitude processing sequence.

As outlined above, the factors affecting AVO unrelated to the geology include data collection, wave propagation effects, and data processing artifacts. In the data collection category, instrumentation, noise, and acquisition geometry all influence AVO. Further, source and receiver directionality, coupling, and array effects also influence the amplitude variations with offset and must be addressed prior to, or as part of the AVO inversion. The amplitude variations with offset are also influenced due to wave propagation factors such as focusing due to reflector curvature or lateral velocity heterogeneity, spreading losses, transmission losses, anelastic attenuation, surface boundary effects, interference from converted waves, intrabed, pegleg, and surface-multiple reflections. To address these factors the seismic data must properly conditioned or processed to meet the assumptions inherent in the AVO model. Unfortunately, processing artifacts can also distort the AVO response. These include NMO stretch, residual NMO, and artifacts introduced by spatial filters used to suppress noise, In addition, trace to trace operations, such as scaling or deconvolution, can modify the AVO response.

1.3.2 Migration

Ideally, migration is performed on the seismic data to address wave propagation issues. This is the approach of Berkhout (1997), and Wapenaar and Berkhout (1989), where the elastic inversion is treated as a two-step inversion. The first step removes the effects of wave propagation to obtain the angle-dependent reflectivity which the second step uses to estimate the elastic parameters. Wave propagation effects such as surface boundary conditions, directivity, source and receiver ghosts, and wave propagation between layers (including multiples) can be written as a series of linear operators. These can then be removed either using the inverse or conjugate operator as part of the migration. These corrections change the AVO behavior of the data. Only after these corrections may AVO be used to infer the elastic properties of the rock. For the case of horizontal layers, the corrections are outlined in Appendix A. Note that these corrections are a function of $\alpha(x, y, z)$, so that the corrections vary as a function of spatial position. Failure to account for this may introduce systematic errors into the AVO estimates which vary spatially. For the more complex case of arbitrary reflector geometry these corrections become increasingly complex. The issue of how to perform true-amplitude migration in complex media is an area of active research (Bleistein et al., 2001).

Just as importantly, migration improves the lateral resolution of the seismic data, for example, the lateral extent of a bright or dim spot and the positioning of the faults. Prestack migration properly maps the energy originating from interfering diffractions and collapses it. These interfering events would otherwise be treated as noise by the AVO inversion. Migration helps with dip related issues of reflector point dispersal, dip-dependent NMO, and the calculation of the angle of incidence. The output of the migration may be in terms of horizontal slowness p (Wapenaar and Berkhout, 1989; Rickett and Sava, 2001; and Brandsberg-Dahl et al., 2003) thus implicitly mapping the data from offset to angle as part of migration removing the necessity of ray tracing as part of the AVO inversion.

Q-compensation and multiples can also be formulated as part of migration but these are typically addressed prior to migration as preprocessing steps for simplicity and quality control reasons. Q-compensation implicitly gains the data, alters the wavelet in a time and frequency-dependent fashion and must be quality controlled along with other wavelet processing procedures. This is discussed further in the next section. Surface-multiple attenuation although formulated using linear mi-

gration operators (Verschuur et al., 1992) is an adaptive process and so is also best quality controlled as a separate processing step. Further, migration makes certain assumptions about the seismic data requiring additional pre-conditioning. For AVO inversion this effectively means the seismic data are processed with migration being one element in the processing sequence. This is the approach advocated by Mosher et al. (1996).

This approach works well for uniform acquisition geometry, such as marine recording. It works less well when the geometry changes in an irregular and sparse fashion such as found in land seismic data acquisition. Both Brandsberg-Dahl et al. (2003), and Ronen and Liner (2000) note that the acquisition footprint may create AVO distortions. In addition, seismic data are recorded with finite offset leading to amplitude distortions in the transform to horizontal slowness space p, which distorts the AVO. These distortions may be addressed by performing least squares migration instead (Kuehl and Sacchi, 2001). However this is extremely computationally expensive and still an area of active research. So in practice, for land data with horizontal layers, AVO is often performed on NMO corrected data (Mazzotti et al., 1990; Spratt et al., 1993; Soroka et al., 2002; and Shen et al., 2002). Prestack migration is left out of the land processing sequence as it causes more problems than it solves. In this case, corrections for wave propagation must be done explicitly as part of the processing sequence.

1.3.3 Controlled amplitude processing

The objectives of creating a good stacked section (relative amplitude processing) and conditioning the gathers for AVO (controlled amplitude processing) are somewhat at odds. The typical objective in creating poststack data is to maximize the frequency bandwidth and signal-to-noise ratio for interpretation on a workstation. The prestack data may be scaled without the concern of preserving the AVO so that trace balancing or an AGC may be used. This is particularly useful since noisy traces may be scaled down so they contribute less to the stack. Further, poststack deconvolution may be applied to increase the bandwidth of the data. Poststack deconvolution implicitly applies a trace to trace scalar since the operator is scaled arbitrarily. Finally, spatial filters may also be used to attenuate coherent and random noise, again potentially introducing edge effects that alter the AVO. To address these concerns, data that might undergo AVO analysis may be dual streamed. This consists of processing the

data up to a certain stage in a manner suitable for both AVO and poststack analysis. At this point the data are dual streamed. In one flow the data are scaled in manner suitable for AVO. In the other, the prestack data may be scaled or spatially filtered so as to achieve the best looking stacked section. This dual streaming reduces the costs of performing the AVO analysis so that reprocessing is not required. This is the processing sequence described below and shown in Figure 1.2. Since most of the examples shown in this thesis are land seismic data, the land processing sequence is described in detail. Subsequent to this, the modifications necessary for marine data are described.

Land processing sequence

The first step in this sequence is to remove the shots and traces that are contaminated with too much noise. Next, the data are prepared for wavelet processing and scaling. To address source and receiver coupling issues, surface consistent scaling is applied to the data (Taner and Koehler, 1981). Surface consistent deconvolution (Cambois and Stoffa, 1992; and Cary and Lorentz, 1993) may then be performed to remove the source and receiver signatures and the earth filter or intrabed multiples. Prior to the design and application of the deconvolution, the seismic data must be made stationary. This may be achieved by performing some type of removable gain correction such as an exponential or t^n correction. More complex gain functions necessary for AVO analysis are performed later in the flow if the interpreter chooses to go to the expense of doing an AVO analysis. In addition, coherent noise such as ground roll may be removed to improve the operator design. If the filter introduces AVO distortions, this may be done in an off-line fashion. That is, only the data going into the deconvolution design is filtered. The seismic data on which the operator is applied is not filtered.

If the earth is anelastic, the stationarity assumption is not justified. In the assumptions outlined in Section 1.1, the earth is assumed to be elastic. This is generally not the case, so attenuation and dispersion must be addressed. This may be done as part of migration, as a separate processing step or as part of the AVO inversion. Generally, this is done as a separate processing step (Hargreaves and Calvert, 1991), prior to the deconvolution since Q-compensation alters the source wavelet and scales the data in a time and offset-dependent manner. The Q-compensation may be used to address both the intrinsic and effective attenuation. It is difficult to separate

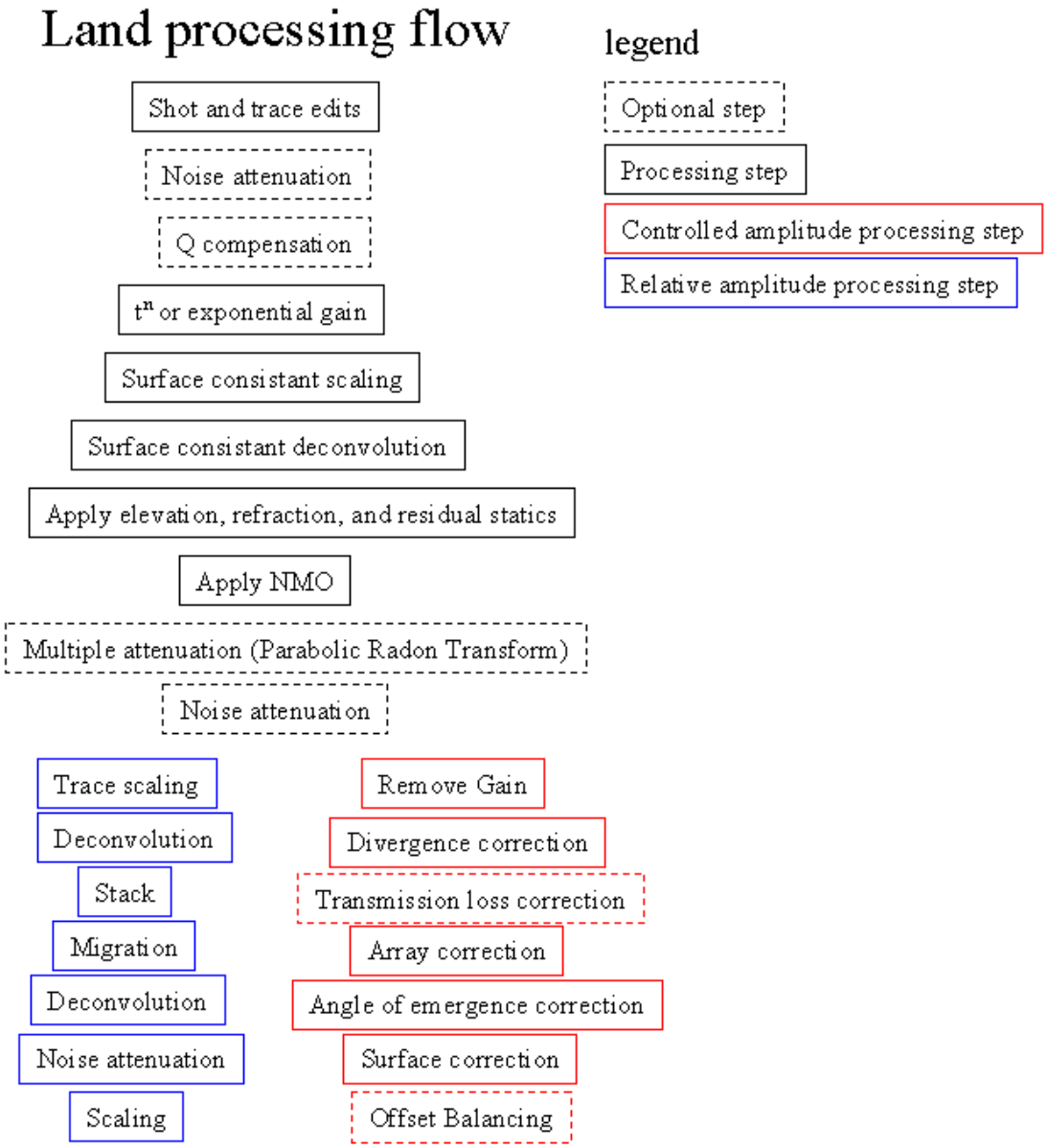

Figure 1.2: Controlled amplitude land processing flow. Steps in red are used to prepare gathers for AVO. Steps in blue are used to create the optimal migrated stack. Common processing steps are in black while optional steps are shown with dashes.

the influence of attenuation effects and the transmission effects of interbed multiples (O'Doherty and Anstey, 1971). Intrinsic attenuation is due to mechanisms such as squirt-flow (Mavko and Jizba, 1991). Effective attenuation is due to scattering or the combined effects of multiples and transmission losses (O'Doherty and Anstey, 1971). Shapiro and Hubral (1999) use statistical models to correct for effective attenuation. For hetrogeneous intrinsic Q models, the Zoeppritz equations are modified, necessitating the use of another linearized approximation (Ursin and Stovas, 2002).

After wavelet processing, the velocity and statics models must be created and applied. This is typically done in a bootstrap fashion. For time processing, RMS velocities are needed, however, interval velocities are required for the AVO inversion. The interval velocity model may be constructed from geostatistical methods, RMS velocities or some type of traveltime inversion such as traveltime tomography. This velocity model is low resolution and susceptible to error. To help stabilize this process, interval velocities from nearby well control may be used to constrain this model. If anisotropy is also present, this process becomes more complex, necessitating reviewing the AVO approximations.

If long offsets relative to the target depth are to be used, higher order terms and corrections are required (de Bazelaire, 1988; Castle, 1994; and Alkhalifah and Tsvankin, 1995). This requires the construction of additional velocity fields. Ignoring these corrections may introduce residual NMO, degrading the subsequent AVO inversion. As explained in the Section 1.2.3, AVO inversion is typically performed one time sample at a time, so the moveout needs to be removed. The gathers may be flattened either by performing NMO or prestack migration.

Before or after NMO, additional spatial filters may be applied to attenuate multiples, converted waves and other forms of coherent noise not included in the AVO data model. Care must be taken not to introduce AVO artifacts with these filters. For example, Marfurt et al. (1996) show that multiple attenuation using a parabolic Radon transform introduce AVO artifacts. The actual AVO inversion is quite robust and can tolerate some forms of coherent noise as will be shown in Chapter 8. If the coherent noise is not dominant, it is often better to leave this noise in and let the AVO inversion deal with it rather than filter it and introduce systematic errors (Downton and Mackidd, 1997).

At this point, the flow becomes dual streamed. One flow leads to a stacked section unconcerned with preserving AVO relationships. The other prepares gathers

for AVO inversion. At this point the gathers must be processed to remove the effects of wave propagation. This may be done either prestack migration of the gathers or NMO correcting the gathers followed by a series of wave propagation corrections (Appendix A). For the reasons listed in Section 1.3.2, the data in this thesis are NMO corrected thus necessitating the application of wave propagation corrections. Prior to their application, the removable gain function used to make the time series stationary prior to the deconvolution must be removed. The geometrical spreading (Ursin, 1990), array, angle-of-emergence, and surface boundary corrections (Spratt et al., 1993) may then be applied. If applicable the correction for offset-dependent transmission losses may also be included at this stage. These corrections assume the earth is composed of a series of flat, homogeneous, isotropic layers. These gain corrections modify the data in a time-dependent and offset-dependent fashion and are a function of $\alpha(x, y, z)$. Lateral velocity changes in the near surface, or changes in the water depth can introduce significant AVO across the seismic survey.

After the application of all these corrections, residual offset-dependent corrections may remain in the data. The data may be calibrated against walk away VSPs or synthetic prestack models generated from nearby well control. A simple offset-dependent scalar may be used to correct for discrepancies. Ross and Beale (1994) call this process offset balancing.

Marine processing sequence

There are several significant differences between processing marine seismic data and processing land seismic data. Surface-multiples are generally more problematic. Generally, some type of surface-multiple attenuation, such as presented by Verschuur et al. (1992), is run prior to the wavelet processing. Surface consistent processes do not make sense, since the receivers are constantly moving. The surface consistent scaling is modified to become channel consistent scaling. Instead of surface consistent deconvolution, designature deconvolution is performed. In terms of the wave propagation corrections, hydrophones are omnidirectional so no angle of emergence correction is needed. The surface boundary correction is modified to account for the source and receiver ghost. Typically, marine data acquisition is quite regular, and the geologic objectives have a strong structural component to them, so prestack migration is usually performed.

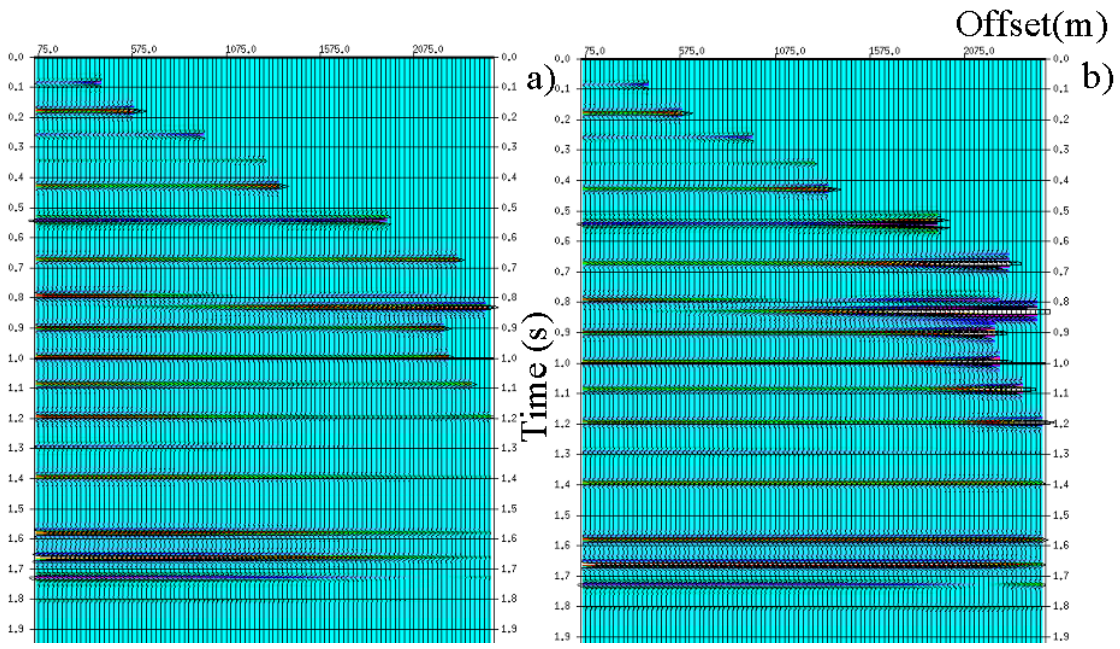

Figure 1.3: Synthetic CMP seismic gather with correct scaling a) and after trace balancing b).

1.3.4 Examples

As noted above, one of the major differences between controlled amplitude processing and relative amplitude processing is that the former allows for trace dependent scaling. Trace scaling normalizes the RMS energy of the trace over some design window. After trace scaling, the amplitude variation with offset relationship is distorted. Figure 1.3 shows a model gather from Downton et al. (2000) before and after trace balancing. The AVO relationship is changed. This results in a systematic error in AVO inversion parameter estimates.

Figure 1.4b shows the AVO inversion for S-wave impedance reflectivity based on the scaled gathers while Figure 1.4a shows the same based on the correctly scaled gathers. Figure 1.4a is identical to the ideal S-wave impedance reflectivity. Figure 1.4b shows the effect of the trace balance on the estimate S-wave impedance reflectivity. The anomalies are still evident, but the background reflectivities are severely distorted. Note that a simple poststack or time variant scalar cannot correct this distortion. For example, the regional reflector at 1.58 seconds on the properly scaled result is absent on the result with relative amplitude scaling. Figure 1.4b cannot be

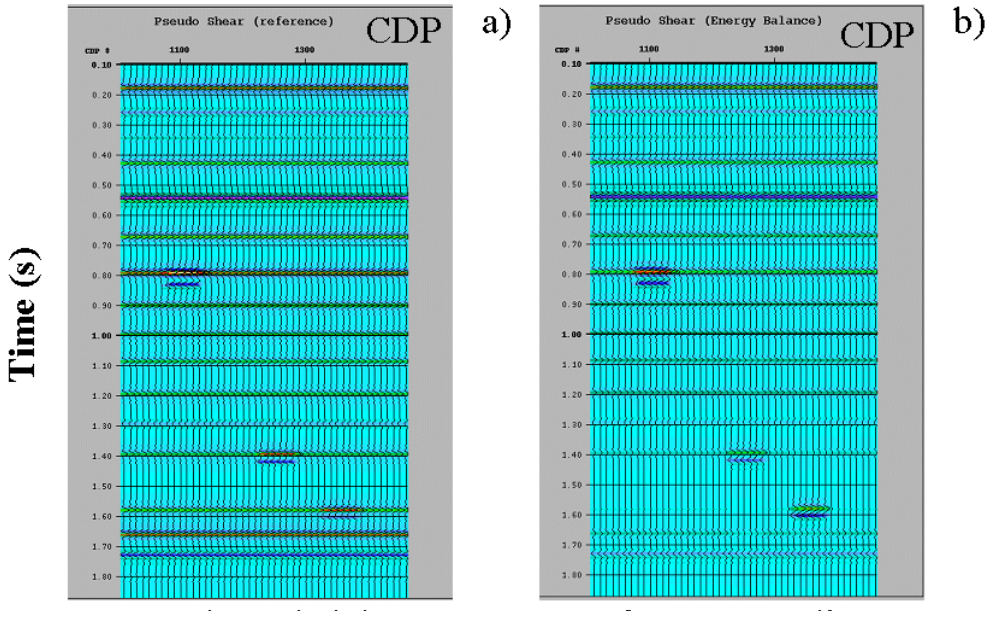

Figure 1.4: S-impedance reflectivity estimate from AVO inversion of synthetic data with correct scaling a) and trace balanced data b).

scaled to look the same as the correct result (Figure 1.4a).

Another example from Downton et al. (2000) shows a comparison of AVO results from relative amplitude and controlled amplitude processing. The seismic data are from northeast British Columbia, Canada with the exploration target being the Halfway sand at 0.79 seconds. The two flows were processed a year apart with both results being best efforts at their respective times. Figure 1.5, shows the S-wave impedance reflectivity estimated from the relative amplitude processing while Figure 1.6 shows the S-wave impedance reflectivity estimated from the controlled amplitude processing. Note the better continuity of events and signal-to-noise ratio of the controlled amplitude processing. Testing showed most of the improvement was due to using surface consistent scaling rather trace dependent scaling. Coherent noise related to surface conditions unduly biased the trace scaling, introducing spatially variant systematic errors. Further, the overall AVO trend was biased by trace scaling further degrading the result. The prestack gathers are shown in Figure 1.7 and 1.8.

These two examples show the importance of controlled amplitude processing. To get good results from AVO inversion, the input seismic data must be properly condi-

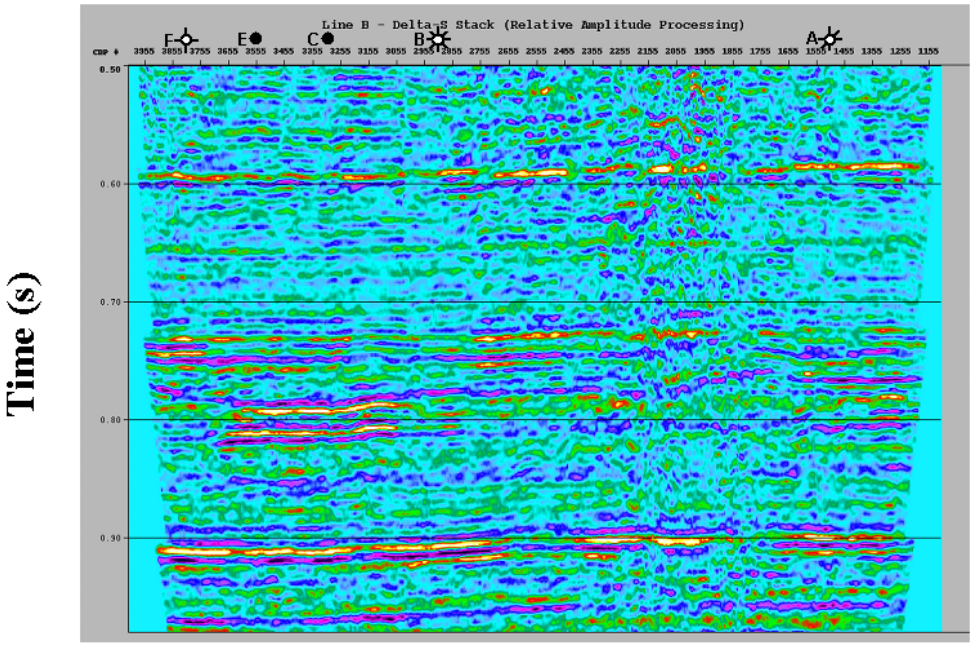

Figure 1.5: S-wave impedance reflectivity estimate for Halfway line based on relative amplitude processing.

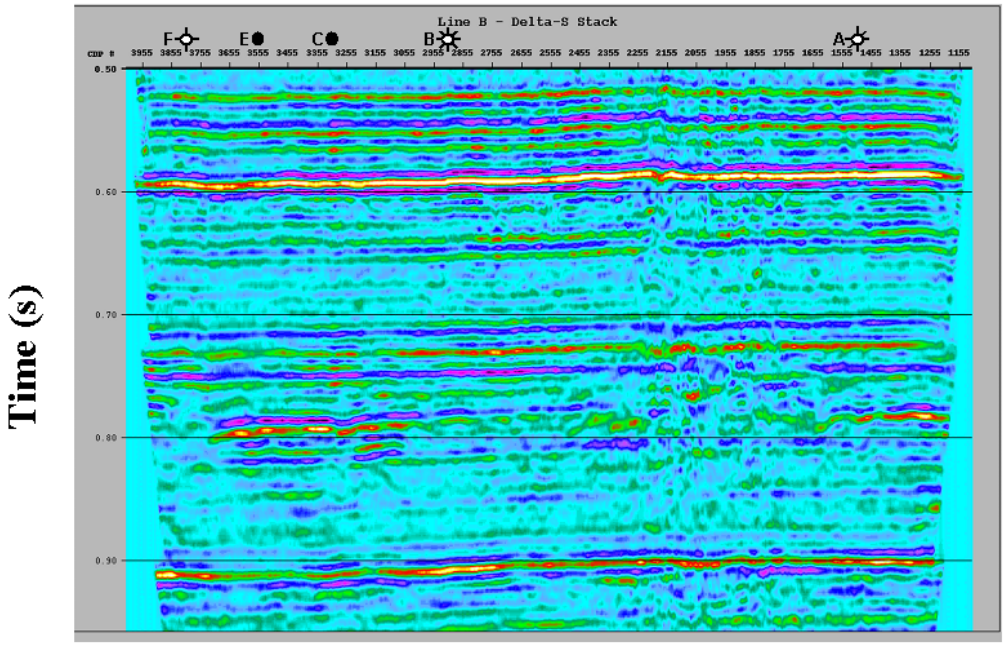

Figure 1.6: S-wave impedance reflectivity estimate for Halfway line based on controlled amplitude processing.

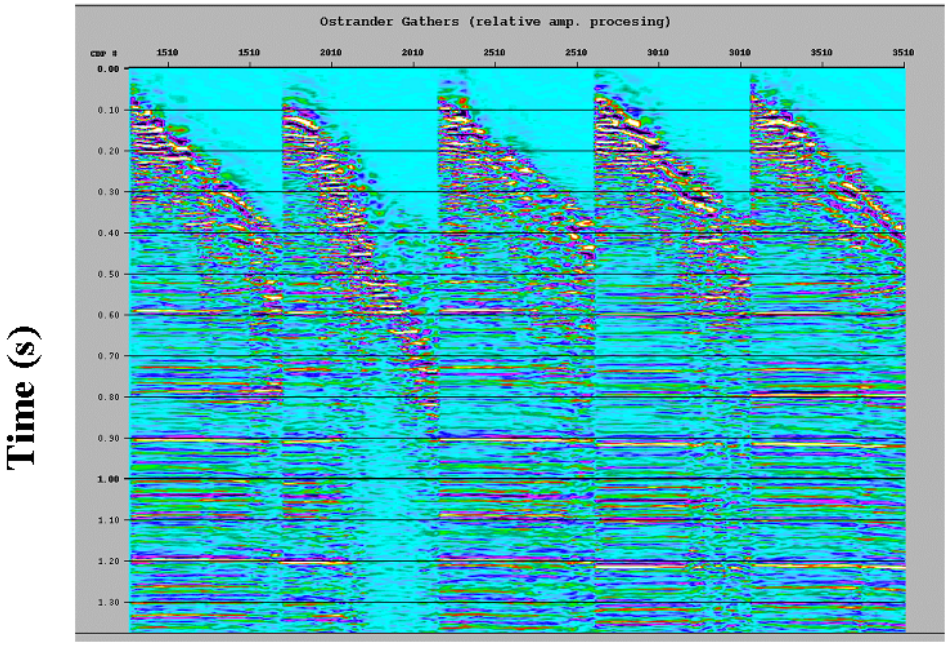

Figure 1.7: Ostrander gathers generated from relative amplitude processing flow for Halfway line.

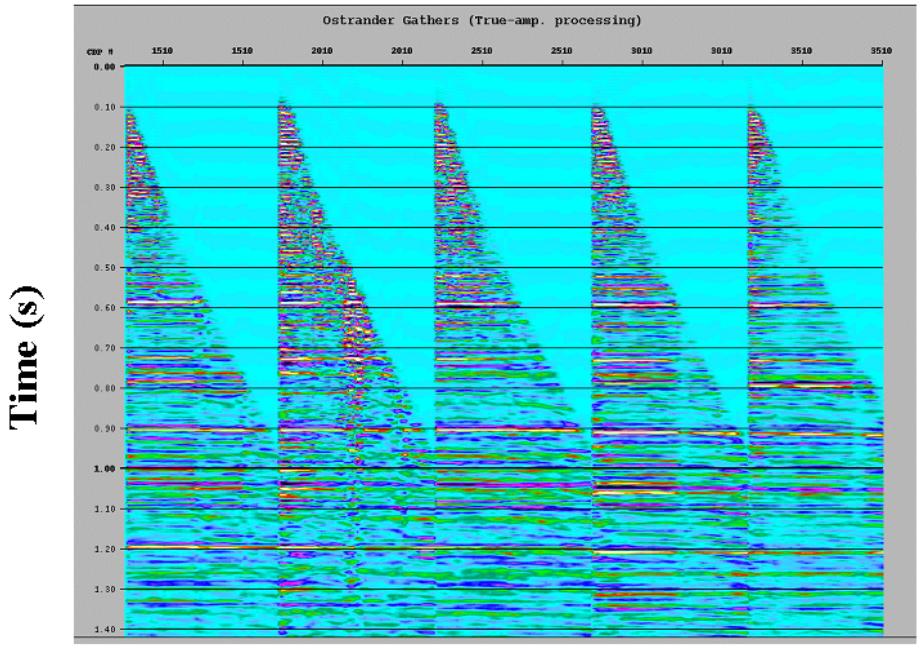

Figure 1.8: Ostrander gathers generated from controlled amplitude processing flow for Halfway line.

tioned.

1.4 Seismic to rock and fluid property mapping

Having estimated the elastic properties, the next question is "what do these mean?" Two broad methodologies exist to answer this question. The first, similar to AVO inversion, estimates the rock and fluid properties from the elastic parameters based on some type of theoretical model based on rock physics relationships. The second broad category of methodologies involves using local core or well log data with known relationships between the elastic parameters and the rock and fluid properties, acting as a template for the interpretation of the seismic data and the resultant elastic parameters. This may be done using the estimated elastic parameters, the band-limited AVO reflectivity attributes and prestack gathers. These two approaches, plus some empirical relationships that may be used to constrain the AVO inversion are reviewed in this section.

1.4.1 Mapping based on theoretical model

Mudrock Relationship

From rock physics studies Castagna et al. (1985) observed, for clastic rocks, that the P-wave velocity α may be estimated from S-wave velocity β (and vice versa) using the empirical mudrock relationship

$$\alpha \approx 1.16\beta + 1360\,\mathrm{m/s}. \tag{1.33}$$

Greenberg and Castagna (1992) extended this to other lithologies including carbonates. Figure 1.9 shows a crossplot for the brine-saturated clastic core measurements made by Han (1986). Note the strong linear trend apparent in Figure 1.9 supporting equation (1.33). Figure 1.10 shows the same data but with porosity as the color overlay. Note that within each porosity unit there is a strong linear relationship suggesting equation (1.33) has a more complex functional relationship. Figure 1.11 shows a subset of the data, but now with the percentage shale as the color overlay. Han et al. (1986) suggested that the P-wave and S-wave velocity of clastic rocks are a function of porosity and shale content. Eberhart-Phillips et al. (1989) further refined the analysis and suggested that there is also minor dependence on effective pressure.

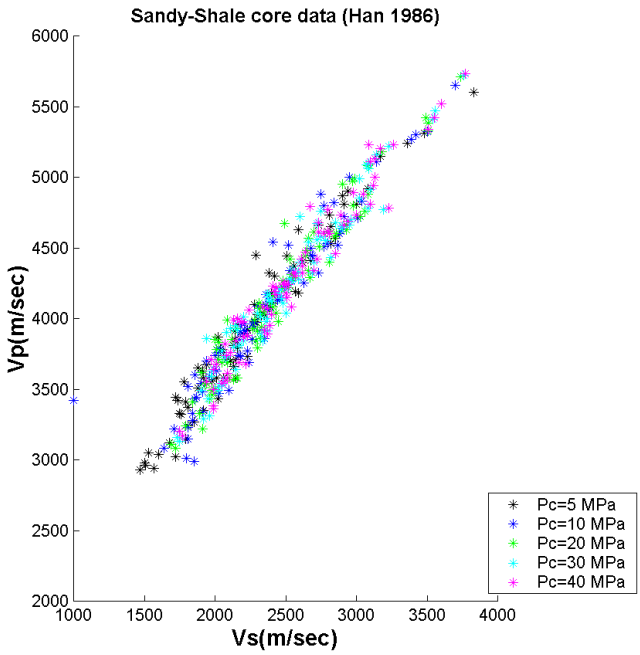

Figure 1.9: Crossplot of P-wave and S-wave velocity measurements for clastic rocks (Han, 1986) with effective pressure as color overlay.

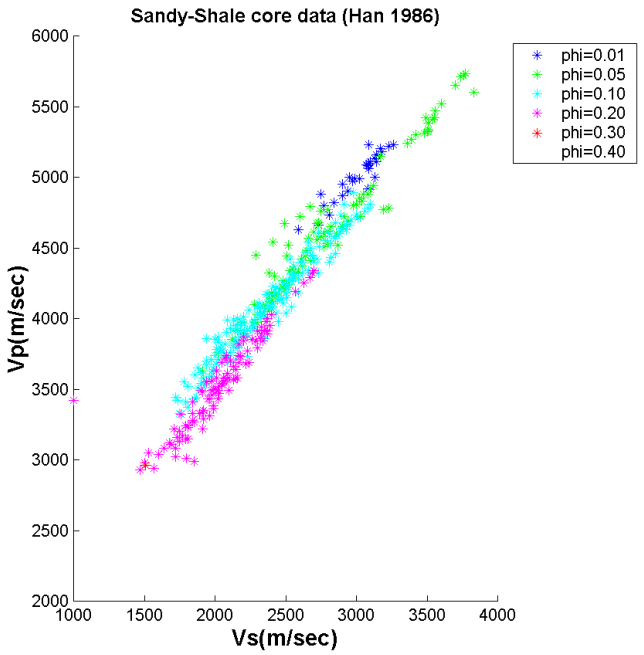

Figure 1.10: Crossplot of P-wave and S-wave velocity measurements for clastic rocks (Han, 1986) with porosity as color overlay.

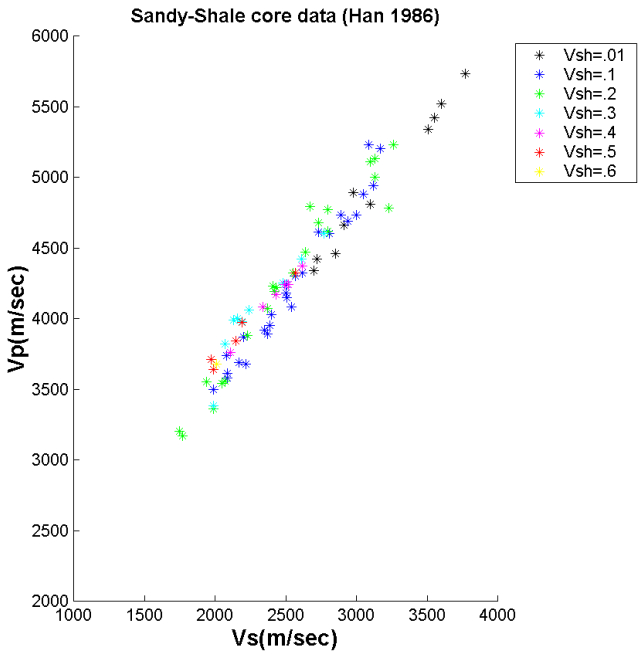

Figure 1.11: Crossplot of P-wave and S-wave velocity measurements for clastic rocks (Han, 1986) with volume of shale as color overlay.

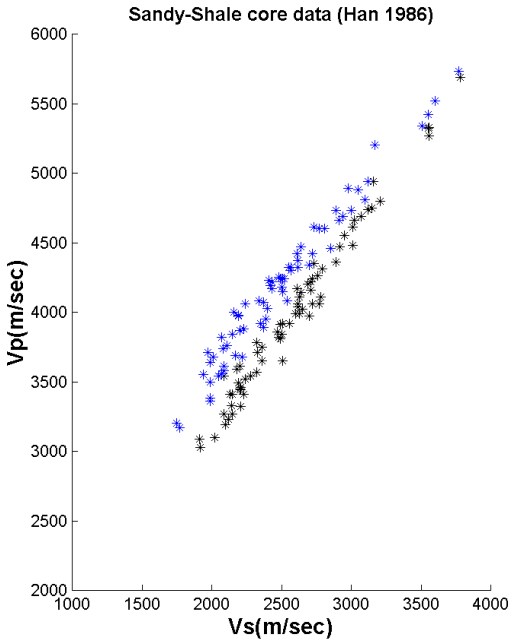

Figure 1.12: Crossplot of P-wave and S-wave velocity measurements for clastic rocks (Han, 1986) for wet samples (blue) and dry samples (black).

This suggests that mudrock relation is a function of all these parameters suggesting that equation (1.33) be generalized to

$$\alpha \approx m\beta + b, \tag{1.34}$$

where the mudrock slope m and intercept b are locally calibrated over the target interval.

Figure 1.12 shows the P-wave and S-wave velocity crossplot of core samples with different fluids. The blue data points are brine saturated while the black data points are dry. Under many circumstances, gas sand should behave in a similar fashion to the dry data points. Note that the wet data points are shifted upwards along the P-wave velocity axis compared to the dry data points. This is due to the fluid stiffening the rock, increasing the incompressibility and the P-wave velocity while largely unaffecting the rock's rigidity and S-wave velocity. The gas filled sandstone is more compressible than the brine filled sandstone. The effect is more pronounced as the porosity increases. This effect is the basis of several fluid detection schemes used in AVO analysis.

Gardner Velocity-Density Relationship

Another relationship used in this thesis is the Gardner et al. (1974) relationship. It is an empirical relationship linking the density and P-wave velocity. If the velocity is expressed in the metric units km/s then the relation is

$$\rho = 1.741\alpha^{0.25}, \tag{1.35}$$

while in the imperial units ft/s for velocity the expression is

$$\rho = 0.23\alpha^{0.25}. \tag{1.36}$$

If the lithology is known this relation may further refined by writing it with lithology dependent coefficients

$$\rho = e\alpha^{g}. \tag{1.37}$$

such as shown in Table 1.1 (Castagna et al., 1993). In general these coefficients are dependent on other factors, in addition to lithology, such as the local burial history. Thus, if given the opportunity, it is best to calibrate these coefficients using local well

control.

Lithology	e	f	α range (km/s)
Shale	1.75	0.265	$1.5 - 5.0$
Sandstone	1.66	0.261	$1.5 - 6.0$
Limestone	1.50	0.225	$3.5 - 6.4$
Dolomite	1.74	0.252	$4.5 - 7.1$
Anhydrite	2.19	0.160	$4.6 - 7.4$

Table 1.1: Gardner coefficients for common lithologies

1.4.2 Mapping based on template data

The second broad category of methodologies used to estimate rock and fluid properties from elastic properties involves using local core and/or well log data with known relationships between the elastic parameters and the rock and fluid properties, to act as a template. For example, well logs from two wells - one with gas, and one with a brine sand - may be used to establish a template for how the elastic parameters respond to changes in fluid content. Additional wells and reservoir zones may be used to understand the influence of lithology, porosity, effective pressure and depth on the elastic parameters. Often the well control is insufficient to adequately describe the full range of desired rock and fluid properties so theoretical relationships are used to help fill in the gaps. For example, the Gassmann (1951) equation may be used to understand the influence of different fluid saturations. Avseth (2000) and Takahashi (2000) describe rock and fluid inversions based on this approach. Often this approach is not implemented formally as an inversion. The core and well log data are used as a template to interpret the seismic data in a subjective manner. This may be done using the estimated elastic parameters (Goodway et al., 1997), the band-limited AVO reflectivity attributes (Verm and Hilterman, 1995; Castagna and Swan, 1997) and prestack gathers (Ostrander, 1984; Rutherford and Williams, 1989).

Interpretation of elastic parameters

Interpreting the seismically derived parameter estimates as elastic parameters is quite natural since the well and core data may be easily manipulated to be displayed in this domain. Figure 1.13 shows the elastic parameters $\lambda\rho$, $\mu\rho$ crossplotted for a Glauconite gas sand from the Blackfoot area in Alberta, Canada. Note the ability to

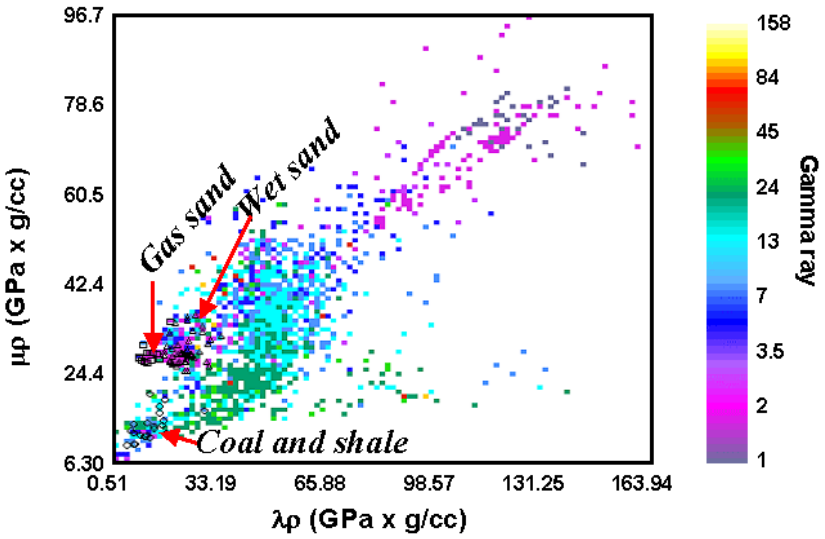

Figure 1.13: Crossplot of Blackfoot well log data showing separation between brine and gas filled Glauconite sands.

discriminate between gas sands, wet sands, coals and shales in the crossplot space. This can be used as a template to interpret the elastic parameters derived from the seismic.

Recall from Section 1.2.4 that these parameters may be estimated by first performing an AVO inversion and then a poststack impedance inversion. Figure 1.14 shows the estimated $\lambda\rho$ section for a seismic line through the Glauconite gas well together with its $\lambda\rho$, $\mu\rho$ crossplot. Using the well control as a template (Figure 1.13), a polygon was picked on the seismic crossplot to highlight the same area on the section correctly identifying the Glauconite gas sand.

Interpretation of prestack gathers

Alternatively, the prestack gathers may be interpreted. Template well logs may be forward modeled to understand the seismic response of the different rock and fluid properties. Figure 1.15a, from Downton et al. (2000) shows the brine filled Bluesky sand's AVO response while Figure 1.15b shows the gas filled gas sand AVO response. For this case, the gas sand's amplitude decreases as a function of offset even reverses polarity at the far offsets. In contrast the brine filled sand has constant AVO. The

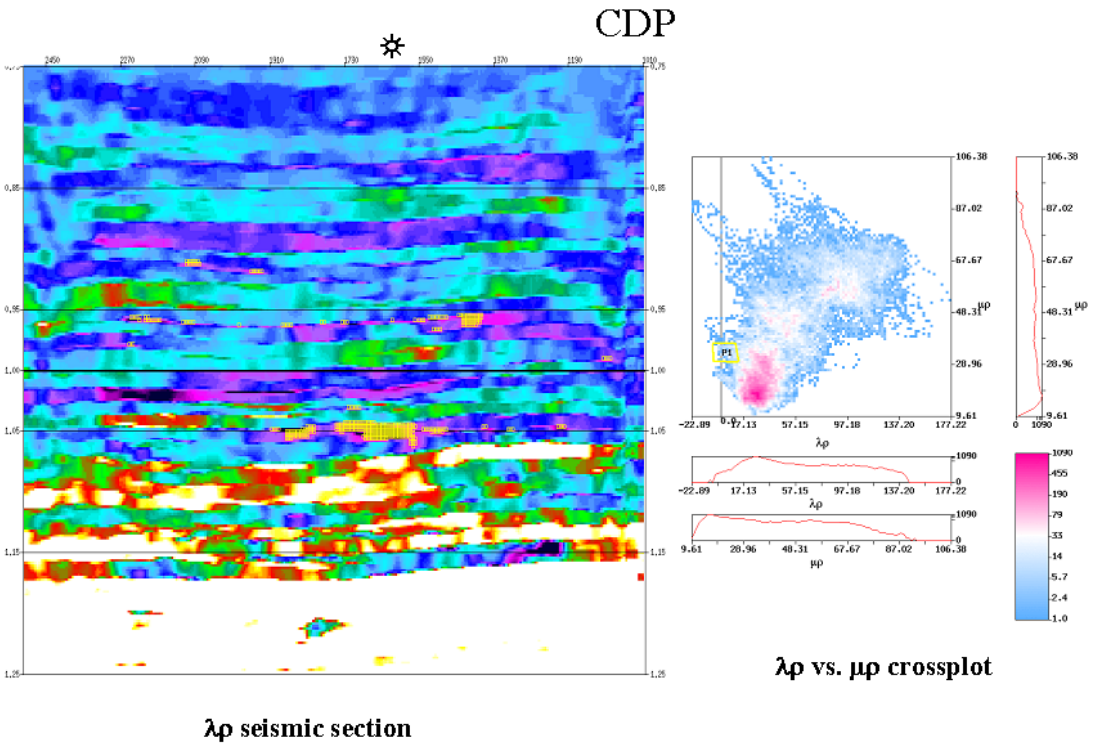

Figure 1.14: Estimate of lambda*rho section for Blackfoot seismic line crossing Glauconite gas well along with crossplotted seismic data. Note that yellow polygon on crossplot is able to identify gas sand on section (shown with yellow squares).

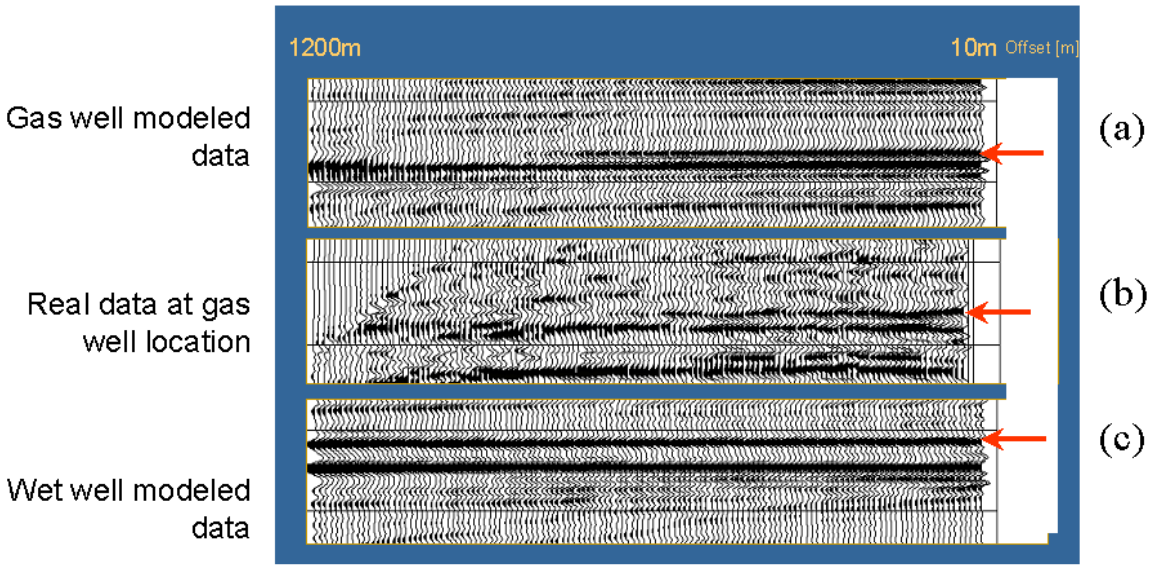

Figure 1.15: Synthetic gather showing wet sand response (a), gas sand response (b) and actual seismic data from gas well (c). Note that this is a Class I gas sand response.

actual seismic data (Figure 1.15c) is compared to the modeled templates.

Rutherford and Williams (1989) systematized these responses by classifying AVO gas sand responses into three classes. A Class I response occurs when the gas sand has a higher impedance relative to the overlying medium. The peak amplitude decreases as a function of offset with the potential of a polarity change at far offsets. Figure 1.15b is an example of a Class I reservoir. For the Class II response, the zero offset response is close to zero. This occurs when impedance of the gas sand is close to that of the overlying material. The amplitude may increase or decrease with offset. The Class III response is what is most often associated with AVO. It occurs when the gas sand has a lower impedance than the overlying material. For this case, the amplitude is a bright trough at near offsets which gets brighter as a function of offset. This is the response typically seen in low velocity clastics typical of unconsolidated sands. More recently, the concept of a Class IV gas sand has been introduced to classify a certain type of response seen in deeper Gulf Coast wells. In this case, the trough is bright at near offsets and dims with offset.

Interpretation of band-limited reflectivity data

Both of the preceding interpretation strategies have their shortcomings. Interpreting gathers for a large 3D volume is too time consuming. Inverting for elastic parameters typically involves introducing a low frequency background trend which can introduce systematic error. Interpreting AVO reflectivity attributes avoids both these shortcomings. However, these reflectivity attributes are band-limited making comparison to the well control more difficult. Further, the reflectivity attributes are local perturbations and the interpretation must be with respect to some other layer. For example, the reflectivity attributes for a particular horizon are crossplotted and then compared to well templates for that horizon. Rutherford and William's classification provides a framework for this type of interpretation and will be used in parts of this thesis.

Fluid stack

Smith and Gidlow (1987) and Fatti et al. (1994) provide another methodology to interpret reflectivity data. Their starting point is the mudrock relationship. They look for local perturbations from the mudrock trend and attribute them to changes in fluid saturation, hence they call this the fluid stack. They reformulate the mudrock

relationship (equation 1.34) in terms of reflectivity to derive this. In water saturated sediments equation (1.34) written in terms of reflectivity is

$$R_\alpha = m\bar{\gamma}R_\beta, \tag{1.38}$$

where m and $\bar{\gamma}$ are the mudrock slope and average β/α ratio across the interface respectively. This follows from taking the differential of equation (1.34) and then dividing by the P-wave velocity. Equation (1.38) suggests that in clastics the S-wave velocity reflectivity is a scaled version of the P-wave velocity reflectivity. This scaling factor changes as a function of γ. This relation holds for brine filled clastics. However, if the fluid of the clastic changes, for example from a gas to brine such as shown in Figure 1.12, then equation (1.38) no longer holds. The P-wave velocity reflectivity responds to the change in compressibility as a result of the change of fluids and rock matrix. The rigidity responds to changes in the rock matrix only. By subtracting a scaled version of the S-wave velocity reflectivity from the P-wave velocity reflectivity, changes in the rock matrix cancel out leaving the difference due to changes in fluid content. Smith and Gidlow defined the difference of the left and right hand sides of equation (1.38) as the fluid stack

$$\Delta F = R_\alpha - \frac{1}{m\bar{\gamma}}R_\beta. \tag{1.39}$$

The fluid stack is approximately zero for brine filled rocks since the equality (equation 1.38) is honored and the S-wave velocity reflectivity is just a scaled version of the P-wave velocity reflectivity. When the fluid changes the difference is non-zero. Going from a wet sand to a gas sand results in a negative response, while going from a gas sand to a wet sand results in a positive response.

1.5 Outline of thesis

This introductory chapter reviewed the objectives, theoretical basis, data preconditioning, and interpretation of AVO. In addition, certain rock physics relationships were reviewed which are useful to help constrain the AVO inversion problem. Chapter 2 explores the stability of the AVO inversion in the presence of random noise. Using a Bayesian formulation, I develop several quality controls that allow an interpreter to judge the validity of the AVO estimates. These quality controls are tested

and verified on synthetic and real data.

Using these quality control displays, I show that for noise levels typical of most seismic acquisition, the linearized three-term AVO inversion problem is unstable. Chapter 2 reviews various methodologies in the literature used to help constrain or stabilize the problem. In Chapter 3, I introduce a new methodology to constrain the three-term AVO inversion problem using probabilistic constraints following a Bayesian framework. In this methodology, the constraints dominate the solution when the data are noisy or ill-conditioned due to the acquisition geometry. Conversely, if the seismic data are adequate, the AVO inversion estimates are obtained primarily from the seismic data, providing additional information about the geology independent from the constraints. Again, quality controls are introduced to help the interpreter understand the validity and usefulness of the results. The algorithm is demonstrated on both synthetic and real data.

The uncertainty estimates demonstrated in both Chapter 2 and 3 are both somewhat overly optimistic. They only show how uncertainty in the data will affect the estimates. However, uncertainty also exists in the linear model that is used for the AVO inversion, namely the β/α ratio and the mapping from offset to angle of incidence. Chapter 4 explores errors associated with both of these parameters. For each of these, the error may be separated into a low frequency background trend and a high frequency error component. I show, provided that the low frequency background trend is accurate, the error in the AVO parameter estimates are negligible. However, if errors in the background trend exist, there will be systematic errors in the AVO estimates.

Chapter 5 examines the theoretical error introduced by NMO stretch and offset dependent tuning. Analytic relationships are developed, extending the work of Dong (1999), to predict how the error of the different Rutherford and Williams (1989) AVO classes will behave. These predictions are tested and verified using synthetic data. I show that errors due to NMO stretch and offset dependent tuning are significant for Class III and IV gas sand anomalies where large offsets are used. Reflectors following the mudrock trend, Class I, and II anomalies are not significantly affected by these distortions. Further, these distortions are insignificant if the maximum angle of incidence used in the AVO inversion is less than 30°.

In Chapter 6, I introduce a two-term AVO waveform inversion that incorporates NMO stretch and offset dependent tuning. The output of this inversion may be

in terms of reflectivity or impedance. Similarities exist with the AVO waveform inversion developed by Simmons and Backus (1996) since both algorithms incorporate offset dependent tuning and NMO stretch. The Simmons and Backus algorithm uses the actual Zoeppritz equations, while I use a linearized approximation. I prefer the linearized approximation, since it is parameterized in terms of relative perturbations (reflectivity) which is important for the constraints and optimization method I employ. I incorporate probabilistic constraints similar to that of Chapter 3. Simmons and Backus (1996) do not constrain the problem, but do a coordinate transform to try and achieve the same thing. I believe my constraint methodology is more rigorous. Further, I believe my algorithm is superior to that of Buland and Omre (2003) in that they do not incorporate NMO stretch and tuning. Like mine, their algorithm is based on a linearized AVO approximation. They have some interesting constraints describing the inter-relationship between different time samples which I would like to investigate in the future. This algorithm is tested on synthetic data constructed to highlight problems due to NMO stretch and offset dependent tuning. Real data results are also shown.

In Chapter 7, the algorithm of the previous chapter is modified to create a stretch-free NMO algorithm. By performing stretch-free NMO the seismic data may be inverted using traditional AVO techniques. Rather than including the NMO stretch and offset dependent tuning in the AVO waveform inversion, the stretch-free NMO preconditions the data so these distortions are addressed. The algorithm is demonstrated and tested on both synthetic and real data. The AVO inversion results of this are better than using NMO corrected gathers, but not as good as that of the AVO waveform inversion.

In Chapter 8 the two-term AVO waveform inversion is extended to three terms. Originally, the algorithm was developed using only two terms for stability concerns. Probabilistic constraints similar to Chapter 3 and 6 are incorporated to help stabilize the problem. In addition, long-tailed misfit distributions are explored to help make the inversion more robust in the presence of coherent noise. An example is shown, demonstrating that the long-tailed misfit distributions improve the AVO estimates in the presence of coherent noise such as multiples. Further, the influence of residual statics is explored, along with a modified formulation which solves for trim statics as part of the AVO waveform inversion. Once again the algorithm is demonstrated on both synthetic and real seismic data.

Chapter 2

AVO inversion and uncertainty analysis

2.1 Introduction

In this chapter the linearized AVO inversion problem is further investigated. First, common parameterizations and approximations used in the literature (Aki and Richards, 1980; Shuey, 1985; Smith and Gidlow, 1987; Gidlow et. al., 1992; Gray et al., 1999) are reviewed with the goal of (1) developing the vocabulary needed for future chapters and (2) showing the inter-relationships and commonality between the different approaches. It is shown that given the background velocity trend, all the reflectivity attributes derived from AVO are just linear combinations of three basic reflectivity attributes. Having established the interrelatedness of the different AVO approaches in the literature, the generalized linearized AVO inverse problem is investigated using Bayes' theorem. Bayes' theorem provides a framework within which the probability of the variables of interest from uncertain data and *a priori* information is obtained. This probabilistic prediction provides a natural way of understanding the uncertainty of the problem. This chapter restricts itself to uncertainties in the data and in particular the assumption that the noise is uniform, uncorrelated and Gaussian. Under these assumptions the parameter covariance matrix may be used to investigate the uncertainty in the problem (Tarantola, 1987). Theoretical error, or error that arises due to error in the linear operator, is investigated in Chapter 4.

The basic approach follows Downton and Lines (2001b) where Bayes' theorem is used to investigate the uncertainty in AVO parameter estimates in the linearized AVO problem. The reader is referred to Riedel et al. (2003) for a Bayesian uncertainty analysis of the nonlinear AVO inversion problem. In that paper, the analysis is performed on a single interface assuming the parameters are known for the first

interface, so this approach is difficult to generalize for the typical problem found in exploration seismology. The AVO uncertainty estimates in this chapter relate to the uncertainty in the elastic parameters In contrast, Houck (2002) proposes a Bayesian approach to estimate the uncertainty in fluid and lithology estimates using AVO. I believe this is actually a two stage problem, as there are two distinct inversions, each with their own data model, and each have their own issues of stability and uniqueness. First there is the elastic parameter estimation problem and then the rock and fluid property estimation problem. Loures and Moraes (2002) provides uncertainty estimates of rock property values following this methodology.

This chapter provides an approach to generate uncertainty estimates for the first part of this problem, that of the band-limited elastic parameters. Various uncertainty displays are developed including an estimate of the prestack seismic noise level in the data set and a stability section indicating how stable a particular AVO reflectivity estimate is for a particular acquisition geometry. From these displays the standard deviation of the parameter estimate and the fractional uncertainty can be calculated. These estimates of uncertainty are tested on synthetic and real data with results consistent with expectations. The primary factors influencing the reliability due to random noise are: the prestack signal-to-noise ratio, the fold, and the range of angles available for the AVO inversion.

The three-parameter AVO inversion problem is ill-conditioned for most common acquisition geometries and noise levels necessitate the use of constraints to help stabilize the problem. In this chapter constraints which reduce the number of variables from three to two are investigated. These "hard constraints" are defined to be some linear relationship linking two or more variables so that one these variables can be eliminated. By constraining the solution, the uncertainty in the parameter estimates is reduced but at the expense of introducing theoretical error.

The basic outline for this chapter is as follows. First, the linearized three-term AVO inversion problem is reviewed. It is noted that the problem is unstable in the presence of noise. Various parameterizations and approximations of the linearized AVO inversion problem from the literature are reviewed. For stability reasons these approximations generally reduce the number of independent variables solved from three to two. Bayes' theorem is then discussed as a way of studying the probability and uncertainty of the elastic parameter estimates for the two-term linearized AVO approximation in the presence of random noise. Note, the three parameter problem

is investigated in Chapter 3, using probabilistic constraints. Under the assumption of uniform, uncorrelated random noise the covariance matrix is introduced as a simple means to estimate the uncertainty of the AVO parameter estimates. To validate this approach, a modeling study is performed to compare the estimates of uncertainty with the actual reliability. Lastly, several data examples are shown where the uncertainty estimates are used to help identify reliability issues in the AVO parameter estimates. Both the modeling study and the data examples show that the uncertainty estimates reasonably predict the error in the inversion giving the interpreter tools to help quantify the reliability of their AVO interpretation.

2.2 Amplitude variation with offset theory

Recall from Chapter 1 (equation 1.25), that linearized AVO model for M offsets may be written as

$$\begin{bmatrix} R(\bar{\theta}_1) \\ \vdots \\ R(\bar{\theta}_M) \end{bmatrix} \approx \begin{bmatrix} \sec^2\bar{\theta}_1 & -8\bar{\gamma}^2\sin_1^2\bar{\theta} & \left(1-4\bar{\gamma}^2\sin^2\bar{\theta}_1\right) \\ \vdots & \vdots & \vdots \\ \sec^2\bar{\theta}_M & -8\bar{\gamma}^2\sin^2\bar{\theta}_M & \left(1-4\bar{\gamma}^2\sin^2\bar{\theta}_M\right) \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{2.1}$$

where $\bar{\theta}$ is the average angle of incidence, $\bar{\gamma}$ is the S-wave to P-wave velocity ratio across the interface. Equation (2.1) may be written more succinctly in matrix notation as

$$\mathbf{d} \approx \mathbf{G}\mathbf{x}, \tag{2.2}$$

where $\mathbf{G}$ is the linear operator, $\mathbf{d}$ the input data vector (offset dependent reflectivity), and $\mathbf{x}$ the unknown parameter vector containing the P-wave and S-wave velocity, and density reflectivity. The matrix $\mathbf{G}$ is completely determined by the acquisition geometry and the background velocity fields. One of the advantages of dealing with linear inverse problems is that the linear operator may be analyzed to determine the stability of the problem. Stability in this context refers to how reliable the estimates are in the presence of a small amount of noise. The problem is considered unstable if a small amount of noise leads to large uncertainties in the estimates. One way to understand this is to look at the condition number of the matrix $\mathbf{G}$. The condition number K is defined (Borse, 1997) as

$$K \equiv \left\|\mathbf{G}^{-1}\right\| \|\mathbf{G}\|, \tag{2.3}$$

where $||\mathbf{G}||$ and $||\mathbf{G}^{-1}||$ are the norm of $\mathbf{G}$ and its inverse. For the Euclidean norm this is equal to the ratio of largest eigenvalue λ_1 to the smallest eigenvalue λ_N. Depending on the noise level, problems become ill-conditioned when the condition number gets into the hundreds. Since the condition number is solely dependent on matrix $\mathbf{G}$, the condition number is only a function of the acquisition geometry and the background velocity of the AVO problem. Through the condition number, the stability of the AVO inverse problem can be investigated for certain ideal geometries and background velocities. Figure 2.1 shows the condition number where the range of angles used in the AVO inversion is varied, while fold is held constant. The horizontal axis shows the maximum angle used in the inversion, assuming the minimum angle used is 0. In Figure 2.1 when the maximum angle used in the inversion is under 20°, the condition number is over 10^6. This is extremely unstable. Only by including angles out to 45° does the condition number go below 1000.

2.3 Rearrangements of the linearized Zoeppritz approximation

Because of these stability concerns, different authors have rearranged the Aki and Richards linearized approximation of the Zoeppritz equations to solve for different reflectivity attributes. Generally, these rearrangements either incorporate constraints or introduce approximations to make the problem more stable. As important, these new reflectivity attributes are designed to highlight geologic information and anomalies in a new manner, ideally identifying the objective more clearly.

2.3.1 Shuey equation

Shuey (1985) rearranged the Aki and Richards equation (1.6) in terms of the polynomial series

$$R(\theta) = A + B\sin^2\theta + C\sin^2\theta\tan^2\theta \tag{2.4}$$

where

$$A = \frac{1}{2}\left(\frac{\Delta\alpha}{\bar{\alpha}} + \frac{\Delta\rho}{\bar{\rho}}\right), \tag{2.5}$$

$$B = \left[B_0 A + \frac{\Delta\nu}{(1-\bar{\nu}^2)}\right], \tag{2.6}$$

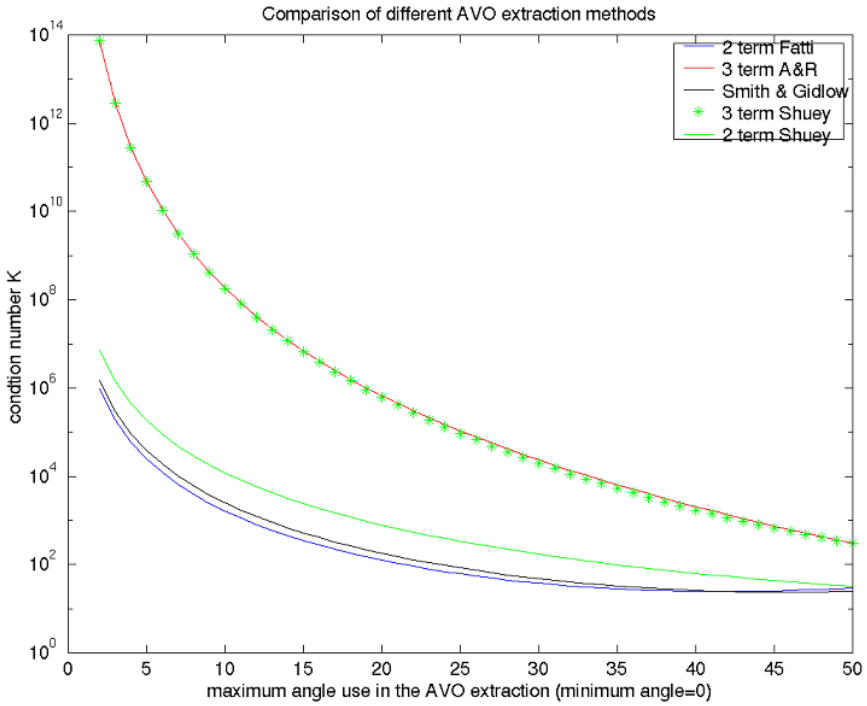

Figure 2.1: The condition number plotted for a number of simulated AVO inversions. The fold is held constant while the maximum angle used for the inversion is varied. The legend indicates the AVO approximation used in calculating the condition number. Note that the condition number improves as the maximum angle used in the inversion increases.

$$B_0 = D - 2\,(1+D)\,\frac{1-2\bar{\nu}}{1-\bar{\nu}}, \tag{2.7}$$

$$D = \frac{\frac{\Delta\alpha}{\bar{\alpha}}}{\frac{\Delta\alpha}{\bar{\alpha}} + \frac{\Delta\rho}{\bar{\rho}}}, \tag{2.8}$$

and

$$C = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}}, \tag{2.9}$$

where $\Delta\nu$ is the change and $\bar{\nu}$ is the average Poisson ratio across the interface. The advantage of this formulation is that each reflectivity attribute is associated with an angle term of different order. The background P-wave velocity must be known in order to do the ray-tracing to calculate the angle of incidence. However, the background S-wave velocity is not required to be known *a priori,* potentially reducing the theoretical error in the problem. For this reason, many practitioners prefer using this formulation over the Aki and Richards formulation.

However, the definition of the B parameter is awkward and theoretically suspect since it includes nonlinear coefficients after already having made a linear approximation. Swan (1993) redefined the B parameter as

$$B = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}} - 2\bar{\gamma}^2\left(2\frac{\Delta\beta}{\bar{\beta}} + \frac{\Delta\rho}{\bar{\rho}}\right), \tag{2.10}$$

which is the one most commonly used in the literature. This definition has the additional advantage that under the assumption $\bar{\gamma} = \frac{1}{2}$ there is a simple relationship between the B term and P-wave and S-wave impedance reflectivity R_p and R_s namely

$$B = R_p - 2R_s. \tag{2.11}$$

Thus, under the assumption that $\bar{\gamma} = \frac{1}{2}$, the S-wave impedance reflectivity can be calculated from the A and B terms

$$R_s = \frac{A - B}{2} \tag{2.12}$$

From these definitions a transform matrix (equation B.1) may be constructed which

links the Aki and Richards parameterization and the "Shuey" parameterization

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -8\bar{\gamma}^2 & -4\bar{\gamma}^2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{2.13}$$

Often for stability reasons the last term in equation (2.4) is truncated

$$R(\theta) = A + B\sin^2\theta \tag{2.14}$$

so only the first two terms are solved for. The form of this approximation is equivalent to that used by Wiggins et al. (1983) but the actual definitions of the parameters are different. Figure 2.1 shows the condition number for both the two- and three-term Shuey equations. The three-term condition number is the same as the Aki and Richards formulation since it is simply a transform and the eigenvalues are the same. The two-term Shuey equation shows a dramatic drop in the condition number. Physically, the two-term Shuey equation (2.4) describes a line with intercept A and slope or gradient B. Unfortunately truncating the 3rd term in the series (2.4) introduces significant theoretical error which biases the estimate of the gradient B.

2.3.2 Smith and Gidlow

Truncating the third term of the Shuey equation (2.4) is equivalent to constraining the P-wave velocity reflectivity to zero

$$C = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}} = 0, \tag{2.15}$$

which is clearly not appropriate in a variable velocity earth. This is what leads to the bias in the estimate of the gradient B. Smith and Gidlow (1987) suggested using a more geologically realistic constraint to improve the stability of the inversion. They suggested using the Gardner et al. (1974) empirical relationship linking density and P-wave velocity to establish a relationship linking the density and P-wave velocity reflectivity. By taking the derivative of equation (1.37) with respect to α and then rearranging the result, the Gardner relationship can be expressed in terms of reflectivity

$$R_d = gR_\alpha, \tag{2.16}$$

This suggests that density reflectivity should be a scaled version of the velocity reflectivity where the scalar is a function of lithology. This thesis calls this relation (equation 2.16) the Gardner reflectivity relation or the Gardner constraint, depending on the context. From this a transform matrix can be constructed

$$\begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ g & 0 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \end{bmatrix}, \tag{2.17}$$

which when introduced into the Aki and Richards equation (2.1) generates the Smith and Gidlow equation

$$\left[R(\bar{\theta})\right] = \begin{bmatrix} \sec^2\bar{\theta} + \left(1 - 4\bar{\gamma}^2\sin^2\bar{\theta}\right) g & -8\bar{\gamma}^2\sin^2\bar{\theta} \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \end{bmatrix}. \tag{2.18}$$

The unknown parameters solved for in this formulation are the P-wave and S-wave velocity reflectivity. In section (2.4.3) it is shown that the velocity reflectivity has a much greater uncertainty than the impedance reflectivity. In this inversion formulation, equation (2.16) is used to make the inversion more stable. Figure 2.1 shows that the Smith and Gidlow equation has a significantly smaller condition number than the Aki and Richards equation, and even smaller than the two-term Shuey equation. However, the accuracy of the estimate is dependent on how accurate the empirical relationship equation (2.16) is. In cases where the density and P-wave velocity reflectivity are poorly correlated or the parameter g is inaccurately known, there is the potential for large error. Note that the P-wave impedance reflectivity does not have this sensitivity built into it, and as a result often gives better estimates when compared to the P-wave velocity reflectivity.

In addition, this method and all the subsequent methods discussed, require more *a priori* information than that of the Shuey equation, in that they require knowledge of the background S-wave velocity. This too can lead to systematic errors in the estimates and is explored further in Chapter 4.

2.3.3 Gidlow equation

The Gidlow et al. (1992) equation follows from rearranging the Aki and Richards equation (2.1) in terms of impedance reflectivity

$$\left[R(\bar{\theta})\right] = \begin{bmatrix} \sec^2\bar{\theta} & -8\bar{\gamma}^2\sin^2\bar{\theta} & 4\bar{\gamma}^2\sin^2\bar{\theta} - \tan^2\bar{\theta} \end{bmatrix} \begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix}, \tag{2.19}$$

where R_p and R_s are the P-wave and S-wave impedance reflectivity respectively. This follows from multiplying the Aki and Richards equation (2.1) by the transform matrix (B.10) found in Appendix A. Again, for stability reasons often only the first two terms are solved for

$$\left[R(\bar{\theta})\right] = \begin{bmatrix} \sec^2\bar{\theta} & -8\bar{\gamma}^2\sin^2\bar{\theta} \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}. \tag{2.20}$$

Figure 2.1 shows the resulting improvement in condition number compared to the three-term equation and other approximations.

Note that the original publication of this work was in a workshop abstract which is not widely available. Fatti et al. (1994) published this equation in a much more widely available journal and hence this equation is often know as the Fatti equation.

Fluid stack based on Gidlow equation

Following a similar derivation to the fluid stack based on the velocity fields (Section 1.4.2), the fluid stack based on impedance reflectivities is defined by Gidlow et al. (1992) as

$$R_g = R_p - m\bar{\gamma}R_s. \tag{2.21}$$

Equation (2.21) responds in a similar fashion as the fluid stack. This may also be written as a transform matrix

$$\begin{bmatrix} R_p \\ R_g \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -m\bar{\gamma} \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}. \tag{2.22}$$

2.3.4 Other formulations

Various other papers define other linearized reflectivity attributes. For example, Gray et al. (1999) solve for λ reflectivity, μ reflectivity, and bulk modulus K reflectivity. Note only two of three of these attributes are independent, but all are shown for completeness. Appendix B derives these, resulting in the transform matrix

$$\begin{bmatrix} R_\lambda \\ R_\mu \\ R_K \end{bmatrix} = \begin{bmatrix} \frac{2}{1-2\bar{\gamma}^2} & -\frac{4\bar{\gamma}^2}{1-2\bar{\gamma}^2} & 1 \\ 0 & 2 & 1 \\ \frac{6}{3-4\bar{\gamma}^2.} & -\frac{8\bar{\gamma}^2}{3-4\bar{\gamma}^2} & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{2.23}$$

Any of these reflectivity attributes may be calculated by first solving the Aki and Richards equation (2.1), and then transforming to the desired reflectivity attributes using the suitable transform matrix. These may be modified to $\lambda\rho$ and $\mu\rho$ reflectivity with (Appendix B)

$$\begin{bmatrix} R_{\lambda\rho} \\ R_{\mu\rho} \end{bmatrix} = \begin{bmatrix} \frac{2}{1-2\bar{\gamma}^2} & -\frac{4\bar{\gamma}^2}{1-2\bar{\gamma}^2} & 2 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{2.24}$$

Reflectivity attributes that include density, such as impedance, tend to be more stable than ones that don't, such as velocity. Shuey defined the delta-Poisson ratio reflectivity R_σ

$$R_\sigma = \frac{d\nu}{(1-\nu)^2}, \tag{2.25}$$

which may be transformed to using (equation B.20)

$$R_\sigma = 8\bar{\gamma}^2 \left[R_\alpha - R_\beta\right], \tag{2.26}$$

Verm and Hilterman (1995) use the delta-Poisson reflectivity attribute as a lithology indicator in the Gulf coast with the assumption $\alpha/\beta = 2$. The Poisson reflectivity may also be defined in a manner more consistent with the previous reflectivity attributes (Appendix B, equation B.26).

$$R_\nu = \frac{1}{\frac{3}{2} - \bar{\gamma}^2 - \frac{1}{2\bar{\gamma}^2}} \left[R_\alpha - R_\beta\right]. \tag{2.27}$$

Summarizing these transform matrices results in

$$\begin{bmatrix} A \\ B \\ C \\ R_p \\ R_s \\ R_\gamma \\ R_\sigma \\ R_\nu \\ R_{\lambda\rho} \\ R_{\mu\rho} \\ R_k \\ R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -8\bar{\gamma}^2 & -4\bar{\gamma}^2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \\ 8\bar{\gamma}^2 & -8\bar{\gamma}^2 & 0 \\ \frac{1}{\left(\frac{3}{2}-\bar{\gamma}^2-\frac{1}{2\bar{\gamma}^2}\right)} & \frac{-1}{\left(\frac{3}{2}-\bar{\gamma}^2-\frac{1}{2\bar{\gamma}^2}\right)} & 0 \\ \frac{2}{1-2\bar{\gamma}^2} & 4\frac{\bar{\gamma}^2}{2\bar{\gamma}^2-1} & 2 \\ 0 & 2 & 2 \\ \frac{6}{3-4\bar{\gamma}^2.} & -\frac{8\bar{\gamma}^2}{3-4\bar{\gamma}^2} & 1 \\ \frac{2}{1-2\bar{\gamma}^2} & -\frac{4\bar{\gamma}^2}{1-2\bar{\gamma}^2} & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{2.28}$$

This transform matrix allows one to work with any three independent reflectivity attributes one desires. Note the resulting reflectivity attributes on the left hand side of equation (2.28) are all just linear combinations of velocity and density reflectivity. Therefore there are only 3 independent reflectivity attributes. Some reflectivity attributes listed here are just scalar multiples of each other such as R_γ and R_σ. Note, that in order to do the majority of these transforms $\bar{\gamma}$ must be known. Further, note that a linear approximation is made in the derivation of these reflectivity attributes, which is questionable for some of these attributes. For example, in unconsolidated sands, the change in λ, $\Delta\lambda$ can almost be as large as the average $\bar{\lambda}$ resulting in reflectivity attribute values approaching 1.

2.4 Bayes' theorem and uncertainty

Having introduced various rearrangements of the three-term linearized AVO inversion and two-term approximations we are now in a position to discuss the uncertainty of these parameter estimates. Bayes' theorem is used to develop these uncertainty estimates. It provides a theoretical framework to make probabilistic estimates of the unknown reflectivity attributes $\mathbf{x}$ from uncertain data and *a priori* information. The

resulting probabilistic parameter estimates are called the *Posterior* Probability Distribution function (*PPDF*). The *PPDF* written symbolically as $P(\mathbf{x}|\mathbf{d},I)$ indicates the probability of the parameter vector $\mathbf{x}$ given the data vector $\mathbf{d}$ (offset dependent reflectivity) and information I. Bayes' theorem, which can be expressed as

$$P(\mathbf{x}|\mathbf{d},I) = \frac{P(\mathbf{d}|\mathbf{x},I)\,P(\mathbf{x}|I)}{P(\mathbf{d}|I)}, \tag{2.29}$$

calculates the *PPDF* from the likelihood function $P(\mathbf{d}|\mathbf{x},I)$ and *a priori* probability function $P(\mathbf{x}|I)$. The denominator $P(\mathbf{d}|I)$ is a normalization function which may be ignored if only the shape of the *PPDF* is of interest

$$P(\mathbf{x}|\mathbf{d},I) \propto P(\mathbf{d}|\mathbf{x},I)\,P(\mathbf{x}|I)\,. \tag{2.30}$$

The most likely estimate occurs at the maximum of the *PPDF*. The uncertainty of the parameter estimate is proportional to the width of the *PPDF*.

To simplify the analysis of equation (2.30) it is initially assumed that the prior probability distributions are uniform. In this case, the Bayesian inversion is equivalent to maximum likelihood inversion. To simplify the analysis further, it is assumed the noise is independent and Gaussian. In this case, the likelihood function may be written as (Sivia, 1996)

$$P(\mathbf{d}|\mathbf{x},I) \propto \sigma_N^{-N} \exp\left[-\frac{\sum_{m=1}^{M}\left(\sum_{n=1}^{N} G_{mn}x_n - d_m\right)^2}{2\sigma_N^2}\right], \tag{2.31}$$

where σ_N^2 is the variance of the noise, M is the number of data, and N is the number of parameters used in the inversion.

2.4.1 Unconstrained AVO inversion

If we include no *a priori* information this is equivalent to employing a uniform distribution so that $P(\mathbf{x}|\mathbf{d},I) \propto \mathbf{P}(\mathbf{d}|\mathbf{x},I)$ and

$$P(\mathbf{x}|\mathbf{d},I) \propto \sigma_N^{-N} \exp\left[-\frac{\sum_{m=1}^{M}\left(\sum_{n=1}^{N} G_{mn}x_n - d_m\right)^2}{2\sigma_N^2}\right]. \tag{2.32}$$

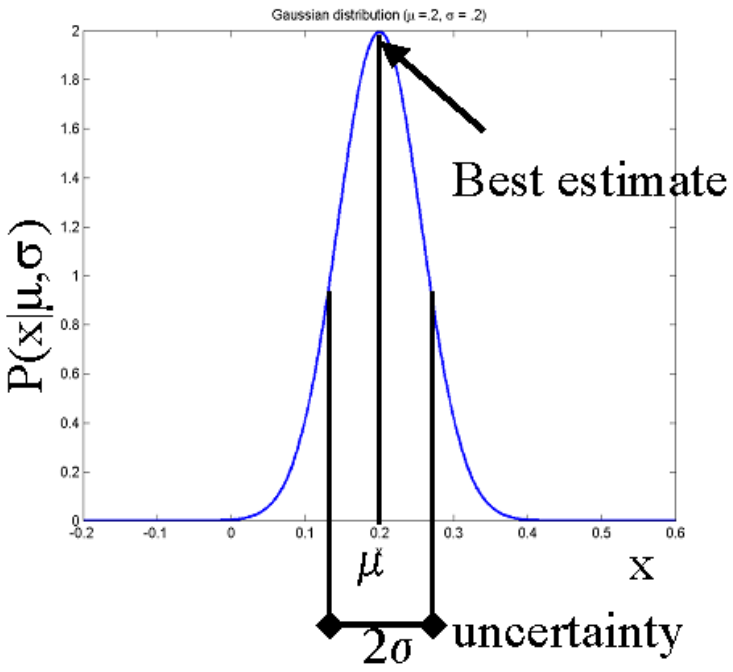

Figure 2.2: One dimension Gaussian distribution. The best estimate occurs where the probability is maximum. The uncertainty is related to the width of the distribution.

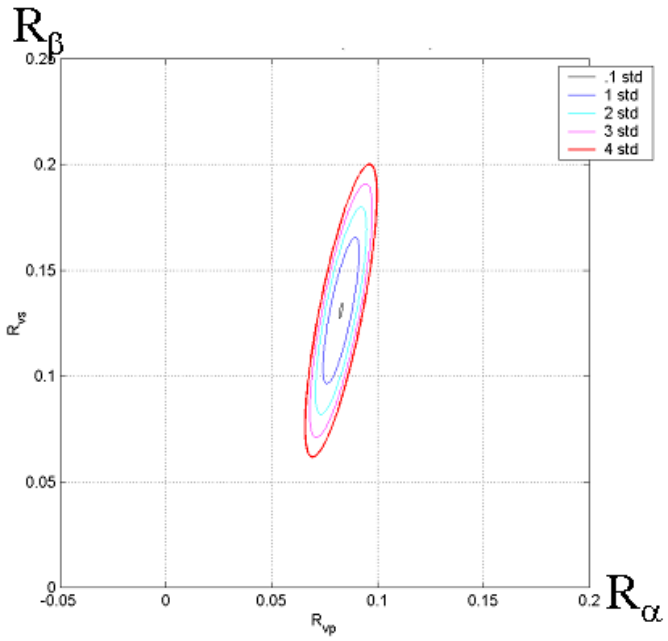

Figure 2.3: Bivariate Gaussian distribution. Each ellipse represents an equiprobable solution whose standard deviation is indicated by the legend. The best estimate occurs on the smallest ellipse.

Under these assumptions the optimization problem is equivalent to least squares with the most likely estimate occurring at

$$\hat{\mathbf{x}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}. \tag{2.33}$$

where "$\hat{}$ " indicates the variable $\mathbf{x}$ is an estimate. For AVO inversion, because of the small number of parameters solved for, it is possible to visualize the PDF generated by equation (2.32). If the parameter vector $\mathbf{x}$ has only one element, such as in the case of the stack model, the PDF is a Gaussian function (Figure 2.2). The most likely solution is located at the maximum of the probability distribution. The uncertainty is proportional to the width of the distribution. Since this is a Gaussian distribution, standard statistical parameterizations such as standard deviation or variance can be used to characterize the uncertainty.

If the parameter vector $\mathbf{x}$ has two elements, the PDF is a bivariate Gaussian function and an equiprobable solution is an ellipse. This is the case for the two-term

Gidlow et al. (1992), Shuey (1985), and the Smith and Gidlow (1987) approximations. Figure 2.3 shows the probability distribution solved for the Smith and Gidlow approximation, for a hypothetical reflector with a signal-to-noise ratio of 2:1. Each contour represents an equiprobable solution. In this example, the most likely solution occurs in the middle of the smallest contour ellipse at $R_\alpha = 0.06$ and $R_\beta = 0.13$ which is the maximum probability.

The bivariate Gaussian distribution

$$P(\mathbf{x}|I) = \frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}-\langle\mathbf{x}\rangle\right)^T \mathbf{C}_\mathbf{x}^{-1}\left(\mathbf{x}-\langle\mathbf{x}\rangle\right)\right]}{(2\pi)^{\frac{3}{2}}\sqrt{\det|\mathbf{C}_\mathbf{x}|}^3}, \tag{2.34}$$

is parameterized by the 2×2 covariance matrix

$$\mathbf{C}_\mathbf{x} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R_\beta} \\ \sigma_{R_\alpha R_\beta} & \sigma^2_{R_\beta} \end{bmatrix}. \tag{2.35}$$

The diagonal elements of the parameter covariance matrix are the variance of each of the parameters estimates. The off diagonal elements describe the amount of correlation between the variables. The standard deviation, which is the square root of the variance, is a measure of reliability of each variable.

Appendix C shows that for least squares inversion that the uncertainty of the parameter estimates may be estimated by (Menke, 1994)

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \mathbf{G}^{-g}\mathbf{C}_d\left(\mathbf{G}^{-g}\right)^T, \tag{2.36}$$

where $\mathbf{G}^{-g}$ is the generalized inverse of $\mathbf{G}$, and $\mathbf{C}_d$ is the data covariance matrix. The diagonal term contains the variance of noise for each datum while the off-diagonal elements contain the degree of correlation between the data. Assuming uncorrelated uniform noise this reduces to (equation C.34)

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \sigma^2_N\left[\mathbf{G}^T\mathbf{G}\right]^{-1}. \tag{2.37}$$

Equation (2.37) suggests that the parameter covariance matrix for the estimated variables is only a function of the prestack noise variance σ^2_N, and the linear operator

G. The prestack noise variance may be estimated from the misfit of the data

$$\hat{\sigma}_N^2 = \frac{\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}}{(M-1)}, \tag{2.38}$$

where

$$\boldsymbol{\varepsilon} = \mathbf{G}\hat{\mathbf{x}} - \mathbf{d}. \tag{2.39}$$

For the two-term Gidlow equation this leads to the analytic relationship (equation D.20)

$$\begin{bmatrix} \tilde{\sigma}_{Rp}^2 & \tilde{\sigma}_{RpRs} \\ \tilde{\sigma}_{RpRs} & \tilde{\sigma}_{Rs}^2 \end{bmatrix} = \frac{\hat{\sigma}_N^2}{D} \begin{bmatrix} \sum_{m=1}^{M} \sin^4 \bar{\theta}_m & \frac{-1}{8\bar{\gamma}^2} \sum_{m=1}^{M} \tan^2 \bar{\theta}_m \\ \frac{-1}{8\bar{\gamma}^2} \sum_{m=1}^{M} \tan^2 \bar{\theta}_m & \frac{1}{64\bar{\gamma}^4} \sum_{m=1}^{M} \frac{1}{\cos^4 \bar{\theta}_m} \end{bmatrix}, \tag{2.40}$$

where

$$D = \left(\sum_{j=1}^{M} \sec^4 \bar{\theta}_j \right) \sum_{m=1}^{M} \sin^4 \bar{\theta}_m - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m \right)^2. \tag{2.41}$$

This equation is studied in greater detail in section (2.4.3).

If we are interested in the uncertainty of the R_α estimate then we need to know the one dimensional probability distribution for R_α. In calculating this we have no interest in the S-wave velocity reflectivity. The S-wave velocity reflectivity may be eliminated by marginalizing the Bivariate Gaussian probability. Marginalization consists of integrating over the range of all possible values for that particular variable

$$P(R_\alpha|I) = \int P(R_\alpha, R_\beta|I) dR_\beta. \tag{2.42}$$

The result of marginalizing a bivariate Gaussian probability function is a Gaussian probability distribution. The uncertainty is related to the total width of the ellipse projected onto the appropriate axis. Figure 2.4 shows this projection for both the P-wave and S-wave velocity reflectivity. It is evident that S-wave velocity reflectivity has greater uncertainty than the P-wave velocity reflectivity.

For the case of equation (2.1) there are 3 parameters so the PDF is a multivariate Gaussian function where an equiprobable solution is an ellipsoid. Typically the ellipsoid is quite elongated (Figure 2.5). This is a result of the small eigenvalue associated with the introduction of the 3rd term. Non-physical solutions, such as those in which the magnitude of the reflectivity is greater than 1, are evident. Figure 2.6 shows a

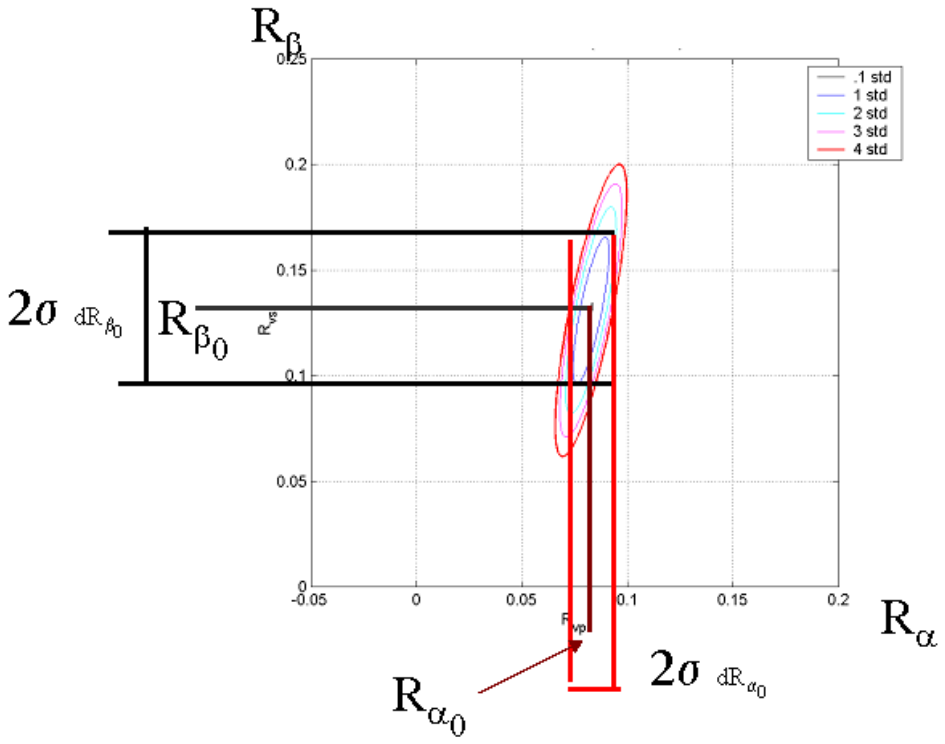

Figure 2.4: Uncertainty of P- and S-wave velocity reflectivity. The uncertainty for each variable is the marginalized one dimension distribution for that variable. This is a Gaussian distribution whose width is characterized by its standard deviation.

view of Figure 2.5 in which only two of the three axes are evident. In this figure the ellipsoid lies along the diagonal and is quite elongated. This implies that the uncertainty for both the S-wave velocity reflectivity and density reflectivity is large. This is also true for the P-wave velocity reflectivity. The reflectivity attributes may be transformed to other attributes, such as impedance reflectivity, each with different uncertainty. Figure 2.7 shows one face of the cube, for density and S-wave impedance reflectivity. Note that even though the misfit ellipse is still elongated in the density direction, there is little variance along the S-wave impedance reflectivity axis. Thus, the estimate of the S-wave impedance reflectivity is more certain than the S-wave velocity reflectivity. It seems that reflectivity attributes which combine density are more reliable than those that do not. Two examples of this are impedance versus velocity reflectivity and $\lambda\rho$ versus λ reflectivity.

Figure 2.8 shows a comparison between the misfit ellipsoid for the density and S-wave impedance reflectivity parameterization, and the density and fluid stack reflectivity parameterization. Note that the fluid stack has less variance than the S-wave impedance reflectivity.

2.4.2 Constrained two-term AVO inversion

Up to this point it has been assumed that unknown reflectivity attributes are uniformly distributed. As seen in the last section, this allows for non-physical solutions. Constraints can be used to eliminate these. One way of doing this is to place bounds or limits on the *a priori* distributions so the reflectivity magnitude must be less than one. This reduces the solution space and uncertainty. However, for noise levels typical in real data, the parameter estimates, even with constraints, are still not accurate enough to make reliable predictions. Additional constraints must be employed to reduce the reflectivity attribute uncertainty to acceptable levels. This is at the expense of potentially introducing into the inversion theoretical error arising from incorrect assumptions.

For example, Smith and Gidlow (1987) used the Gardner constraint (equation 2.16) which defines a plane in the 3D solution space. Using Bayes' theorem, the solution is the intersection of this plane and the 3D ellipsoid (Figure 2.9). This constraint greatly reduces the uncertainty but at the expense of effectively removing one of the variables. The density reflectivity is no longer an independent variable. The inverse problem now has only two parameters with the *PPDF* of these variables

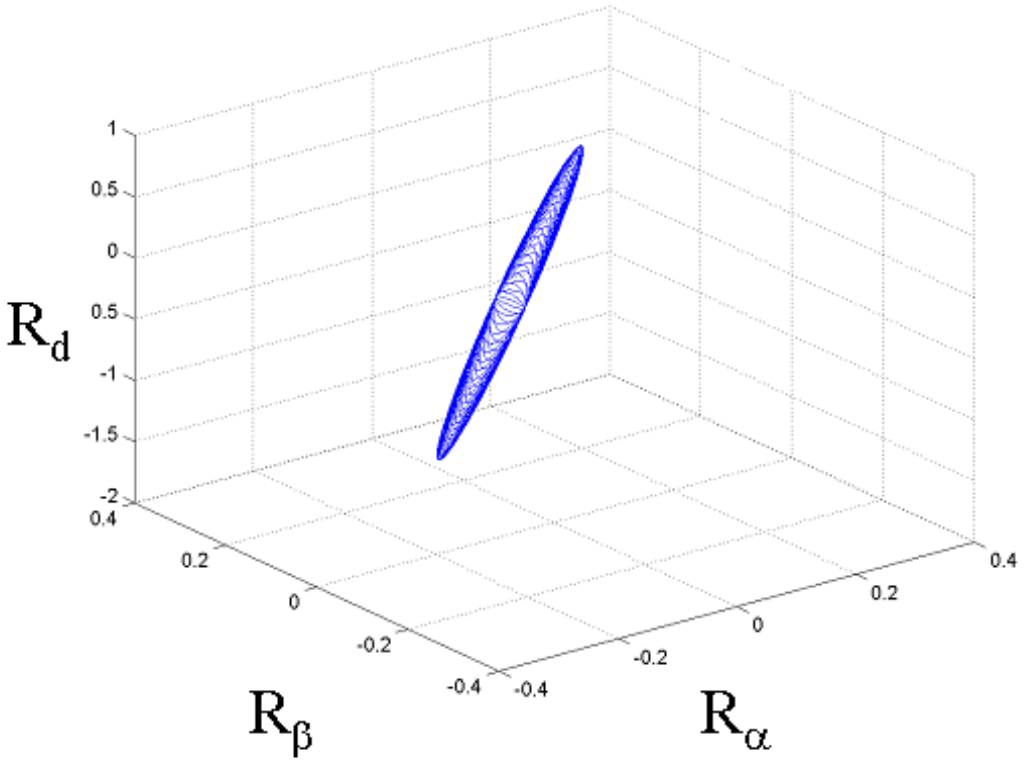

Figure 2.5: Misfit ellipsoid generated from likelihood function based on equation (2.1) for the acquisition geometry used in the Blackfoot synthetic and a noise variance of 1. Note that the ellipsoid is quite elongated especially along the density axis. The ellipsoid represents the equiprobable solution surface for a particular noise level.

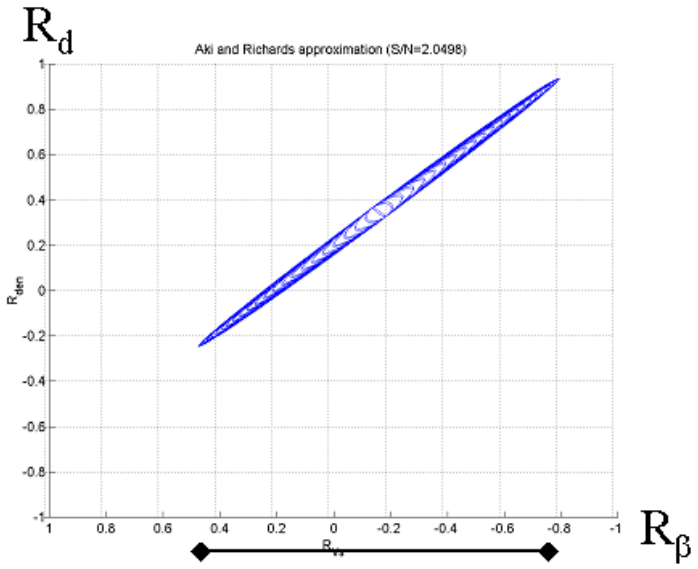

Figure 2.6: One face of the cube shown in Fig. 2.5. Note the large amount of uncertainty in both the S-wave velocity and density reflectivity.

a bivariate Gaussian distribution.

Both the Gidlow and Shuey reformulations of equation (2.1) are often solved using only the first two terms. This is equivalent to setting the third term to zero. For Shuey's equation this sets the P-wave velocity reflectivity to zero and for the Gidlow equation this sets the density reflectivity to zero. Figure 2.10 shows the application of these different constraints in three parameter solution space. In each case the optimal solution is the intersection of the plane defined by the constraint with the misfit ellipsoid. Each constraint results in a slightly different optimal solution (Figure 2.11). The solution is as good as the prior knowledge leading to the constraint. Determining which solution is most realistic involves answering the question, "Is it more realistic that the velocity reflectivity is zero, that the density reflectivity is zero or that the Gardner equation holds?" By introducing the constraint, the uncertainty has been reduced but at the expense of introducing error arising from the correctness of the *a priori* information.

It has been suggested that to avoid the bias introduced by using the two-term Shuey equation, all three terms should be used and then the third term thrown away.

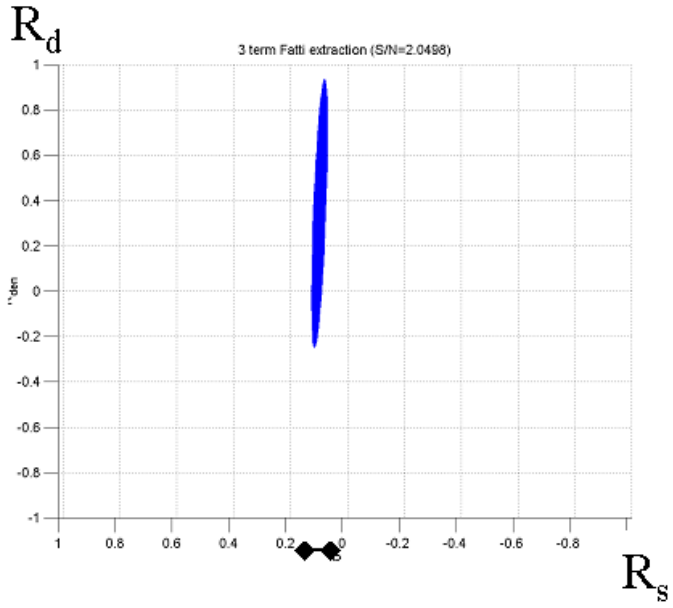

Figure 2.7: One face of the cube shown in Fig. 2.5 after variables transformed to impedance. Note that because of the transform, the uncertainty of variable along the x-axis is much less than in Fig. 2.5.

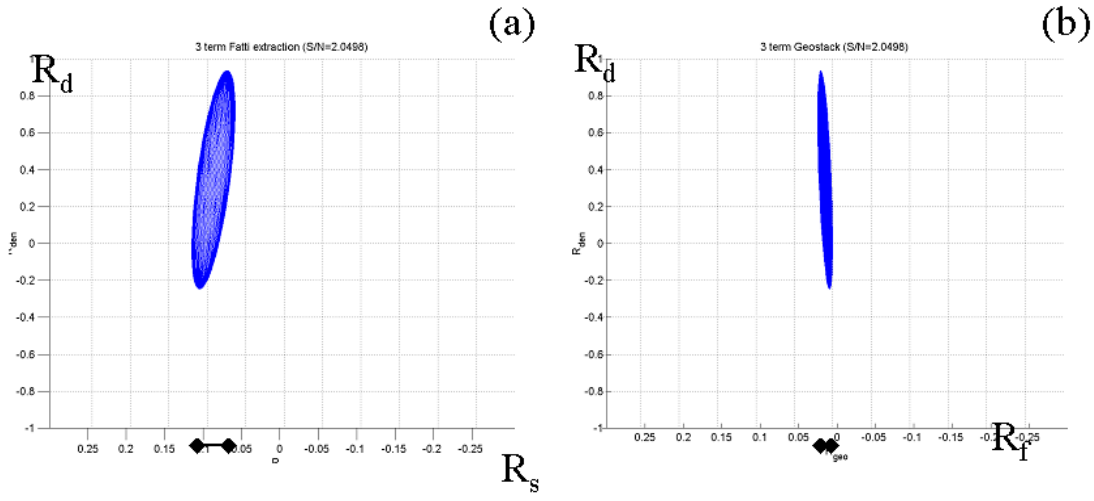

Figure 2.8: Comparison of misfit ellipsoids for two alternative parameterizations. The S-impedance reflectivity (a) has greater uncertainty than the fluid stack parameterization (b).

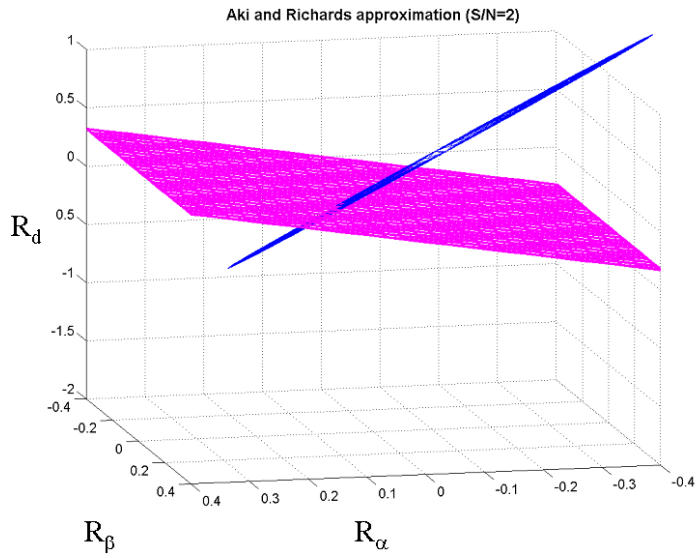

Figure 2.9: The blue surface is the misfit ellipsoid generated from likelihood function based on equation (2.1) for the acquisition geometry used in the Blackfoot synthetic. The red surface is the solution space defined by the *a prior* constraints based on the Gardner equation. The probability of this is a delta function. The optimal solution occurs where the combined probability is maximum.

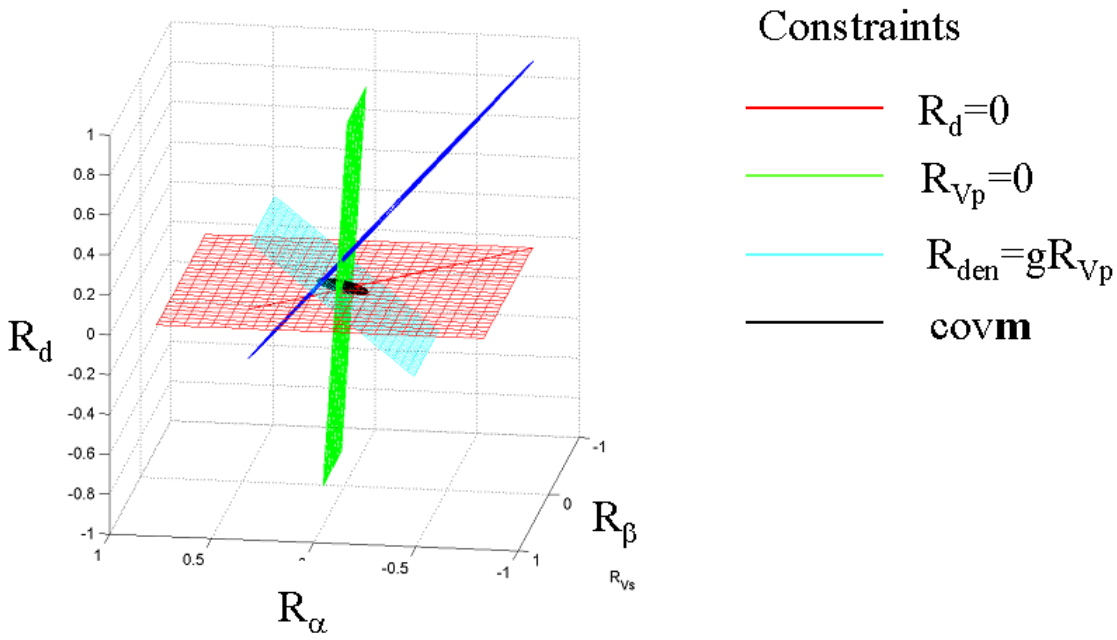

Figure 2.10: The blue surface is the misfit ellipsoid generated from likelihood function based on equation (2.1) for the acquisition geometry used in the Blackfoot synthetic seismogram. Various constraints are shown. The optimal solution in each case occurs where the combined probability is maximum. Note that each constraint has a different optimal solution.

This does remove the bias but adds a great deal of uncertainty to the parameter estimates. The unconstrained estimate of the gradient term is much more sensitive to noise.

2.4.3 Two-term parameter uncertainty

In practice, most AVO inversion being done in the industry today uses some form of two-term constrained inversion. This being the case, the rest of this chapter considers how to quantify the influence of noise for this particular inversion problem. This analysis, however, ignores the theoretical error associated with the application of inappropriate constraints and underestimates the total error. The estimation and uncertainty analysis for the three-term problem is dealt with in the next chapter. In the previous section, it was shown that the probability distribution for the two term AVO inversion is described by a bivariate Gaussian distribution parameterized using covariance matrix (2.37). In studying the covariance matrix (equation 2.37) it

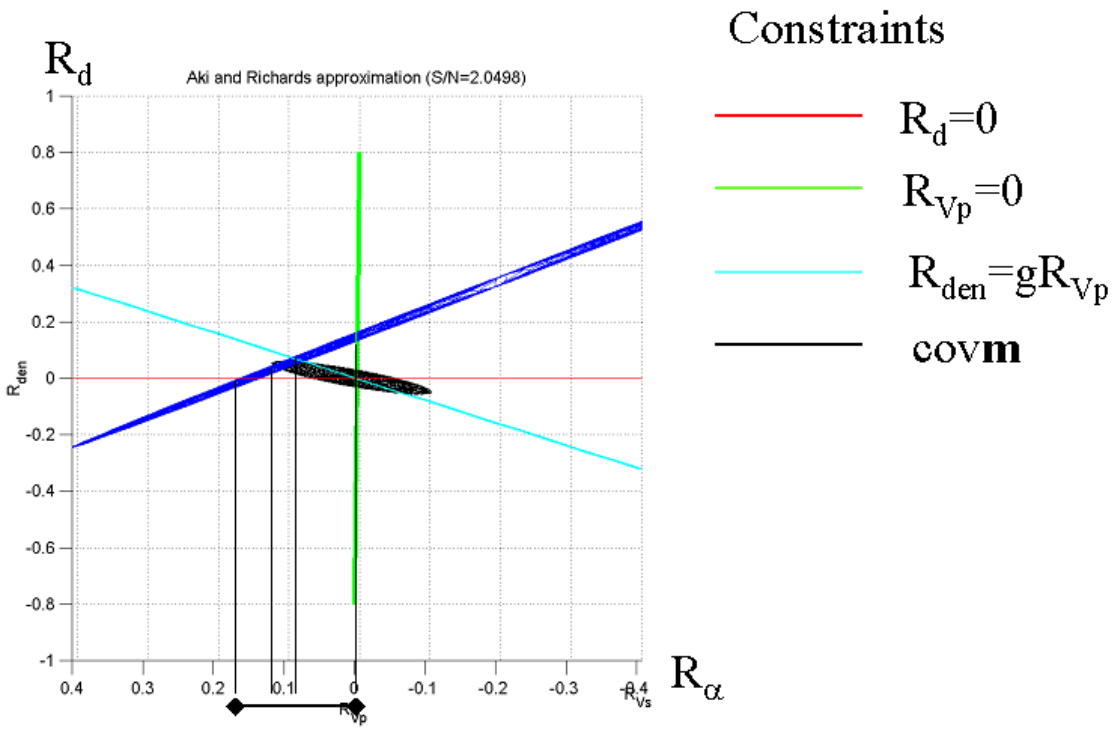

Figure 2.11: The blue surface is the misfit ellipsoid generated from likelihood function based on equation (2.1) for the acquisition geometry used in the Blackfoot synthetic. Various constraints are shown. The optimal solution in each case occurs where the combined probability is maximum. Note that each constraint has a different optimal solution.

is instructive to separate the influence of the prestack noise $\hat{\sigma}_N^2$ and the inverse of the normal equations, $[\mathbf{G}^T\mathbf{G}]^{-1}$ since both contain unique information. The prestack noise acts as a scalar: the more noise the greater the uncertainty. The normal equations are solely determined by the acquisition geometry and the background velocity fields. For future reference, this thesis refers to the square root of the diagonal terms of this matrix as stability sections. Displaying these diagonal elements gives an indication of the stability of each variable due to the acquisition geometry and background velocity which is separate from the estimate of the noise.

For the case of the two-term Gidlow equation, the inverse of the normal equations, $[\mathbf{G}^T\mathbf{G}]^{-1}$ is described by the analytic function (equation D.13)

$$[\mathbf{G}^T\mathbf{G}]^{-1} = \frac{1}{D}\begin{bmatrix} \sum_{m=1}^{M}\sin^4\bar{\theta}_m & \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M}\tan^2\bar{\theta}_m \\ \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M}\tan^2\bar{\theta}_m & \frac{1}{64\bar{\gamma}^4}\sum_{m=1}^{M}\sec^4\bar{\theta}_m \end{bmatrix}, \tag{2.43}$$

where D is defined by equation (2.41). The inverse is solely a function of the ratio γ, α, and the acquisition geometry. The influence of different acquisition parameters such as maximum and minimum offset can be studied using equation (2.43). Figure 2.12 shows the results of such a study. In the first set of examples (Figure 2.12a, b, c) the fold is held constant, while the angles used in the AVO inversion are varied. In Figure 2.12a the maximum angle included in the AVO inversion is varied while the minimum angle is held constant at zero. The uncertainty in the S-wave impedance reflectivity decreases in a nonlinear fashion as the maximum angle used increases. For small maximum angles, increasing the angle range by a small amount greatly decreases the uncertainty. For large maximum angles, the effect of increasing the angle range by a small amount is less noticeable. Because of this, for typical prestack signal-to-noise ratios, it is desirable to have a minimum maximum angle of 20 ° and preferably 30 °. The standard deviation of the P-wave impedance, unlike the standard deviation of the S-wave impedance, seems to be insensitive to the maximum angle used.

The influence of changing the minimum angle while holding the maximum angle constant is shown in Figure 2.12b. Unlike Figure 2.12a both the uncertainty for the P-wave and S-wave impedance reflectivity increase as the minimum angle is increased. Intuitively, this is easy to understand if one thinks of the AVO problem in terms of the two-term Shuey relationship of fitting a line to data points. If there are many points close to the intercept, the intercept can be estimated with great certainty.

If the data points are far from the intercept, greater uncertainty comes into the estimate. This issue is important for marine data where it is difficult to record the near angles especially in the shallow section. This is also an issue with 3D land acquisition where there is a predominance of far offsets relative to near offsets. This can be partially compensated by including greater angles as shown in Figure 2.12c. Figure 2.12c shows the case of where the total range of angles used in the AVO inversion is held constant but the minimum angle used is varied. Note that the variance for the S-wave impedance reflectivity remains relatively constant. The P-wave impedance variance increases but is still an order of magnitude smaller than the S-wave impedance variance. This suggests that what is really important for the stability of the problem is the usable angle range or aperture.

Lastly the effect of fold is considered. In this analysis the angle range is held constant. In Figure 2.12d it is observed that if the fold increases, the reliability of the estimates increases. This is due to the fold effectively increasing the poststack signal-to-noise ratio and thereby increasing the reliability of the estimates. This is an interesting observation since this suggests that one way of increasing the S-wave impedance reflectivity reliability is to increase the fold acquired in data acquisition.

The ratio $\bar{\gamma}$ also influences the stability of the inversion as Figure 2.13 shows. As $\bar{\gamma}$ is decreased from 3 to 1.7 the condition number decreases by a factor of 8. Typically shallow unconsolidated clastics have higher α/β ratios while deeper consolidated sands or carbonates have lower values. Since $\bar{\gamma}$ often changes as a function of depth, this implies the that the uncertainty of the parameter estimates also will change as a function of depth.

Change of variables

In order to analyze two reflectivity sections simultaneously, interpreters often look at linear combinations of reflectivity attributes. The fluid stack (Smith & Gidlow 1986, Gidlow et al., 1992) is an example of this. When crossplotted, the P-wave and S-wave impedance reflectivity should cluster along a line whose slope is defined by the mudrock relationship (Castagna et al., 1985) and the background α/β ratio. Significant departure from the linear trend may indicate the presence of gas. Recall that the fluid stack is calculated by using the transform matrix (equation 2.22). The

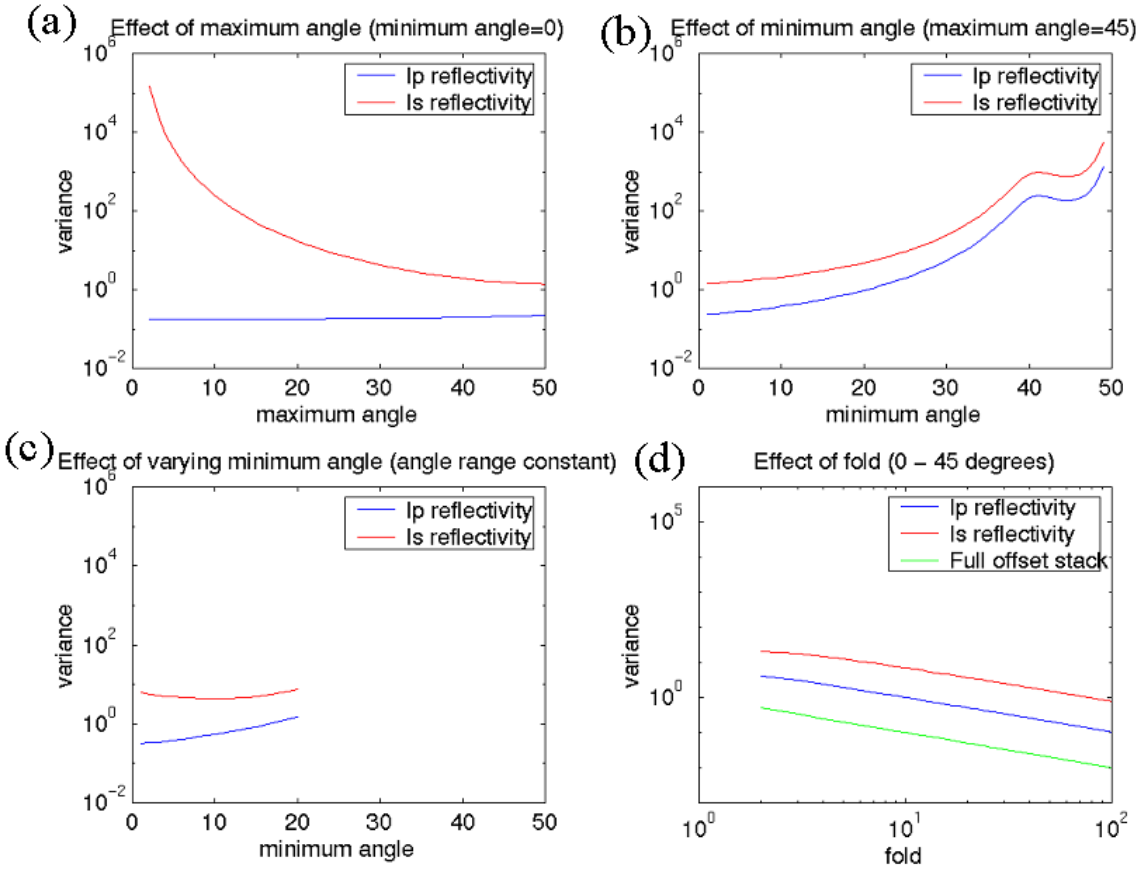

Figure 2.12: The estimated variance based on the two-term Fatti covariance matrix. Figure (a) shows the effect of varying the maximum angle while holding the fold constant. Figure (b) shows the effect of varying the minimum angle while holding the fold and maximum angle constant. Figure (c) shows the affect of holding the range of angles used to do the inversion constant. Figure (d) shows the effect of varying fold while holding the minimum and maximum angles constant.

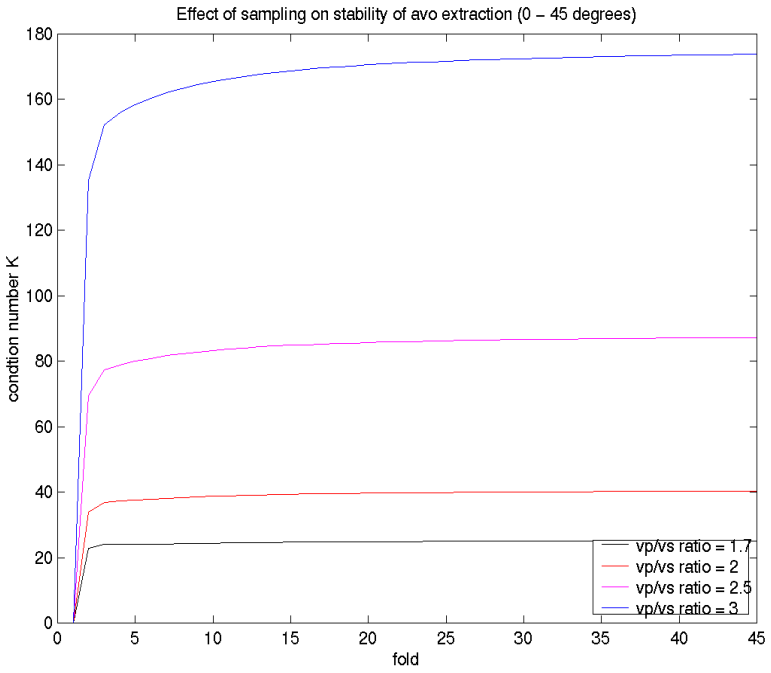

Figure 2.13: The condition number of the AVO inversion problem as both the α/β ratio and fold are varied, while the minimum and maximum angles are held constant. Note that as α/β increases, the condition number increases.

variance of the transformed covariance matrix is

$$\hat{\mathbf{C}}_{\mathbf{x}'} = \mathbf{T}\hat{\mathbf{C}}_{\mathbf{x}}\mathbf{T}^T, \tag{2.44}$$

where $\mathbf{T}$ is the transform matrix. Thus, the covariance matrix for the transformed variables

$$\hat{\mathbf{C}}_{\mathbf{x}'} = \begin{bmatrix} \hat{\sigma}^2_{Rp} & \hat{\sigma}_{RpRg} \\ \hat{\sigma}_{RpRg} & \hat{\sigma}^2_{Rg} \end{bmatrix}, \tag{2.45}$$

is

$$\hat{\mathbf{C}}_{\mathbf{x}'} = \frac{\hat{\sigma}^2_N}{D} \begin{bmatrix} 1 & 0 \\ 1 & -m\bar{\gamma} \end{bmatrix} \begin{bmatrix} \sum_{m=1}^{M} \sin^4\bar{\theta}_m & \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M} \tan^2\bar{\theta}_m \\ \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M} \tan^2\bar{\theta}_m & \frac{1}{64\bar{\gamma}^4}\sum_{m=1}^{M} \frac{1}{\cos^4\bar{\theta}_m} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -m\bar{\gamma} \end{bmatrix}^T, \tag{2.46}$$

and where upon the transformation

$$\hat{\sigma}^2_{Rp} = \frac{\hat{\sigma}^2_N}{D}\sum_{m=1}^{M} \sin^4\bar{\theta}_m, \tag{2.47}$$

$$\hat{\sigma}_{RpRg} = \frac{\hat{\sigma}^2_N}{D}\sum_{m=1}^{M} \sin^4\bar{\theta}_m - \frac{m}{8\bar{\gamma}}\sum_{m=1}^{M} \tan^2\bar{\theta}_m, \tag{2.48}$$

and in particular the variance for the fluid stack is

$$\hat{\sigma}^2_{Rg} = \frac{\hat{\sigma}^2_N}{D}\left(\sum_{m=1}^{M} \sin^4\bar{\theta}_m - \frac{m}{4\bar{\gamma}}\sum_{m=1}^{M} \tan^2\bar{\theta}_m + \frac{m^2}{128\bar{\gamma}^2}\sum_{m=1}^{M} \frac{1}{\cos^4\bar{\theta}_m}\right). \tag{2.49}$$

Figure 2.14 shows the variance before and after the transformation where the original variables P-wave and S-wave impedance reflectivity variance are transformed to the P-wave impedance reflectivity and fluid stack variance. It is interesting to note that the fluid stack has a lower variance than the S-wave impedance reflectivity. The fluid stack is an interesting and useful reflectivity attribute because of its low uncertainty in the presence of noise and its ability to highlight fluid changes.

2.5 Feasibility and Uncertainty analysis

The reliability of the AVO attributes can be calculated using the above methodology prior to doing an AVO analysis, as a feasibility study, to investigate whether

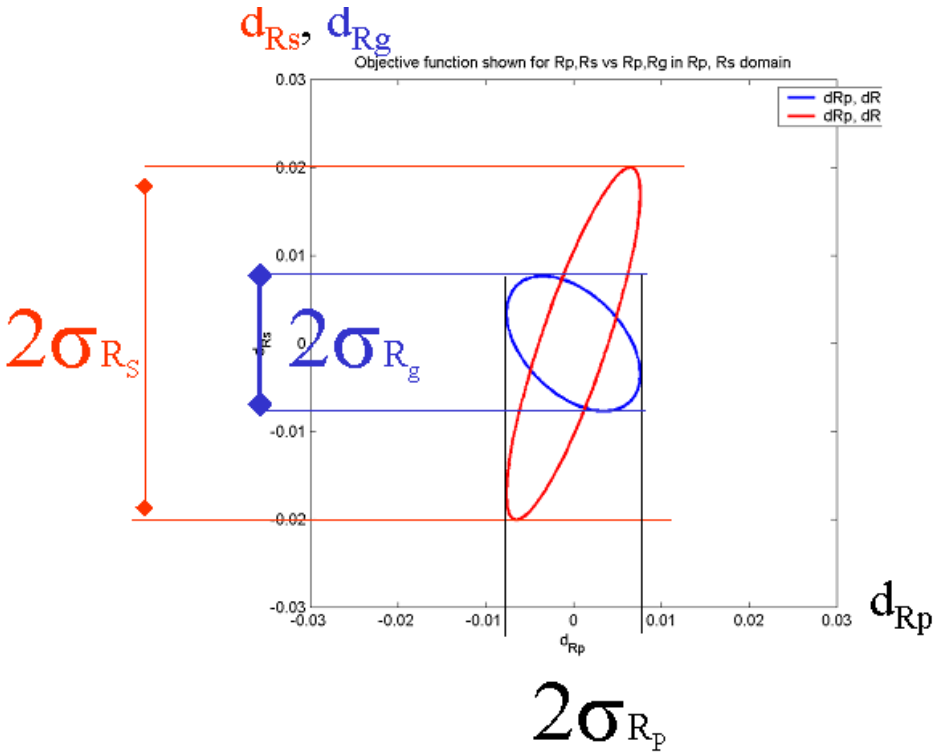

Figure 2.14: The misfit function of the P- and S-wave impedance reflectivity (red) compared to the P-impedance, fluid stack parameterization (blue). Note the uncertainty of the fluid stack is less than that of the S-wave impedance reflectivity. Note the axis correspond to $dR = R_{ideal} - R_{estimate}$.

the proposed seismic survey has the required fold and aperture to get usable results. Some sort of estimate of the prestack noise level is needed to do this analysis. This analysis could also be performed as part of the AVO inversion to provide reliability quality controls to help with the interpretation of the AVO reflectivity attributes. In this case, the prestack noise level can be estimated by calculating the misfit of the model to the data. To validate that this methodology gives reasonable estimates of uncertainty a modeling study was performed testing the key variables identified above.

2.5.1 Modeling study

A suite of well logs from the Blackfoot field in western Canada including P-wave velocity, S-wave velocity, and density information, was used to generate a series of primary only convolutional AVO models. These models were created using the Zoeppritz equations and ray tracing with a 10/14-70/80 Hz zero phase wavelet. The models were generated with a number of different folds, aperture, and prestack signal-to-noise ratios so the sensitivity of each of these variables could be tested independently. The models were created in such a manner that one variable was varied while the others were held constant. Figure 2.15 is displayed with a signal-to-noise ratio of 2:1 and offset out to 2000 m. The zone of interest is the upper and lower Glauconite gas sand which corresponds to the low velocity and density interval from 0.98 to 1.01 seconds.

To allow for this testing, the low frequency P-wave velocity trend of the well was modified so that over the window from 0.88 to 1.1 seconds, the angle of incidence is roughly constant for each offset. This allows the fold to be varied independently of angle of incidence in the testing. The α/β ratio and density curves were retained from the original log.

Two sets of gathers were created. The first set of gathers was designed so that 25 separate AVO inversions could be performed over different angle ranges while keeping the fold constant at 64 and the minimum angle constant at zero. The maximum angle was allowed to vary from 2° to 50° in 2 degree increments. Noise was added to these gathers with the following signal-to-noise ratios: 64/1, 16/1, 4/1, 1/1, 1/4 and 1/16. To understand how the extraction behaved at the end points, one set of gathers had no noise and another set was all noise. This resulted in 200 permutations (25 different angle records times 8 noise levels).

The second set of gathers was designed such that the angle range in the AVO

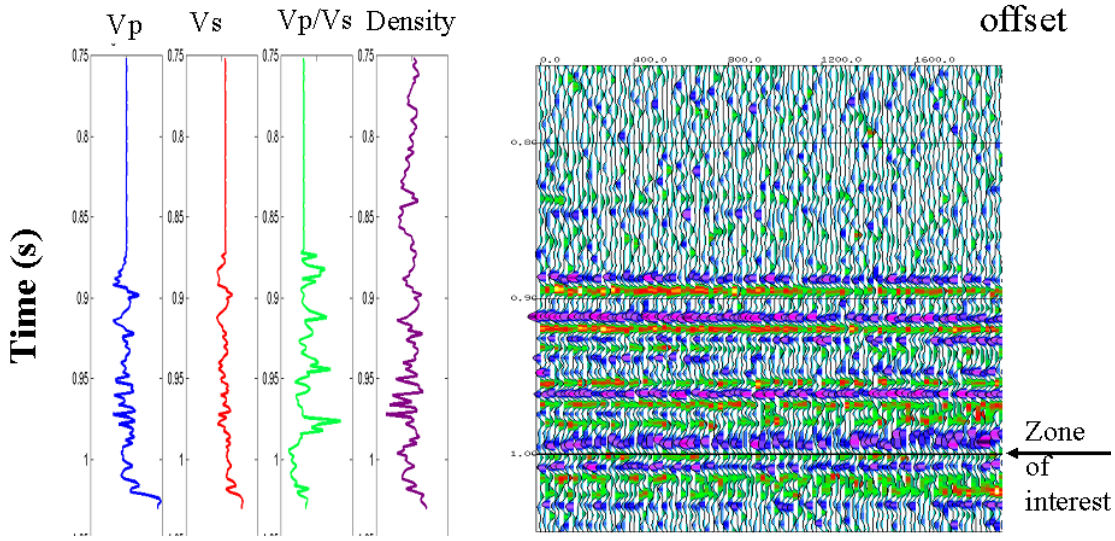

Figure 2.15: Wireline logs displayed in time for the Blackfoot well log. The synthetic gather was generated using reflectivity generated from the Zoeppritz equation. The reflectivity is shown without moveout, filtered and with noise added to give a S/N=2.

extraction could be held constant (0° to 45°) while the fold was allowed to vary over the following range: 4, 8, 16, 32, 64, 128, 256 and 512. The same noise levels were combined with this data resulting in 64 permutations (8 different fold records times 8 noise levels).

AVO extractions were then performed on both these sets of records to calculate estimates of P-wave and S-wave impedance reflectivity using the two- term Gidlow equation. A simple low frequency velocity model was used to ray trace the data. The α/β ratio was smoothly varying. These were constructed as one would normally do in performing an AVO inversion on real data.

2.5.2 Modeling Results

The top panel of Figure 2.16 shows the estimates of S-wave impedance reflectivity $\hat{R}_s$ when the angular range is varied for different signal-to-noise ratios. These are compared to the ideal zero offset S-wave impedance reflectivity displayed in terms of P-wave travel time (Figure 2.16b). The fractional error (Figure 2.16c) is calculated

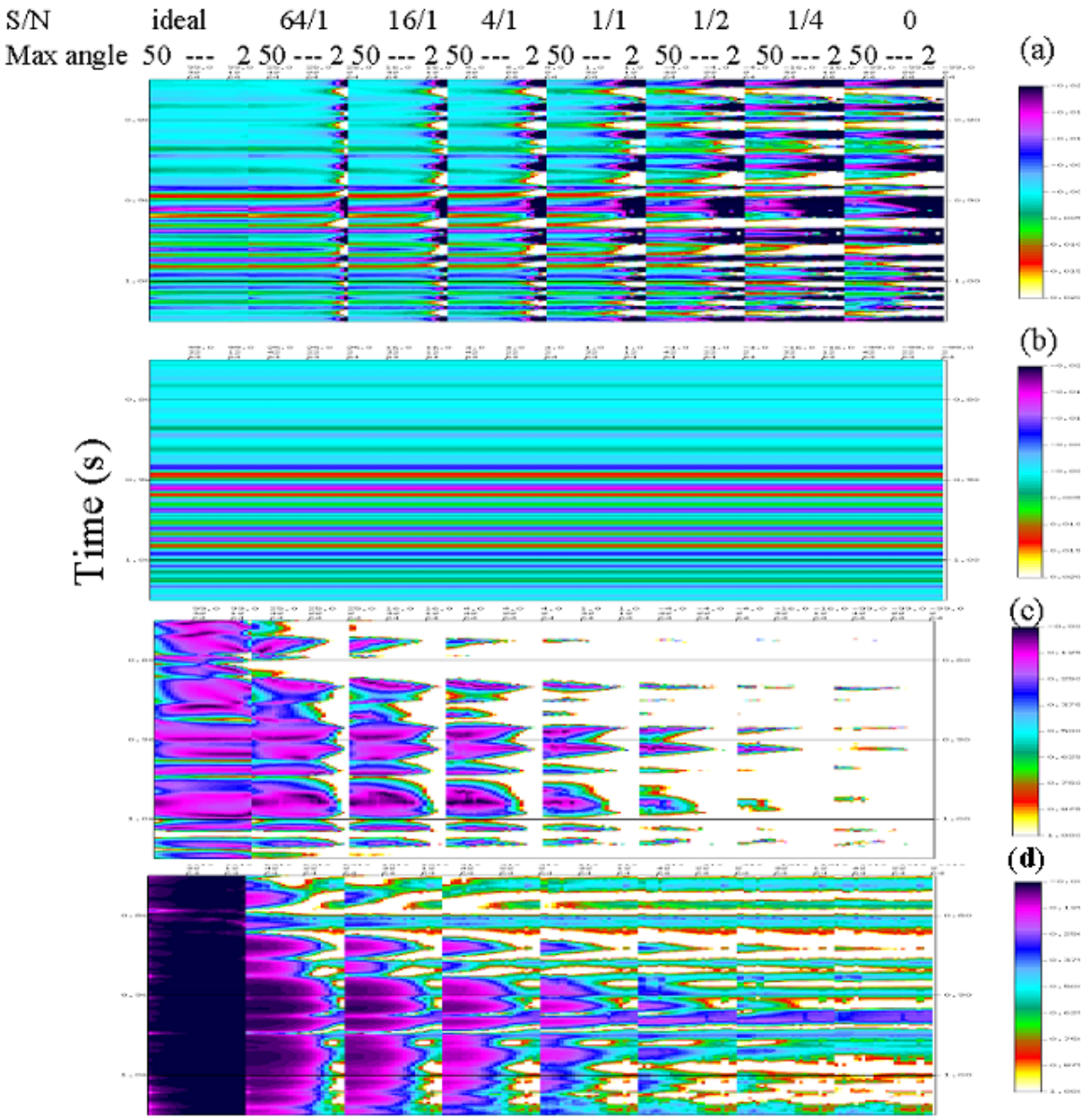

Figure 2.16: Subplot (a) is the S-impedance reflectivity generated from AVO modeling experiment described in text. Fig. (b) shows the ideal reflectivity while (c) shows fractional uncertainty calculated from (a) and b). Fig. (d) shows the predicted fractional uncertainty calculated from the covariance matrix.

using

$$FE\left(R_s\right) = \frac{I_{amp}\left(\hat{R}_s - R_s\right)}{I_{amp}\left(R_s\right) + \varepsilon}, \tag{2.50}$$

where ε is a small stability factor and the functional $I_{amp}\left(\right)$ is the operation of taking the instantaneous amplitude (Taner and Sheriff, 1977). Figure 2.16d shows the estimate of the fractional uncertainty of the S-wave impedance reflectivity predicted by the covariance matrix. This agrees well with the ideal fractional error (Figure 2.16c) calculated using equation (2.50).

Figure 2.17 shows the uncertainty of the P-wave and S-wave impedance reflectivity attributes calculated based on the actual error over a 0.100 second window, centered on 0.950 seconds. These measurements of the actual error are consistent with the predictions made by the covariance matrix in the preceding section (Figure 2.12). The P-wave impedance reflectivity is insensitive to angle range while the S-wave impedance reflectivity is strongly sensitive to aperture in the same manner as suggested by Figure 2.12a. It is interesting to note that for high angles and fold the fractional error actually increases. This is not a misfit error, but is due to theoretical error introduced by the *a priori* information used to constrain the problem.

The data misfit is used to estimate the noise variance, which is plotted versus signal-to-noise in Figure 2.17. This was done to see if the misfit could be used to predict the noise level and hence the uncertainty in the AVO extraction on noisy data. The variance calculated using equation (2.38) shows a linear relationship versus signal-to-noise (Figures 2.17a and 2.17b). There is little scatter when the aperture is allowed to vary (Figure 2.17a). However, there is significantly more scatter when the fold is varied (Figure 2.17b). This means that the misfit can be used to estimate uncertainty, but the error in this estimate increases as the fold decreases.

2.5.3 Colony data example

Figure 2.20a shows the P-wave impedance reflectivity extraction over a Colony gas field. This Colony gas sand is a Class III AVO anomaly. It is the purple amplitude around 0.5 seconds. The feature at CDP 4070 was drilled on the AVO analysis of this line and resulted in a successful well. The background α/β ratio is quite high, around 2.5. The estimated uncertainty in the AVO inversion for this line is expected to be small since a successful gas well was drilled based on the AVO analysis. The fluid stack (Figure 2.19) shows an anomaly at the Colony level over quite a wide area.

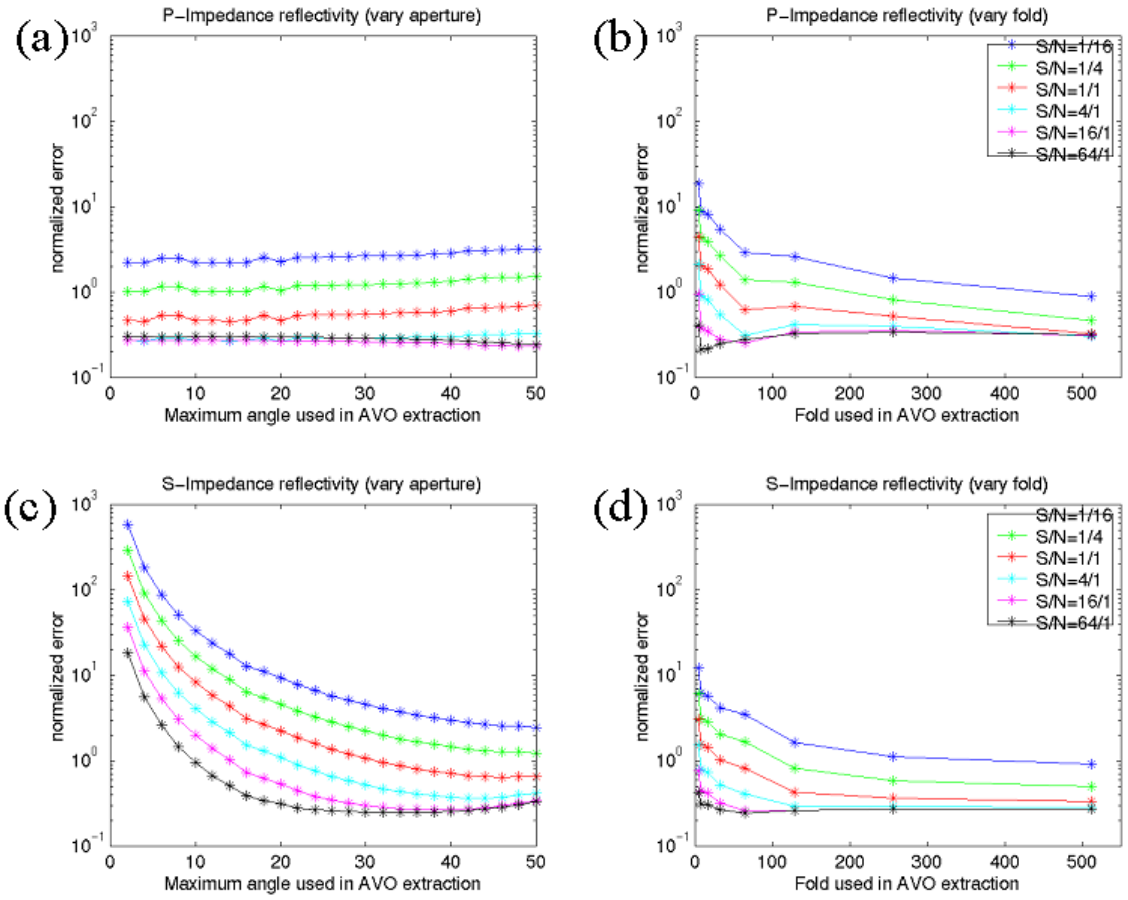

Figure 2.17: Normalized error calculated as described in text. Fig. a) shows P-impedance reflectivity error when the maximum angle is varied while b) shows the error while the fold is varied. Fig. c) shows S-impedance reflectivity error when the maximum angle is varied while d) shows the error while the fold is varied.

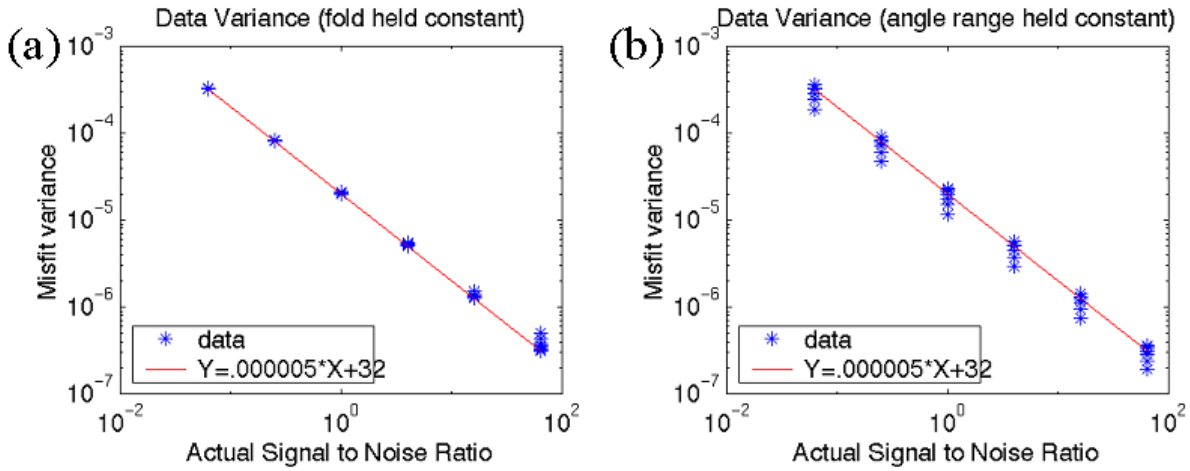

Figure 2.18: Predicted noise versus actual noise variance. Fig (a) shows calculation when fold held constant while (b) shows the calculation when angle range is held constant.

Note that at the beginning and end of the line the fluid stack appears unstable.

The S-wave impedance stability section (Figure 2.21b) shows that this is indeed the case. This is due to the lack of far offsets in the data acquisition at the ends of the line. As well, local instabilities in the shallow section are evident. These instabilities are due to missed source locations in the data acquisition, and disappear with depth as the fold increases and the α/β ratio decreases. In comparison, the P-wave impedance stability section (Figure 2.20b) is quite uniform and insensitive to the variations in acquisition geometry.

The fractional uncertainty (equation 2.50) calculated for both the P-wave and S-wave impedance sections is shown in Figure 2.20c and 2.21c respectively. The fractional uncertainty for the P-wave impedance reflectivity is less than the S-wave impedance reflectivity as expected. Both have uncertainty less than 0.2 in the zone of interest. This is significantly less than the expected change in zero-offset P-wave impedance reflectivity due to the introduction of gas.

2.5.4 Heavy oil seismic data example

To illustrate the use of reliability predictions on real data, a fluid stack (Figure 2.22) is shown for a 2D seismic line acquired over a heavy oil field undergoing steam injection. Modeling studies (Downton and Lines, 2001c) suggest the fluid stack should

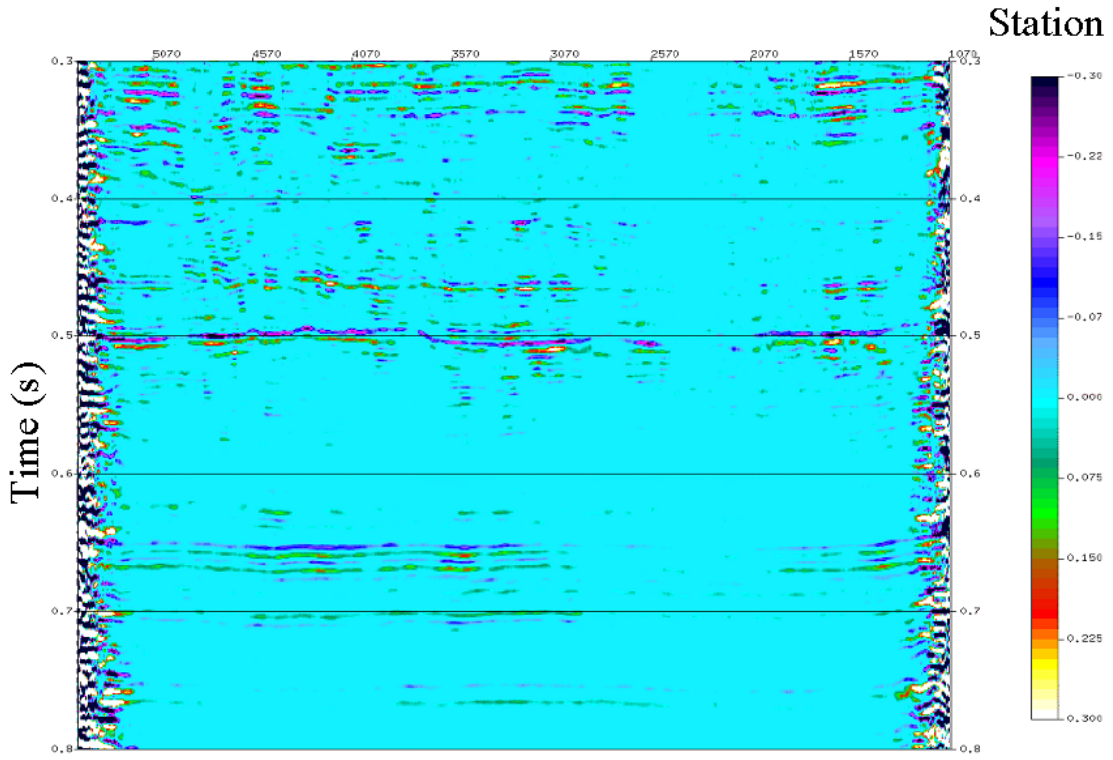

Figure 2.19: Fluid stack for colony example. Note Colony anomalies at 0.5 seconds.

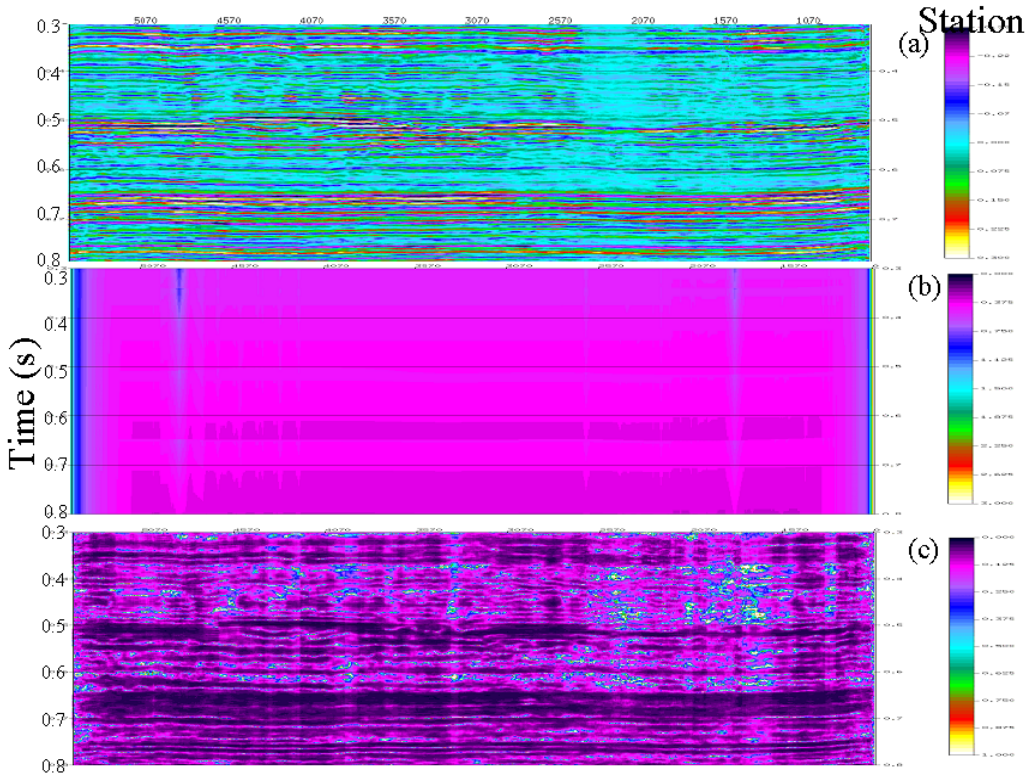

Figure 2.20: Colony P-impedance reflectivity (a), stability section for P-wave impedance reflectivity (b), and fractional uncertainty (c).

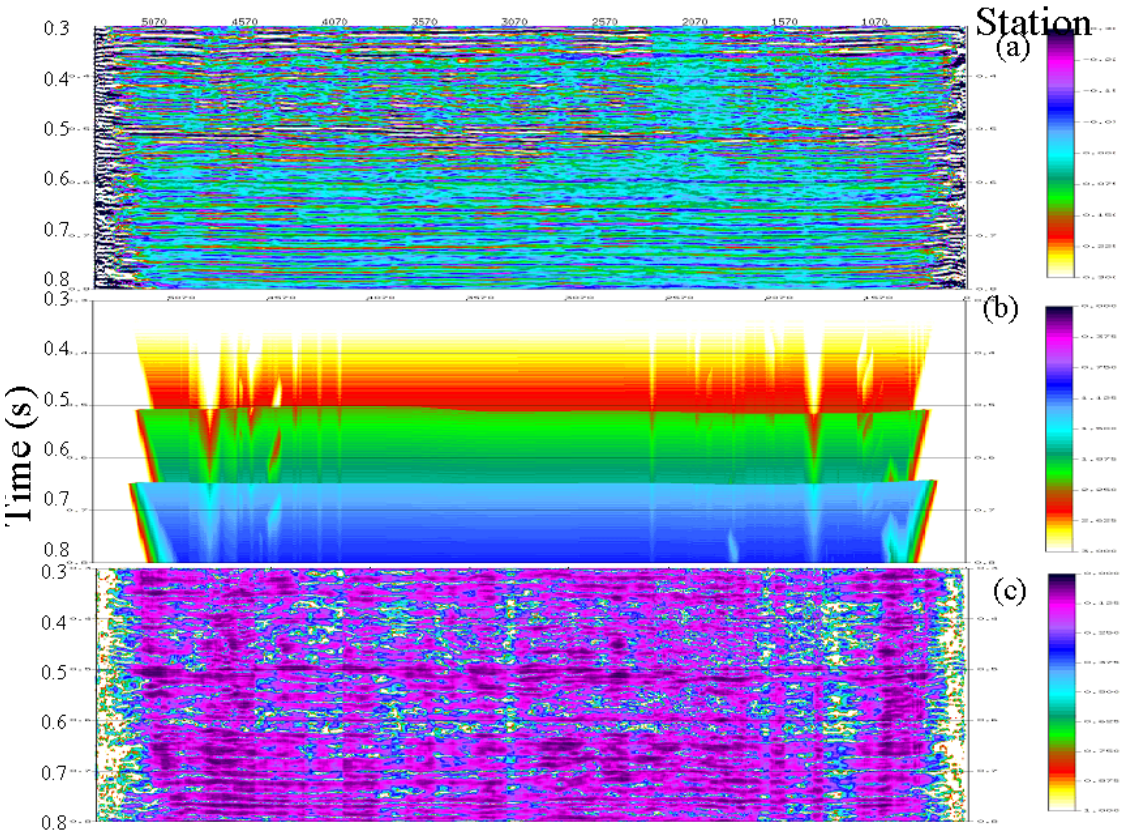

Figure 2.21: Colony S-wave impedance reflectivity (a), stability section for S-impedance reflectivity (b), and fractional uncertainty (c).

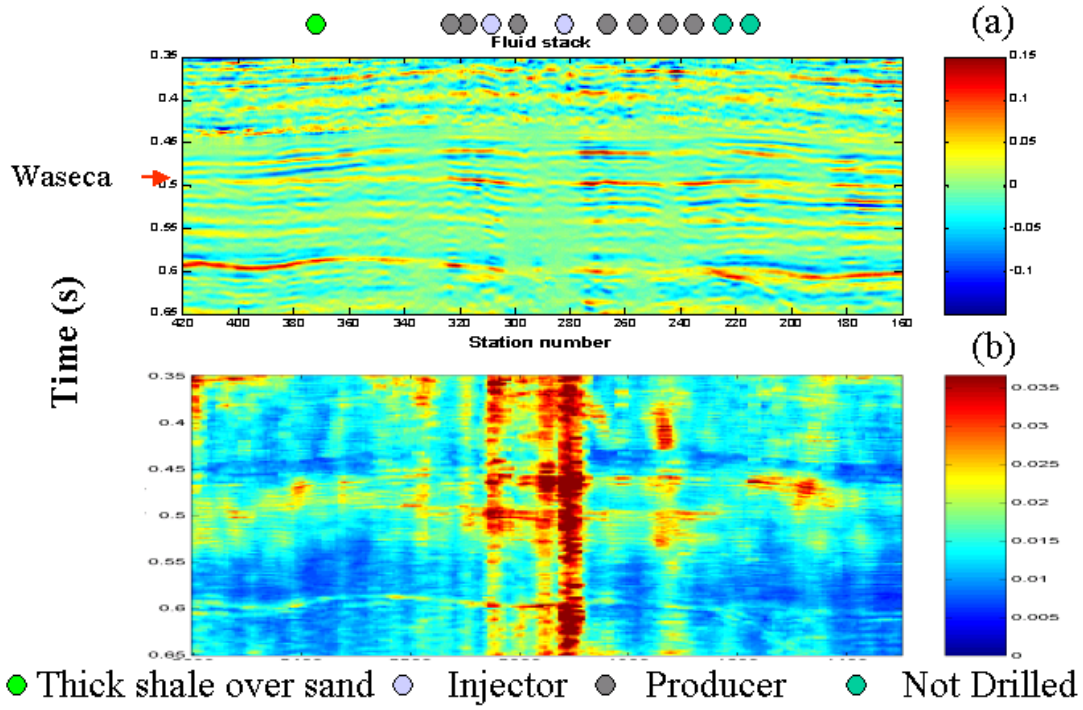

Figure 2.22: Fluid stack for A.O.S.T.R.A. example (a), and fractional uncertainty display (b).

show an anomaly where the reservoir is at elevated temperatures. The line was acquired while production was ongoing. As a result there is considerable noise from the pumps causing reliability issues with the reflectivity attributes at several places along the line. The fluid stack shows anomalies at the base of the Waseca sand which corresponds well with most of the producing wells and injectors. However, the injector at station 280 does not show an anomaly. It is interesting to note that the standard deviation of the fluid stack at station 280 is several orders of magnitude greater than the fluid stack estimate. This suggests that the fluid stack attribute is unreliable at this point and the predicted fluid stack value should be ignored. With the incorporation of the reliability information, the fluid stack is consistent with the geologic control. Greater confidence can be placed in the AVO interpretation by knowing where the AVO estimates are reliable and just as importantly where they are less reliable.

2.6 Discussion

It is important to note that this chapter has not considered the impact of coherent noise, theoretical error, and systematic data errors. Theoretical error and non-Gaussian error are discussed in future chapters. Systematic data errors arise due to incorrect preconditioning or processing of the seismic prior to the AVO inversion (section 1.3.3). Having ignored theoretical and systematic errors, the uncertainty estimates generated in this chapter are unduly optimistic. This being the case, the reliability estimates are still valuable in that they identify problem areas in the survey due to acquisition geometry and noise.

It is difficult to identify systematic data errors. One approach is to crossplot the reflectivity estimates, to at least gain some indication whether the results are geologically plausible. For example, the P-wave and S-wave impedance reflectivity may be crossplotted. The data should cluster around a line whose slope is defined by the mudrock relationship (equation 1.38)

$$R_\alpha = m\bar{\gamma}R_\beta. \tag{2.51}$$

This is discussed further in Chapter 4. The gathers may also be compared to synthetic models or walkaway VSPs (Ross and Beale, 1994). To avoid systematic data errors it is best to process the data in as careful a fashion as possible following the key steps outlined in Chapter 1 (section 1.3.3).

2.7 Conclusions

This chapter reviewed various three-term linearized approximations of the Zoeppritz equations. Provided $\bar{\gamma}$ is known *a priori* these parameterizations are mathematically equivalent. There is no point doing statistical regression or neural network analysis on more than three of these attributes since there are only three independent pieces of information available. The interpreter may wish to display the reflectivity attributes which best highlight the geologic objective using the transform matrix (2.28). However, these reflectivity attributes have different uncertainty associated with them. Given two parameterizations that discriminate the geology equally well, it is desirable to use the parameterization with the lesser amount of uncertainty. For example, impedance reflectivity and velocity reflectivity often convey similar geo-

logic information, but the impedance reflectivity has dramatically lower uncertainty. Thus, it is preferable to use impedance reflectivity over velocity reflectivity. Better yet, it is preferable to use the fluid stack based on impedance reflectivity as the second attribute over S-wave impedance reflectivity.

It was shown that the three-term AVO inversion problem is ill-conditioned necessitating the use of constraints. In general the problem can be made more stable by including a greater range of angles in the inversion. However, the theoretical error also increases as maximum angle used in the inversion increases, so that in practice there is an upper limit that may be used. In addition, the background α/β ratio influences the parameter uncertainty with a larger α/β ratio implying greater uncertainty.

To address the ill-conditioned nature of the linearized AVO inversion problem, constraints are used to stabilize the problem. In this chapter, "hard constraints" were used. In this case, they reduce the number of parameters to be solved for from three to two, but at the expense of potentially introducing theoretical error into the problem. The correctness of the different approximations of the Zoeppritz equations published in the literature can be viewed in terms of how geologically plausible the applied constraints are and what bias they introduce into the problem. The Gardner constraint used in the Smith and Gidlow formulation is clearly more physical than the two-term Shuey approximation where the P-wave velocity reflectivity is set to zero. However, both lead to errors when their respective equalities are not met. It would be preferable to use a probabilistic expression rather than an equality to describe an empirical relationship. In the next chapter, probabilistic constraints are developed.

The uncertainty of the AVO inversion problem was investigated with the aid of Bayes' Theorem. In the case of independent uniform Gaussian noise equation (2.37) may be used to estimate the uncertainty of each parameter. In this case, the reliability is only a function of the prestack noise, the line geometry and the background α/β ratio. Based on this, a series of quality control displays can be generated. The standard deviation of the noise may be estimated from the misfit function. This may be accurately estimated as long as the fold is sufficient. The other key displays are the stability sections generated from the diagonal of matrix (2.43). This shows the stability of the inversion due to changes in the acquisition geometry and α/β ratio. From this the standard deviation of each parameter may be calculated and with the inclusion of the parameter, the fractional uncertainty calculated. The uncertainty

may be transformed to any parameterization using equation (2.44). The ability of this methodology to predict the uncertainty was tested on synthetic data with good results. In addition, when applied to real data the predicted uncertainty was in accordance with known sources of error. The calculated uncertainty was greatest where the angle range was smallest and/or noise the greatest.

The covariance matrix can be created before or as part of the AVO inversion. If it is done before as part of a feasibility study, it can help determine if the seismic data set is a suitable input to AVO analysis. If it is run as part of the AVO inversion, the quality controls can be used to help appropriately weight the value of the AVO information relative to other forms of information in the overall interpretation.

Chapter 3

Constrained three-term AVO inversion and uncertainty analysis

3.1 Introduction

In the preceding chapter, it was shown that the three-term linearized AVO equation is typically too ill-conditioned to be reliably solved for. The problem was made stable by solving for two parameters rather than three. This was done by introducing hard constraints, either explicitly, through the use of rock physical relationships or implicitly by truncating the third term in the various linearized approximations of the Zoeppritz equations. In doing so, there is a trade-off between increased reliability and a loss of information in the form of one of the parameters. In contrast, this chapter incorporates probabilistic constraints and uses Bayes' theorem to find the optimal solution. The degree to which the constraints influence the solution is dependent on the signal-to-noise ratio, the fold and the range of angles available to the inversion. When the signal-to-noise ratio and geometry are sufficient, the data dominates the solution rather than the constraints and reliable predictions may be made for all three terms. When the conditions are inadequate, the constraints dominate the solution providing less information, yet the solution behaves in a geologically believable fashion. Uncertainty displays provided by the method help the interpreter understand the reliability of the estimates. The approach has some similarities to the approach of Lortzer and Berkhout (1993) in that both implement statistical constraints using Bayes' theorem. Lortzer and Berkhout use empirical relationships and implement them as data constraints. In this approach the constraints are treated as *a priori* information and much more detail is provided in the construction of them.

The advantage of three-term AVO inversion is that density reflectivity may also

be estimated. Density reflectivity provides the potential to quantify fluid saturation within the reservoir. The P-wave velocity reacts to gas saturation in a nonlinear fashion (Figure 3.1) with gas saturations greater than 10% causing most of the response (Domenico, 1976). Thus, the P-wave impedance stack reacts to trace amounts of gas. Two-term AVO analysis (Gidlow et al., 1992; Shuey, 1985) can predict the presence of gas within a reservoir, but not whether gas present, is in quantities large enough to be commercial. In contrast, density reacts to gas in a linear fashion. By comparing the velocity with the density response, more accurate estimates of gas saturation may be made. Recently, there has been a series of papers demonstrating three-term AVO inversion. Kelly and Skidmore (2001) developed a nonlinear approximation to the Zoeppritz equations, but do not discuss how the solution is stabilized. In a subsequent paper, Van Koughnet et al. (2003) publish a series of examples from the Gulf Coast showing that density reflectivity can be practically solved for and used in an exploration environment. Downton and Chaveste (2004) show an example demonstrating the use of three-term AVO inversion to predict fluid properties in a reservoir where the sand quality varies complicating the density response and analysis. The preceding paper and this chapter are based on the work and are an extension of Downton and Lines (2001a).

This chapter describes how probabilistic constraints may be developed either from local well control or empirical rock physical relationships and how these constraints may be used to stabilize the solution. The algorithm incorporates constraints calibrated and suitable for local geologic conditions over some target interval. Bayes' theorem is used to develop the theory, algorithm and uncertainty analysis resulting in a constrained three-term nonlinear AVO inversion algorithm with reliability estimates. The inversion solves for P-wave and S-wave impedance, and density reflectivity but can be transformed to a variety of other attributes popular in the literature in a simple manner. Both synthetic and real seismic data examples are shown. The synthetic example demonstrates that even when the P-wave velocity and density are uncorrelated, the density reflectivity may be accurately predicted under situations typical for real seismic. The real seismic data example shows the ability of the density reflectivity to differentiate a non-commercial reservoir from a commercial reservoir.

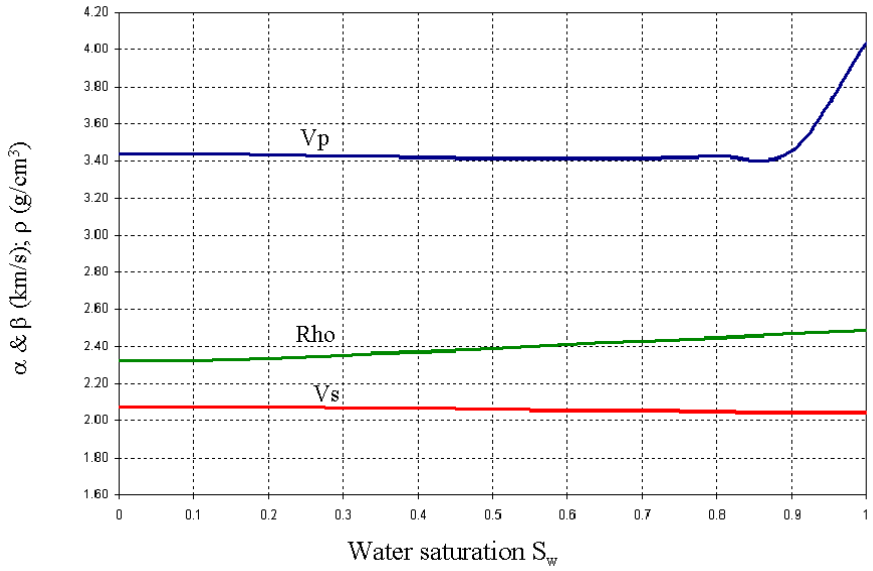

Figure 3.1: Response of density, P-wave, and S-wave velocity to changing water and gas saturation assuming a two phase fluid. The Gassmann equation is used to model the velocities.

3.2 Theory

3.2.1 Amplitude variation with offset model

Following the approach outlined in Chapter 2, the elastic parameters may be estimated, using a linearized approximation of the Zoeppritz equations, such as the three-term Gidlow et al. (1992) (equation 2.19)

$$\begin{bmatrix} R(\bar{\theta}_1) \\ \vdots \\ R(\bar{\theta}_M) \end{bmatrix} = \begin{bmatrix} \sec^2\bar{\theta}_1 & -8\bar{\gamma}^2\sin_1^2\bar{\theta} & 4\bar{\gamma}^2\sin^2\bar{\theta} - \tan^2\bar{\theta} \\ \vdots & \vdots & \vdots \\ \sec^2\bar{\theta}_M & -8\bar{\gamma}^2\sin^2\bar{\theta}_M & 4\bar{\gamma}^2\sin^2\bar{\theta}_M - \tan^2\bar{\theta}_M \end{bmatrix} \begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix}, \quad (3.1)$$

where $R\left(\bar{\theta}\right)$ is the angle dependent reflectivity and the parameters R_P, R_S, R_d respectively are the P-wave and S-wave impedance, and density reflectivity as defined in Chapter 2. The parameter $\bar{\gamma}$ is the average ratio of the S- to P-wave velocity while the variable $\bar{\theta}$ is the average angle of incidence across the interface. If $\bar{\theta}$ and $\bar{\gamma}$ are considered to be known *a priori*, equation (3.1) may be solved using linear inverse techniques. Equation (3.1) assumes that there are M offsets and that ray tracing is done to map the offsets to the average angle of incidence.

This model assumes the earth is composed of a series of flat, homogeneous, isotropic layers. Transmission losses, converted waves, and multiples are not incorporated in this model and so must be corrected for through prior processing. In theory, gain corrections such as spherical divergence, absorption, directivity, and array corrections can be incorporated into this model, but are not considered for brevity and simplicity, so must be previously applied. The inversion is performed on NMO corrected or prestack migrated gathers as the case may be. In this chapter the band-limited nature of the seismic data are ignored along with NMO stretch and offset dependent tuning which will be discussed in later chapters. By ignoring the band-limited nature of the seismic data, each time sample can be considered an interface which may be independently inverted for. This is the typical assumption made in AVO inversion. The inversion is performed on each time sample for each CMP gather.

I choose to use a linearized approximation of the Zoeppritz equations, rather than a higher order approximation, such as done by Kelly and Skidmore (2001), since the latter includes squared reflectivity terms. If these terms are included, then first order intrabed multiples should be as well, since they are of the same order leading to a much more complex model such as considered by Sen and Stoffa (1995).

Equation (3.1) may be written more succinctly in matrix notation as

$$\mathbf{G}\mathbf{x} = \mathbf{d}, \tag{3.2}$$

where $\mathbf{G}$ is the linear operator, $\mathbf{x}$ the unknown parameter vector and $\mathbf{d}$ the input data vector (offset dependent reflectivity). The parameter vector is composed of the density, P-wave and S-wave impedance reflectivity attributes. These are called reflectivity attributes since they share the same mathematical form as reflectivity but not the same physical significance. For conciseness and brevity, in the rest of this thesis I shall refer to $\mathbf{x}$ as reflectivity.

3.2.2 Bayes' theorem

Bayes' theorem provides a theoretical framework to make probabilistic estimates of the unknown reflectivity attributes $\mathbf{x}$ from uncertain data and *a priori* information. The resulting probabilistic parameter estimates are called the Posterior Probability Distribution Function (*PPDF*). The *PPDF* written symbolically as $P(\mathbf{x}|\mathbf{d},I)$ indicates the probability of the parameter vector $\mathbf{x}$ given the data vector $\mathbf{d}$ (offset

dependent reflectivity) and information I. Bayes' theorem

$$P\left(\mathbf{x}|\mathbf{d},I\right) = \frac{P\left(\mathbf{d}|\mathbf{x},I\right) P\left(\mathbf{x}|I\right)}{P\left(\mathbf{d}|I\right)}, \tag{3.3}$$

calculates the *PPDF* from the likelihood function $P\left(\mathbf{d}|\mathbf{x},I\right)$ and an *a priori* probability function $P\left(\mathbf{x}|I\right)$. The denominator $P\left(\mathbf{d}|I\right)$ is a normalization function which may be ignored if only the shape of the *PPDF* is of interest so

$$P\left(\mathbf{x}|\mathbf{d},I\right) \propto P\left(\mathbf{d}|\mathbf{x},I\right) P\left(\mathbf{x}|I\right). \tag{3.4}$$

In this dissertation rather than characterizing each reflectivity attribute by its complete *PPDF,* parametric measures are used to characterize the *PPDF.* The maximum of the *PPDF* or mode is used as the estimate of the reflectivity attribute while some measure of dispersion or width of the *PPDF* is used characterize the uncertainty of the parameter estimate.

3.2.3 Likelihood function

If the noise associated with the m^{th} experimental measure is assumed to be Gaussian with variance σ_m then the probability of the m^{th} offset dependent reflectivity d_m, given the parameter vector $\mathbf{x}$ is

$$P(d_m|\mathbf{x},\sigma_m, I) = \frac{1}{\sigma_m\sqrt{2\pi}} \exp\left(-\frac{\left(F_m - d_m\right)^2}{2\sigma_m^2}\right), \tag{3.5}$$

where F_m describes the functional relationship between the data and parameters $\mathbf{x}$ for offset m as described by equation (3.1) $F_m = \sum_{p=1}^{3} G_{mp}x_p$. If it is further assumed that the noise in the data are independent then

$$P(\mathbf{d}|\mathbf{x},\sigma_m, I) = (2\pi)^{-\frac{M}{2}} \left(\prod_{m=1}^{M} \sigma_m^{-1}\right) \exp\left(-\sum_{m=1}^{M} \frac{\left(F_m - d_m\right)^2}{2\sigma_m^2}\right). \tag{3.6}$$

This follows from the repeated application of the product rule (Sivia 1996, equation 1.2) $P(d_m, d_n|\mathbf{x},\sigma_m, I) = P(d_m|d_n,\mathbf{x}, I) \times P(d_n|,\mathbf{x}, I)$ and from the definition of independent data $P(d_m|d_n,\mathbf{x}, I) = P(d_m|\mathbf{x}, I)$ so that $P(d_m, d_n|\mathbf{x},\sigma_m, I) = P(d_m|\mathbf{x}, I) \times$

$P(d_n|, \mathbf{x}, I)$. Assuming uniform noise $\sigma_m = \sigma_N$ then

$$P(\mathbf{d}|\mathbf{x}, \sigma_N, I) \propto (2\pi)^{-\frac{M}{2}} \sigma_N^{-M} \exp\left(-\sum_{m=1}^{M} \frac{\left(\sum_{n=1}^{3} G_{mn} x_n - d_m\right)^2}{2\sigma_N^2} \right). \tag{3.7}$$

In the case of uniform priors, Bayesian inversion is equivalent to maximum likelihood inversion. For AVO inversion, because of the small number of parameters solved for, it is possible to visualize the *PPDF*. If the parameter vector $\mathbf{x}$ had only one element, the *PPDF* is a Gaussian function. If the parameter vector $\mathbf{x}$ has two elements, the *PPDF* is a bivariate Gaussian function and an equiprobable solution is an ellipse. For the case of AVO inversion, where there are three parameters, the *PPDF* is a multivariate Gaussian function where the equiprobable solution is an ellipsoid (Figure 2.5). Typically the ellipsoid is quite elongated along the density reflectivity axis. The solutions are nonphysical when the reflectivity is greater than $+1$ or less than -1.

3.2.4 *A priori* constraints

One way to reduce the uncertainty and to exclude nonphysical solutions is to impose constraints on the solution. For example, nonphysical solutions (e.g. $|r_p, r_s, r_d| > 1$) can be excluded from the solution space by assigning a uniform probability distribution where physical solutions are equiprobable and nonphysical solutions given zero probability. However, it is not necessarily desirable to assign uniform probabilities over the range of physically valid reflectivity. The stratigraphic sequence is a result of cyclic geologic processes that result in reflectivity probability functions, which may be reasonably described by common probability functions. It is desirable to select a probability function which contains greater information content.

The Gaussian probability function is highly desirable from this point of view. It is a good compromise between increased information content and computational simplicity. Compared to the uniform distribution, it contains additional information about the mean and dispersion of the reflectivity. The mean describes the overall background trend, which may be used to reconstruct the impedance trend, while the dispersion describes the range of the expected reflectivity values. This information is obtainable from nearby well control. Todoeschuck et al. (1990) argue, based on well control, that over a short enough window, the statistics of a well log are Gaussian. However, Todoeschuck et al. (1990) point out that this is not true over the length of

the entire well log where large velocity and density boundaries are crossed repeatedly. This tends to increase the kurtosis or the degree of peakedness of the distribution of the sequence. For our purposes over a small enough window, typical of a target oriented inversion, the Gaussian probability function is applicable. In later chapters, more computationally complex *a priori* probability distributions are explored.

In actuality, there are three reflectivity series, thus a multi-variate Gaussian probability function

$$P(\mathbf{x}|I) = \frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}-\langle\mathbf{x}\rangle\right)^T \mathbf{C}_\mathbf{x}^{-1}\left(\mathbf{x}-\langle\mathbf{x}\rangle\right)\right]}{(2\pi)^{\frac{3}{2}}\sqrt{\det|\mathbf{C}_\mathbf{x}|^3}}, \tag{3.8}$$

must be used to describe the probability distribution where $\langle\mathbf{x}\rangle$ is the mean reflectivity and $\mathbf{C}_\mathbf{x}$ is the parameter covariance matrix. In this case, the band-limited reflectivity is estimated so the mean $\langle\mathbf{x}\rangle = \mathbf{0}$, thus the multi-variate Gaussian probability function simplifies to

$$P(\mathbf{x}|I) = \frac{\exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{C}_\mathbf{x}^{-1}\mathbf{x}\right]}{(2\pi)^{\frac{3}{2}}\left(\det|\mathbf{C}_\mathbf{x}|\right)^{\frac{3}{2}}}, \tag{3.9}$$

which is parameterized by the parameter covariance matrix

$$\mathbf{C}_\mathbf{x} = \begin{bmatrix} \sigma^2_{R_p} & \sigma_{R_pR_s} & \sigma_{R_pR_d} \\ \sigma_{R_pR_s} & \sigma^2_{R_s} & \sigma_{R_sR_d} \\ \sigma_{R_pR_d} & \sigma_{R_sR_d} & \sigma^2_{R_d} \end{bmatrix}. \tag{3.10}$$

The diagonal elements of the parameter covariance matrix are the variances $\sigma^2_{R_p}, \sigma^2_{R_s}, \sigma^2_{R_d}$ of the P-wave and S-wave impedance and density reflectivity respectively. The off-diagonal elements or covariances describe how correlated the P-wave and S-wave impedance and density reflectivity are. For a particular covariance, the subscripts indicate for which two variables the covariance has been calculated. An alternative parameterization is in terms of the correlation coefficient. For example, the correlation coefficient $r_{R_pR_s}$ measures the how closely the P-wave and S-wave impedance reflectivities are correlated. The range of the correlation coefficient is $0 \leq |r| \leq 1$. If the magnitude of the correlation coefficient $\left|r_{R_pR_s}\right|$ approaches 0 then the S-wave impedance reflectivity does little to help predict the P-wave impedance reflectivity. Conversely, if the magnitude of the correlation coefficient $\left|r_{R_pR_s}\right|$ approaches 1 then the two reflectivity series are nearly identical. The relationship between the covari-

ance and the correlation coefficient is

$$r_{R_pR_s} = \frac{\sigma_{R_pR_s}}{\sigma_{R_p}\sigma_{R_s}}. \tag{3.11}$$

For future reference, the inverse covariance matrix or weighting matrix is defined as

$$\mathbf{W_x} = \mathbf{C_x^{-1}}, \tag{3.12}$$

so equation (3.9) transforms to

$$P(\mathbf{x}|I) = \frac{\exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{W_x}\mathbf{x}\right]}{(2\pi)^{\frac{3}{2}}\sqrt{\det|\mathbf{W_x^{-1}}|^3}}. \tag{3.13}$$

Calculation of the covariance matrix

Statistical calculation The simplest way to generate the covariance matrix in equation (3.10) is to estimate the statistics from nearby well control. This is done by transforming the appropriate logs from depth to time and then generating the reflectivity. The covariance matrix is calculated from the density, P-wave and S-wave velocity reflectivity series over a time interval similar to the window which is to be inverted. Assuming this window has N time samples and that these reflectivity series are stationary, then the $3 \times N$ replicate matrix $\mathbf{X}$ is constructed from the N parameter vectors $\mathbf{x}_i = [R_{p_i}, R_{s_i}, R_{d_i}]^T$ where the subscript represents the time sample. For zero mean variables, the covariance matrix is estimated using

$$\hat{\mathbf{C}}_\mathbf{x} = \frac{\mathbf{X}\mathbf{X}^T}{N}, \tag{3.14}$$

The *a priori* information should be indicative of the regional statistics of the geology. Large reflectivity anomalies, such as those due to Class III AVO anomalies, bias these trends. To prevent outliers dominating the calculation, the reflectivity may be weighted to reduce the influence of outliers. For example, the weighting matrix made up of the diagonal elements

$$W_{nn} = \begin{cases} \frac{1}{\epsilon} & \left|\frac{x_n}{\sigma_p}\right| < \epsilon \\ \frac{1}{\left|\frac{x_n}{\sigma_p}\right|} & \left|\frac{x_n}{\sigma_p}\right| > \epsilon \end{cases}, \tag{3.15}$$

accomplishes this. The weighted covariance matrix is

$$\hat{\mathbf{C}}_{\mathbf{x}} = \frac{\mathbf{X}\mathbf{W}\mathbf{X}^T}{N}. \tag{3.16}$$

It is shown in Chapter 6 that these weights may be used to mimic distributions other than Gaussian or ℓ^2 norm.

In reality the calculation is more complex than equation (3.14) or (3.16) suggests since at least one of the reflectivity series is nonstationary. The following section shows that the S-wave velocity reflectivity is nonstationary due to the fact that the P-wave to S-wave velocity ratio changes as a function of depth. This implies that the covariance matrix should change as a function of γ and be calculated with this nonstationarity taken into account. Further complicating the analysis is the fact that the explorationist is often dealing with incomplete information. Often, one or more of the sonic, density or dipole logs are not available. The following section describes a methodology on how to construct a time variant covariance matrix when there is missing information.

Covariance matrix parameterized in terms of rock physical relationships

Two-term formulation Recall from Chapter 1 that the mudrock relationship (1.34) relates the P-wave velocity α to the S-wave velocity β for fluid filled clastics using the empirical mudrock relationship

$$\alpha \approx m\beta + b. \tag{3.17}$$

Further, Smith and Gidlow reformulated equation (3.17) in terms of reflectivity to derive the fluid stack. In water saturated sediments equation (3.17) written in terms of reflectivity is

$$R_\alpha \approx m\bar{\gamma}R_\beta. \tag{3.18}$$

Assuming the P-wave velocity reflectivity R_α is stationary and γ changes as a function of depth, equation (3.18) implies the S-wave velocity reflectivity R_β is nonstationary. It is quite typical that the value of γ changes as a function of depth. This is evident when the mudrock equation (3.17) is rewritten in terms of γ,

$$\gamma = 0.862 \; - \frac{1172}{\alpha}, \tag{3.19}$$

using the typical parameters $m = 1.16$ and $b = 1360\ m/s$ (Castagna et al., 1985). Soft uncompacted formations have a low P-wave velocity whose lower limit approaches that of water. Thus, for soft unconsolidated sediments the second term is almost equal to the first resulting in the ratio being close to zero resembling an acoustic material. For more compact sediments the P-wave velocity approaches that of the dominant mineral. For example, if the upper limit of the P-wave velocity is 6500 m/s this implies a γ ratio of around 0.68. Since the value of γ changes as function of depth, equation (3.18) implies that R_β is nonstationary, even if R_α is stationary. If R_α is stationary over the inversion window, then by performing a change of variables $R'_\beta = \bar{\gamma} R_\beta$ it is possible to write equation (3.18) in terms of stationary variables

$$R_\alpha = m R'_\beta. \tag{3.20}$$

where R'_β is the scaled S-wave velocity reflectivity. The slope may be calculated as a least squares problem assuming that R_α and R'_β are stationary and that the mudrock slope is constant over the inversion window. In this case, equation (3.20) written in matrix notation is

$$\mathbf{r}_\alpha = m \mathbf{r}'_\beta, \tag{3.21}$$

for which the least squares solution is

$$m = \frac{\mathbf{r}'^T_\beta \mathbf{r}_\alpha}{\mathbf{r}'^T_\beta \mathbf{r}'_\beta} = \frac{\sigma_{R_\alpha R'_\beta}}{\sigma^2_{R'_\beta}}, \tag{3.22}$$

where $\sigma_{R_\alpha R'_\beta}$ is the covariance between the P-wave and scaled S-wave impedance reflectivity and $\sigma^2_{R'_\beta}$ is the variance of the scaled S-wave impedance reflectivity. If m is not constant over the whole inversion window then the analysis may be broken into smaller windows for which m is constant, and then the constraints may be applied in a time dependent fashion.

It is more intuitive to write equation (3.22) in terms of the correlation coefficient $r_{R_\alpha R'_\beta}$ rather than the covariance $\sigma_{R_\alpha R'_\beta}$. Rewriting equation (3.22) written in terms of the correlation coefficient (equation 3.11) results in a relationship between the P- and S- wave velocity reflectivity variance

$$\sigma_{R'_\beta} = \frac{r_{R_\alpha R'_\beta}}{m} \sigma_{R_\alpha}. \tag{3.23}$$

After substituting equation (3.11) and (3.23) the covariance matrix for the transformed variables $\mathbf{C}_{\mathbf{x}'}$ written in terms of the mudrock slope m and the correlation coefficient $r_{R_\alpha R'_\beta}$ is

$$\mathbf{C}_{\mathbf{x}'} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R'_\beta} \\ \sigma_{R_\alpha R'_\beta} & \sigma^2_{R'_\beta} \end{bmatrix} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m} \\ \frac{r^2_{R_\alpha R'_\beta}}{m} & \frac{r^2_{R_\alpha R'_\beta}}{m^2} \end{bmatrix}. \tag{3.24}$$

As long as positive values for m and $r_{R_\alpha R'_\beta}$ are chosen the parameter covariance matrix $\mathbf{C}_{\mathbf{x}'}$ is positive semi-definite. This is important later, since the inverse of the parameter covariance matrix is required to find the solution.

Change of variables Note that the parameter covariance matrix in equation (3.24) is written in terms of scaled S-wave velocity reflectivity R'_β rather than the S-wave velocity reflectivity R_β. In the next section, the parameter covariance matrix must be written in terms of S-wave velocity reflectivity R_β, so a change of variables

$$\begin{bmatrix} R_\alpha \\ R'_\beta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \bar{\gamma} \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \end{bmatrix}, \tag{3.25}$$

must be performed. Equation (3.25) written more generally in matrix notation is

$$\mathbf{x}' = \mathbf{T}\mathbf{x}, \tag{3.26}$$

where $\mathbf{x}'$ is the transformed variable and $\mathbf{T}$ is the transform matrix. The parameter covariance matrix (equation 3.14) under the change of variables (equation 3.26) is

$$\mathbf{C}_{\mathbf{x}} = \mathbf{T}^{-1}\mathbf{C}_{\mathbf{x}'}\left(\mathbf{T}^{-1}\right)^T. \tag{3.27}$$

Thus, the parameter covariance matrix equation (3.24) written in terms of the S-wave velocity reflectivity rather than the scaled S-wave velocity reflectivity is

$$\mathbf{C}_{\mathbf{x}} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R_\beta} \\ \sigma_{R_\alpha R_\beta} & \sigma^2_{R_\beta} \end{bmatrix} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\bar{\gamma}} \end{bmatrix} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m} \\ \frac{r^2_{R_\alpha R'_\beta}}{m} & \frac{r^2_{R_\alpha R'_\beta}}{m^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\bar{\gamma}} \end{bmatrix}^T, \tag{3.28}$$

or

$$\mathbf{C_x} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R_\beta} \\ \sigma_{R_\alpha R_\beta} & \sigma^2_{R_\beta} \end{bmatrix} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} \\ \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} & \frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} \end{bmatrix}. \tag{3.29}$$

From this the useful relationships

$$\sigma_{R_\alpha R_\beta} = \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} \sigma^2_{R_\alpha}, \tag{3.30}$$

and

$$\sigma_{R_\beta} = \frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}} \sigma_{R_\alpha}, \tag{3.31}$$

follow.

Three-term formulation The three-term parameter covariance matrix may be written in a similar fashion. Taking the derivative of the Gardner relationship (1.35) $\rho = e\alpha^g$ provides a relationship linking density to the P-wave velocity reflectivity,

$$R_d = gR_\alpha. \tag{3.32}$$

In a similar fashion to equation (3.22) the parameter g is found by least squares fitting the well log reflectivity

$$g = \frac{\sigma_{R_\alpha R_d}}{\sigma^2_{R_\alpha}}. \tag{3.33}$$

Substituting the definition of the correlation coefficient between the density and P-wave velocity reflectivity

$$r_{R_\alpha R_d} = \frac{\sigma_{R_\alpha R_d}}{\sigma_{R_\alpha}\sigma_{R_d}}, \tag{3.34}$$

into (3.33) results in a relation linking the density and P-wave velocity reflectivity

$$\sigma_{R_d} = \frac{g}{r_{R_\alpha R_d}} \sigma_{R_\alpha}. \tag{3.35}$$

Potter et al. (1998) observed a similar relationship to equation (3.32) between S-wave velocity and density. Thus, the reflectivity relationship

$$R_d = fR_\beta. \tag{3.36}$$

may be derived. In a similar fashion to equation (3.22) the parameter f is found by

least squares fitting the well log reflectivity

$$f = \frac{\sigma_{R_\beta R_d}}{\sigma^2_{R_\beta}}. \tag{3.37}$$

These parameters and their correlation coefficients can be calculated from the local well control. Upon substitution of equation (3.11), (3.23), (3.33), (3.37) and (3.35) into

$$\mathbf{C_x} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R_\beta} & \sigma_{R_\alpha R_d} \\ \sigma_{R_\alpha R_\beta} & \sigma^2_{R_\beta} & \sigma_{R_\beta R_d} \\ \sigma_{R_\alpha R_d} & \sigma_{R_\beta R_d} & \sigma^2_{R_d} \end{bmatrix}. \tag{3.38}$$

results in the multi-variate parameter covariance matrix

$$\begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R_\beta} & \sigma_{R_\alpha R_d} \\ \sigma_{R_\alpha R_\beta} & \sigma^2_{R_\beta} & \sigma_{R_\beta R_d} \\ \sigma_{R_\alpha R_d} & \sigma_{R_\beta R_d} & \sigma^2_{R_d} \end{bmatrix} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} & g \\ \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} & \left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 & f\left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 \\ g & f\left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 & \left(\frac{g}{r_{R_\alpha R_d}}\right)^2 \end{bmatrix}. \tag{3.39}$$

The advantages of writing the parameter covariance matrix in these terms are: (1) if there is missing information from one or more wells, geologically plausible parameters may be chosen to construct the covariance matrix, (2) the calculations of the statistics are all based on stationary reflectivity series, (3) the nonstationary nature of the covariance matrix is built into the analytic relationship through the γ dependence, and (4) if positive values of parameters are chosen this leads to a positive definite matrix which can be inverted as required by equation (3.12).

Parameter covariance matrix expressed in terms of impedance reflectivity

The likelihood model is expressed in terms of impedance reflectivity (equation 3.1) rather than velocity reflectivity as the parameter covariance matrix (equation 3.39) is. Using the transform matrix

$$\begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}, \tag{3.40}$$

the covariance matrix (3.39) becomes

$$\mathbf{C_x} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} & g \\ \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} & \left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 & f\left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 \\ g & f\left(\frac{r_{R_\alpha R'_\beta}}{m\bar{\gamma}}\right)^2 & \left(\frac{g}{r_{R_\alpha R_d}}\right)^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^T, \quad (3.41)$$

or

$$\mathbf{C_x} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 + 2g + \frac{g^2}{r^2_{R_\alpha R_d}} & \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} + g + f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}} & g + \frac{g^2}{r^2_{R_\alpha R_d}} \\ \frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} + g + f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}} & \frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + 2f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}} & f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}} \\ g + \frac{g^2}{r^2_{R_\alpha R_d}} & f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}} & \frac{g^2}{r^2_{R_\alpha R_d}} \end{bmatrix}. \quad (3.42)$$

Writing this with the matrix normalized with respect to the P-wave impedance reflectivity variance $\sigma^2_{R_p} = \sigma^2_{R_\alpha}\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)$ the covariance matrix (3.42) becomes

$$\mathbf{C_x} = \sigma^2_{R_p}\tilde{\mathbf{C}}_\mathbf{x}. \quad (3.43)$$

where $\tilde{\mathbf{C}}_\mathbf{x}$ is defined as the normalized parameter covariance matrix

$$\tilde{\mathbf{C}}_\mathbf{x} = \begin{bmatrix} 1 & \frac{\frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} + g + f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} & \frac{g + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} \\ \frac{\frac{r^2_{R_\alpha R'_\beta}}{m\bar{\gamma}} + g + f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} & \frac{\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + 2f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} & \frac{f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} \\ \frac{g + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} & \frac{f\frac{r^2_{R_\alpha R'_\beta}}{m^2\bar{\gamma}^2} + \frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} & \frac{\frac{g^2}{r^2_{R_\alpha R_d}}}{\left(2g + \frac{g^2}{r^2_{R_d R_\alpha}} + 1\right)} \end{bmatrix}. \quad (3.44)$$

Sample calculation

Figures 3.2 and 3.3 illustrate this approach based on logs from a well drilled from northeastern British Columbia, for which the drilling target was the Halfway sand. This well is used in a later section to create a synthetic model and to constrain the real

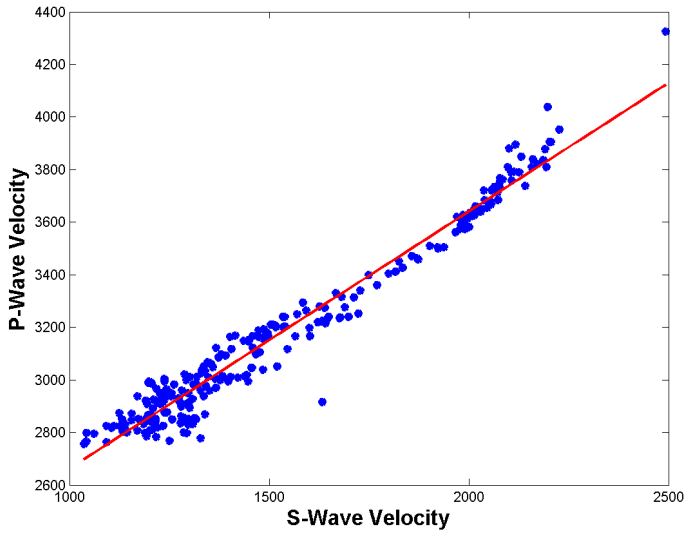

Figure 3.2: Cross-plot of the P- and S-wave velocities in m/s for the Halfway well over the clastic interval. The red line indicates the best fit in a least squares sense of $\alpha = 0.98\beta + 1683$.

seismic data example. Figure 3.2 shows a crossplot of the P- and S- wave velocities. Excluding the carbonates, there is a strong linear relationship relating the velocities, $\alpha = 0.98\beta + 1683$. Figure 3.3a shows the relationship between the P-wave and S-wave velocity reflectivity. The least squares fit found using equation (3.22) is $R_\alpha = 1.01 R'_\beta$ with a correlation coefficient of $r_{R_\alpha R_{\beta'}} = 0.845$. The close agreement between the two slope estimates supports the relationship between equations (3.17) and (3.20). Thus, for $\gamma = 0.5$, $m = 1.01$ and $r = 0.845$ the two-term covariance matrix (3.29) is

$$\mathbf{C_x} = \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{bmatrix} = \sigma_{R_\alpha}^2 \begin{bmatrix} 1 & 1.4139 \\ 1.4139 & 2.7998 \end{bmatrix}. \tag{3.45}$$

Note, in actuality the covariance matrix changes as a function of γ. Figure 3.3b shows the crossplots for the density reflectivity versus the P-wave velocity crossplot. The least-squares relationship for this is $R_d = 0.2559 R_\alpha$ with a correlation coefficient of $r_{R_\alpha R_d} = 0.8483$. This value of $g = 0.2559$ is close to that published by Gardner et al. (1974) of $g = 0.25$. Similarly, Figure 3.3c shows the crossplot between the density and S-wave velocity reflectivity with the least squares fit $R_d = 0.1791 R_\beta$. From these values the parameter covariance matrix (3.44) is calculated to be

$$\tilde{\mathbf{C}}_{\mathbf{x}} = \begin{bmatrix} 1.0 & 1.411\,4 & 0.216\,43 \\ 1.411\,4 & 2.429\,3 & 0.369\,63 \\ 0.216\,43 & 0.369\,63 & 5.677\,6 \times 10^{-2} \end{bmatrix}. \tag{3.46}$$

for $\gamma = 0.5$.

3.2.5 Nonlinear inversion

The likelihood function (equation 3.7) may be combined with the *a priori* probability function (equation 3.13) using Bayes' Theorem (equation 3.4). Figure 3.4 illustrates this. The estimate is taken to be the mode of the *PPDF*. On Figure 3.4, this would be close to the intersection of the two ellipsoids. Contrast this to Figure 2.9 which shows the two-term Smith and Gidlow solution. The Gardner constraint $R_d = gR_\alpha$ defines a plane. The optimal solution is where the misfit function (ellipsoid) intersects this plane. The constraint dramatically decreases the uncertainty but at the expense of introducing theoretical error. For example, if there is no noise in the solution, the estimate must still lie within the plane defined by $R_d = gR_\alpha$. In contrast, the Bayesian solution would be dominated by the likelihood function. It would be able to predict situations where the density behaves in a different fashion than the P-wave velocity reflectivity such as in a low saturation gas sand. Figures 2.10 and 2.11 compare the probabilistic constraint used in the Bayesian solution with explicit and implicit constraints used to generate two-term solutions. The black surface is the equi-probable surface generated by the multivariate Gaussian constraints. It is evident that the different constraints greatly influence the mode of the *PPDF* and therefore the estimate of the parameters with some choices clearly not being appropriate. For example, the two-term Shuey approximation requires that the P-wave velocity reflectivity be set to zero, which is at odds with a variable velocity earth model. This theoretical error introduced by the constraint is what leads to the bias in the estimate of the gradient term in the Shuey equation.

Mathematically, combining the likelihood function (equation 3.7) with the *a priori* probability function (equation 3.13) using Bayes' Theorem (equation 3.4) leads to

$$P(\mathbf{x}, \sigma_N | \mathbf{d}, I) \propto (2\pi)^{-\frac{M}{2}} \sigma_N^{-M} \exp\left(-\sum_{m=1}^{M} \frac{\left(\sum_{n=1}^{3} G_{mn} x_n - d_m\right)^2}{2\sigma_N^2} \right) \exp\left[-\frac{1}{2} \mathbf{x}^T \mathbf{W}_{\mathbf{x}} \mathbf{x} \right]. \tag{3.47}$$

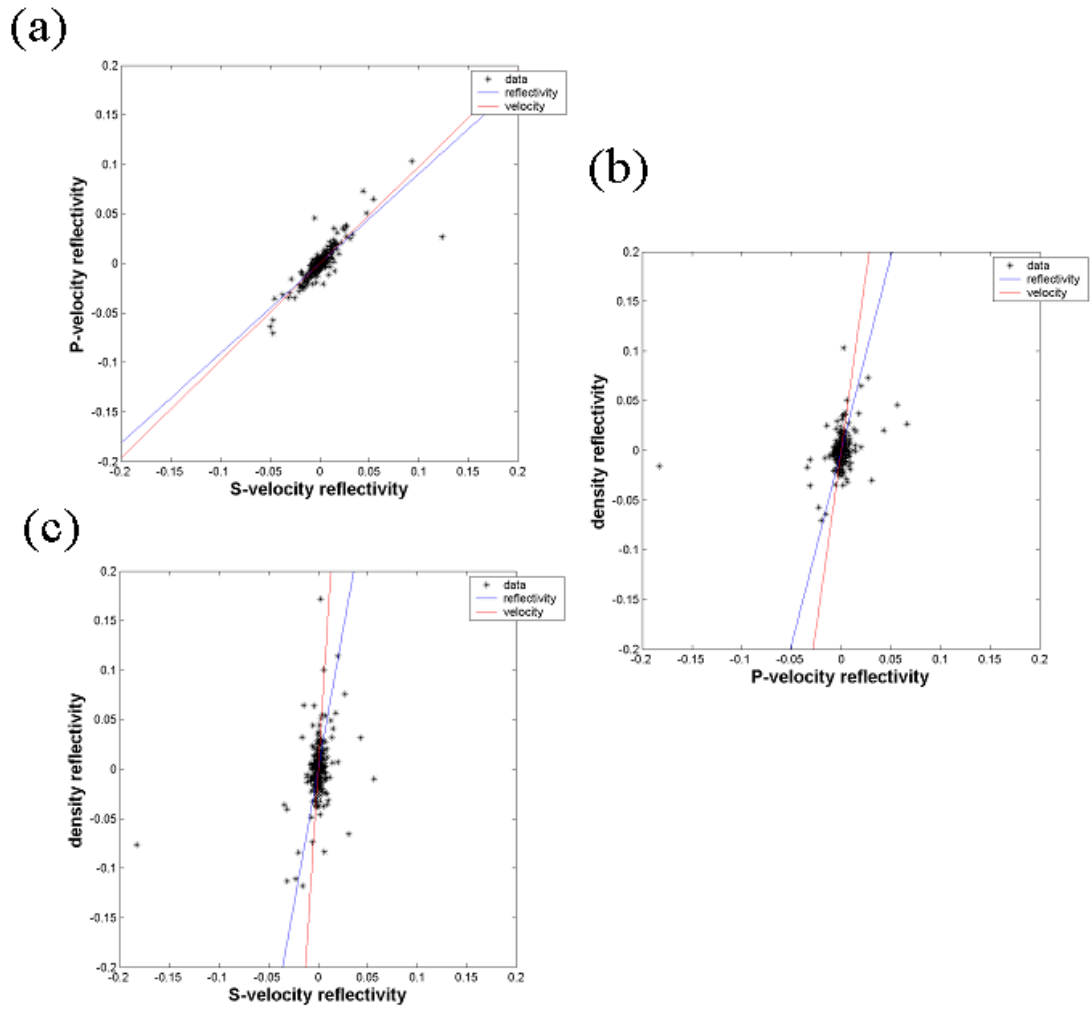

Figure 3.3: Crossplot of the P- and S-wave reflectivity for the Halfway well over the clastic interval. (a) The red line shows the trend generated from Fig. 3.2. The blue line indicates the best fit in a least squares sense between the reflectivity giving a slope of m=1.01. Figure (b) is a crossplot of the P-wave velocity and density reflectivity. The red line shows the parameters calculated using the Gardner equation while blue line indicates the best fit in a least squares sense giving a slope of g=0.2559. Figure (c) is a crossplot of the S-wave velocity and density reflectivity. The red line shows the parameters calculated using the Potter equation while blue line indicates the best fit in a least squares sense giving a slope of f=0. 0.1791.

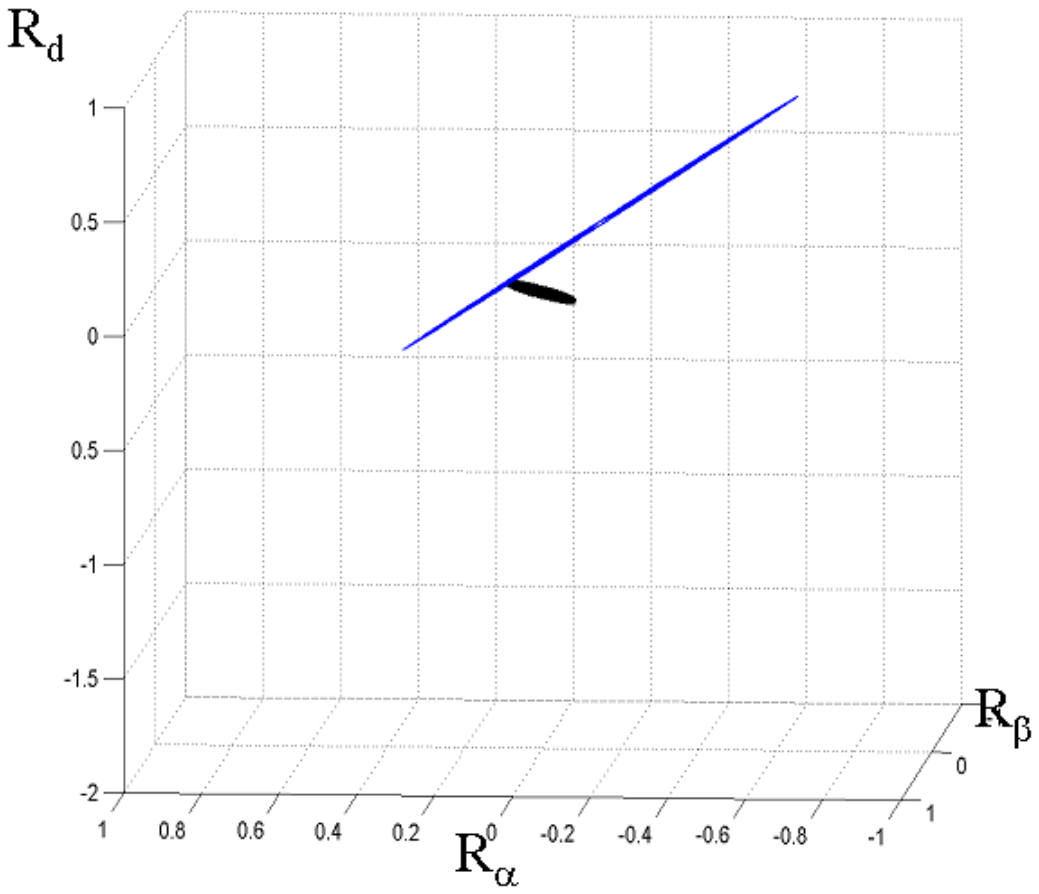

Figure 3.4: Blue surface is misfit ellipsoid generated from likelihood function based on equation (3.1) for the acquisition geometry used in the Blackfoot synthetic. The black surface is equi-probable surface generated by the constraints based on the multivariate Gaussian constraints. The optimal solution occurs where the combined probability is maximum.

Since there is no explicit interest in the variance σ_N, it is marginalized (Appendix E)

$$P(\mathbf{x}|\mathbf{d},I) \propto \exp\left[-\frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j\right]\left(\sum_{m=1}^{M}\left(\sum_{i=1}^{3} g_{mi}x_i - d_m\right)^2\right)^{-\frac{(M-1)}{2}}. \quad (3.48)$$

The most likely solution is found by finding where the probability function is stationary, $\frac{\partial P}{\partial x_i} = 0$, and convex. It is easier to determine this by taking the logarithm of the probability distribution

$$L(\mathbf{x}|\mathbf{d},I) \propto -\frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j - \frac{(M-1)}{2}\log\left(\sum_{m=1}^{M}\left(\sum_{i=1}^{3} g_{mi}x_i - d_m\right)^2\right). \quad (3.49)$$

Taking the gradient of equation (3.49) results in

$$\nabla L = -\mathbf{W_x x} - \frac{1}{\hat{\sigma}_N^2}\left[\mathbf{G}^T\mathbf{G}\mathbf{x} - \mathbf{G}^T\mathbf{d}\right], \quad (3.50)$$

where the estimate of the noise variance estimate is

$$\hat{\sigma}_N^2 = \frac{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{(N-1)}, \quad (3.51)$$

and

$$\boldsymbol{\varepsilon} = \mathbf{G}\hat{\mathbf{x}} - \mathbf{d}. \quad (3.52)$$

Evaluating equation (3.50) at the stationary point $\nabla\mathbf{L} = \mathbf{0}$ and using $\mathbf{W_x} = \frac{1}{\sigma_{r_p}^2}\tilde{\mathbf{C}}_\mathbf{x}^{-1}$ results in the nonlinear equation

$$\left[\mathbf{G}^T\mathbf{G} + \Theta\tilde{\mathbf{C}}_\mathbf{x}^{-1}\right]\mathbf{x} = \mathbf{G}^T\mathbf{d}, \quad (3.53)$$

where

$$\Theta = \frac{\hat{\sigma}_N^2}{\sigma_{r_p}^2}. \quad (3.54)$$

Equation (3.53) is of the form of a constrained least squares problem which is weakly nonlinear. The nonlinearity arises since Θ is a function of the misfit $\boldsymbol{\varepsilon}$ which must be estimated after one has an estimate of the parameters. The ratio Θ acts as a weighting factor determining how much the prior constraints influence the solution. The value of Θ is proportional to the noise-to-signal ratio. If the signal-to-noise is

large, then the weighting factor is small and the constraints add little to the solution and vice versa. In the extreme, when there is no noise, Θ is zero and the equation (3.53) reduces to the least squares formulation.

The simplest case where constraints play a role is when $\tilde{\mathbf{C}}_{\mathbf{x}}^{-\mathbf{1}} = \mathbf{I}$. This implies that the variances of all three reflectivity attributes are the same but uncorrelated. This case is equivalent to damped least squares or Tikhonov regularization (Hansen, 1998). This is what is advocated by Jin et al. (2000). Based on the analysis in the constraints section, it is expected at the very least that the variances of each of the reflectivity attributes are different. More importantly they are correlated. Incorporating this extra information leads to more geologically plausible solutions when the signal-to-noise ratio is poor.

Newton-Raphson solution

It was noted previously that equation (3.53) is weakly nonlinear due to the fact that Θ is dependent on the final solution. The probability distribution (equation 3.49) is a multi-variate Gaussian function multiplied by a multi-variate Student distribution. As long as the number of observations is reasonably high, the Student distribution is close to a Gaussian distribution and can be approximated well by a two-term Taylor series. The Newton-Raphson approach is well suited to solve this problem and converges to the optimal solution Q-quadratically (Nocedal and Wright, 1999) so that in this case the optimal solution is found in 2 or 3 iterations. The form of this solution is

$$\mathbf{x}_{N+1} = \mathbf{x}_N - \left[\boldsymbol{\nabla}\boldsymbol{\nabla} L\left(\mathbf{x}_N\right)\right]^{-1} \boldsymbol{\nabla} L\left(\mathbf{x}_N\right), \tag{3.55}$$

where the subscripts indicate the iteration number. The second derivative may be derived from the first derivative (equation 3.50)

$$\boldsymbol{\nabla}\boldsymbol{\nabla} L\left(\mathbf{x}_N\right) = -\mathbf{W}_{\mathbf{x}} - \frac{1}{\hat{\sigma}_N^2}\mathbf{G}^T\mathbf{G}. \tag{3.56}$$

Combining equations (3.50), (3.52) and (3.56) in (3.55) results in

$$\hat{\mathbf{x}}_{N+1} = \hat{\mathbf{x}}_N - \left[\mathbf{G}^T\mathbf{G} + \Theta\tilde{\mathbf{C}}_{\mathbf{x}}^{-1}\right]^{-1}\left[\mathbf{G}^T\boldsymbol{\varepsilon} + \Theta\tilde{\mathbf{C}}_{\mathbf{x}}^{-1}\mathbf{x}\right]. \tag{3.57}$$

If the incremental parameter step $d\mathbf{x}$ is defined by the equation

$$\mathbf{x}_{i+1} = \mathbf{x}_i - d\mathbf{x}, \tag{3.58}$$

then the incremental step is

$$d\mathbf{x} = \left[\mathbf{G}^T\mathbf{G} + \Theta\tilde{\mathbf{C}}_{\mathbf{x}}^{-1}\right]^{-1}\left[\mathbf{G}^T\boldsymbol{\varepsilon} + \Theta\tilde{\mathbf{C}}_{\mathbf{x}}^{-1}\mathbf{x}\right]. \tag{3.59}$$

Equation (3.58) and (3.59) are iteratively applied to estimate the parameter vector $\mathbf{x}$.

Scaling There is one further complication that needs to be addressed in order to have a working algorithm. This complication is that the processed seismic data are scaled by an unknown global scalar. In recording and processing the seismic data, even if great care is taken in processing the data in an amplitude friendly fashion, the data are still scaled by some arbitrary global scalar. The calculation of Θ (equation 3.54)

$$\Theta = \frac{\hat{\sigma}_N^2}{\sigma_{r_p}^2}, \tag{3.60}$$

is biased by this scalar. The estimate of the variance of the noise contains this scalar since it is estimated from the data. The denominator $\sigma_{r_p}^2$ does not since it is estimated from the well control. This scalar will either make Θ too large or small affecting the influence of the constraints. Other than the Θ term, the global scalar just scales the solution in a manner consistent with the data.

The problem of finding this global scalar is difficult since the problem as posed is underdetermined. The problem is multi-dimensional but reducing it to its most basic form one arrives at an equation of the form $rs = constant$ where r represents the reflectivity and s the scalar. If one increases the scalar s then the reflectivity decreases and vice versa. From the equation, one cannot tell if one solution is right. More information must be brought into the problem, for example specifying the variance of the reflectivity. This requires considering multiple interfaces (time samples) and treating the well information as data rather than *a priori* constraints complicating the analysis tremendously.

A pragmatic solution to this, is to redefine the definition of Θ so that

$$\Theta = \frac{\hat{\sigma}_N^2}{\hat{\sigma}_{r_p}^2}, \tag{3.61}$$

where $\hat{\sigma}_{r_p}^2$ is estimated from the data. We have assumed the P-wave impedance reflectivity is stationary so that an estimate of the P-wave impedance reflectivity variance may be made based on N time samples of the P-wave impedance reflectivity. The P-wave impedance reflectivity is chosen since it is the most reliable of the reflectivity attributes (Chapter 2) and least likely to change due to variations in acquisition geometry or signal-to-noise ratio. Since both the denominator and the numerator contain the scale factor, it divides out becoming a non-issue in the calculation of Θ.

3.2.6 Uncertainty analysis

The uncertainty of the parameter estimate is related to the width of the distribution. This may be estimated from the 2nd derivative evaluated at the parameter estimate (Sivia 1996, equation 3.71)

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = -\left[(\nabla\nabla L)^{-1}\right], \tag{3.62}$$

where $\mathbf{C}_{\hat{\mathbf{x}}}$ is the parameter covariance matrix of the estimate. With the assumption of uncorrelated uniform Gaussian noise the uncertainty is described by the covariance matrix

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \sigma_N^2 \left[\mathbf{G}^T\mathbf{G} + \Theta\tilde{\mathbf{C}}_{\mathbf{x}}^{-1}\right]^{-1}. \tag{3.63}$$

The diagonal of the covariance matrix (equation 3.63) represents the variance of each parameter estimate. The off-diagonal element represents the degree of correlation between the errors (Menke, 1984). Later, in the example section, the standard deviation, rather than the variance, is used to display the uncertainty so that the uncertainty is displayed in same units as that of the reflectivity for ease of comparison.

The case where $\Theta = 0$ is equivalent to the unconstrained uncertainty

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \sigma_N^2 \left[\mathbf{G}^T\mathbf{G}\right]^{-1}, \tag{3.64}$$

as calculated in Chapter 2 (equation 2.36). Since the three-term AVO problem is ill-conditioned the uncertainty of the unconstrained problem will be typically large.

Adding the constraints (equation 3.63) helps stabilize the inverse thus reducing the uncertainty. It is also important to understand how much the constraints influence the solution. Following Wijngarrden and Berkhout (1996), the uncertainty of the constrained solution is compared to the uncertainty of the estimate for the case where constraints are not included. The ratio of these two uncertainty estimates give a sense for how much the solution is being influenced from the data and how much by prior knowledge. To make accurate predictions about the subsurface the parameter estimate of interest should be largely coming from the data.

3.2.7 Transform matrix

In this thesis the P-wave and S-wave impedance, and density reflectivity are estimated. This is a rearrangement of the Aki and Richards (1980) formulation written in terms of P-wave and S-wave velocity, and density reflectivity. As reviewed in Chapter 2, there are numerous other rearrangements of equation (3.1) in the literature including Lamé reflectivity (Gray et al., 1999), and the Shuey parameters A, B, C (Shuey, 1985). This being the case, it is possible to use the transform matrix (equation B.55) to transform from impedance reflectivity to any of these other attributes. Further, the uncertainty of the transformed variables may be estimated using

$$\hat{\mathbf{C}}'_{\hat{\mathbf{x}}} = \mathbf{T}\hat{\mathbf{C}}_{\hat{\mathbf{x}}}\mathbf{T}^T. \tag{3.65}$$

In this way different AVO attributes can be examined to see how they show off some particular geologic feature or anomaly. An attribute can be selected which best highlights the objective. Of equal importance, the reliability of each of these attributes can also be examined to understand whether the anomaly under this variable is reliable or an artifact due to the noise.

3.3 Examples

The three-term AVO inversion developed in this chapter is demonstrated on both synthetic and real data. Both the synthetic data examples are generated based on two wells from Western Canada. The first example demonstrates how the inversion reacts to changes in angle range and signal-to-noise ratio. The second example was constructed so that the density reflectivity is poorly correlated with velocity reflectivity. The objective in this second synthetic example is to show that the

density reflectivity may be accurately estimated, even when not correlated with the velocity reflectivity, given signal-to-noise ratios and angle ranges found in typical seismic data. The real seismic data are in close proximity to the second synthetic example and demonstrates the technique on real data.

3.3.1 Blackfoot synthetic example

The first example is based on a well from the Blackfoot pool from Alberta, Canada introduced in Chapter 2. Based on these wireline logs a synthetic model was generated (Figure 2.15) using the Zoeppritz equations to model the reflectivity. The reflectivity was generated with no moveout so NMO stretch and tuning are not an issue. The reflectivity was convolved with a 10/14 - 70/80 Hz zero phase wavelet. Random noise was added to give a signal-to-noise ratio of 2:1. The zone of interest is the relatively low velocity Glauconitic sand from 0.98 to 1.01 seconds The model was generated with a variety of different acquisition geometries with different angle ranges and signal-to-noise ratios. The constraints were constructed based on a composite of logs in the area. Figure 3.5 shows crossplots between the various reflectivity sections for one of the wells. The following empirical relationships were observed

$$R_\alpha = 1.5R'_\beta \quad \text{with } r_{R_\alpha R_{\beta'}} = 0.79,$$

$$R_d = 0.4620R_\alpha \quad \text{with } r_{R_\alpha R_d} = 0.8633,$$

$$R_d = 0.420R_\beta.$$

from which the the normalized covariance matrix was constructed.

The results of the constrained inversion were output in terms of impedance reflectivity and compared to the corresponding zero offset reflectivity attributes. In order to obtain a reliable estimate of the density reflectivity we need to have seismic data with good signal-to-noise and large range of incidence angles. Figure 3.6 shows the case where the inversion was done with an adequate angle range from $0\,^\circ$ to $45\,^\circ$, with a good signal-to-noise ratio of 8:1. For this case the density is accurately estimated. Displayed with the reflectivity estimate is the uncertainty estimate calculated using equation (3.63). The uncertainty in this case is small. The second inversion (Figure 3.7) is over the same angle range, but the input data have a much poorer signal-to-ratio of 1/4. The estimated uncertainty is greater for all the reflectivity estimates. For this example the estimate of the density is poor above 0.85 seconds. This is

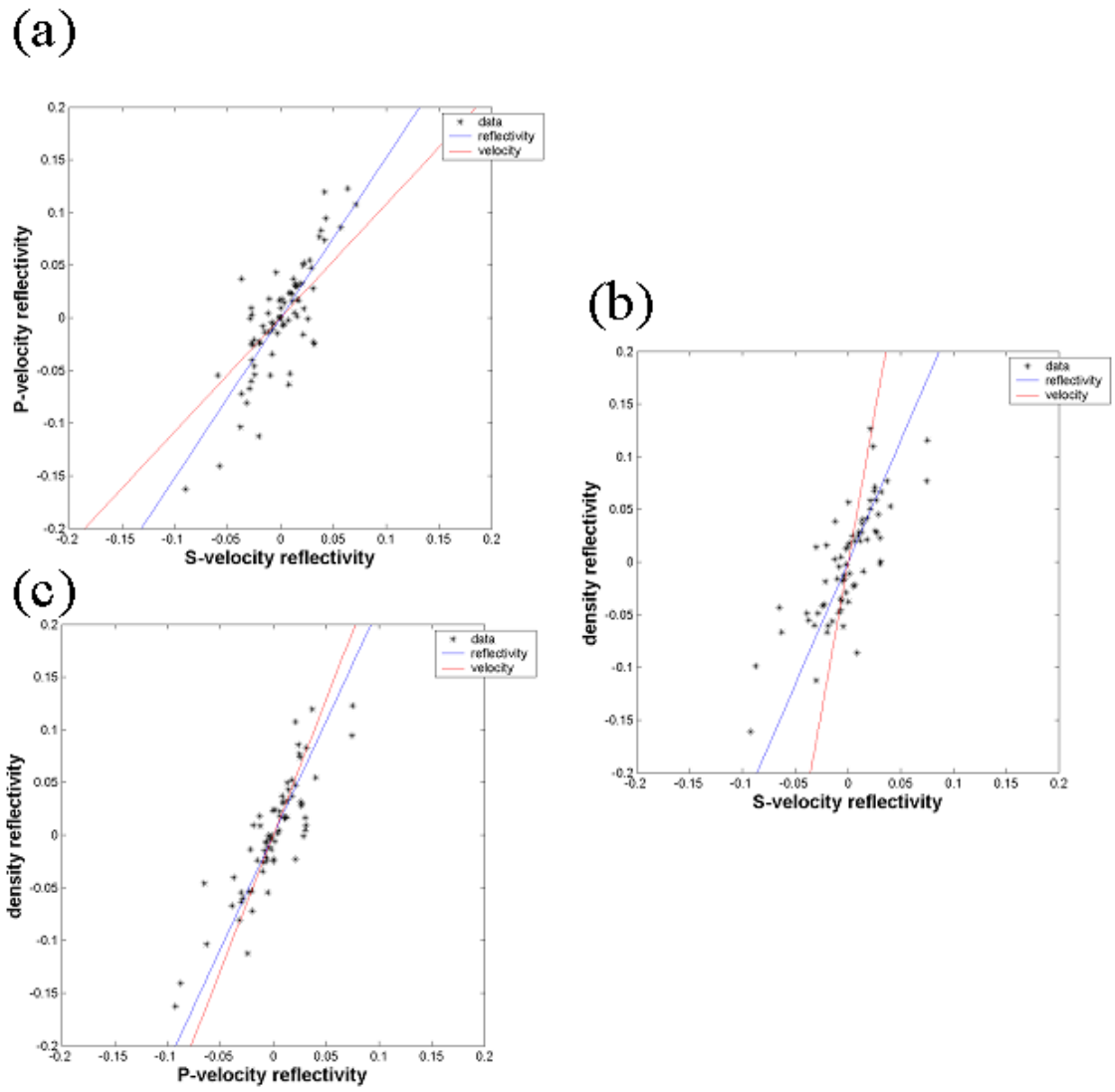

Figure 3.5: Cross-plots reflectivity used to construct constraints for Blackfoot AVO inversion.

due to two reasons. First the signal-to-noise is poorer due to the fact the signal is relatively smaller above 0.85 seconds. More importantly though, the constraints do not provide helpful information since velocity and density are uncorrelated (Figure 2.15). Below 0.85 seconds the density reflectivity estimate is more accurate due to the higher signal-to-noise ratio and the better applicability of the constraints. The last inversion shown (Figure 3.8) is for data with a good signal-to-noise ratio of 8:1, but over a limited angle range of $0^\circ - 28^\circ$. For this limited angle range, the problem is ill-conditioned and relies on the constraints to give a stable estimate. Note that the estimated uncertainty for the P-wave and S-wave impedance reflectivity is small while the density is larger. This is consistent with the errors observed between the ideal and the estimate. For times, $t > 0.9$ seconds, the density estimate is quite accurate. However, the solution is dominated by the *a priori* information with little information coming from the seismic data itself. For this case the density prediction has little predictive value. When interpreting the density reflectivity it is important to also look at the parameter uncertainty.

3.3.2 Halfway synthetic example

The next synthetic example is constructed from a well drilled for the Halfway in northeastern British Columbia, Canada. Once again, density, P-wave and S-wave velocity logs were available from this well. For modeling purposes, the well logs were converted to time based on the P-wave velocity log (Figure 3.9). In this well, the density is uncorrelated with P-wave velocity log at a number of different times. The Bluesky top is the velocity increase and density decrease at 0.67 seconds. The density behaves dramatically different from the P-wave velocity at the Bluesky top and at 0.32 seconds. Additionally, both these events are poorly correlated when looking at the corresponding reflectivity. Figure 3.10 shows the Gardner density estimate (equation 3.32) based on the P-wave velocity reflectivity. This estimate is quite poor as the difference and crossplot (Figure 3.11) shows. This model provides a good test of how well the density is estimated even when the Gardner relationship is violated.

The model is constructed using the Zoeppritz equations to generate the reflectivity. No moveout is applied so NMO stretch and tuning may be ignored. These will be dealt with in a later chapter. The reflectivity is convolved with a 10/14 - 90/110 Hz wavelet and muted so no angles greater than 45° are shown. Noise is added to give a signal-to-noise ratio of 4:1.

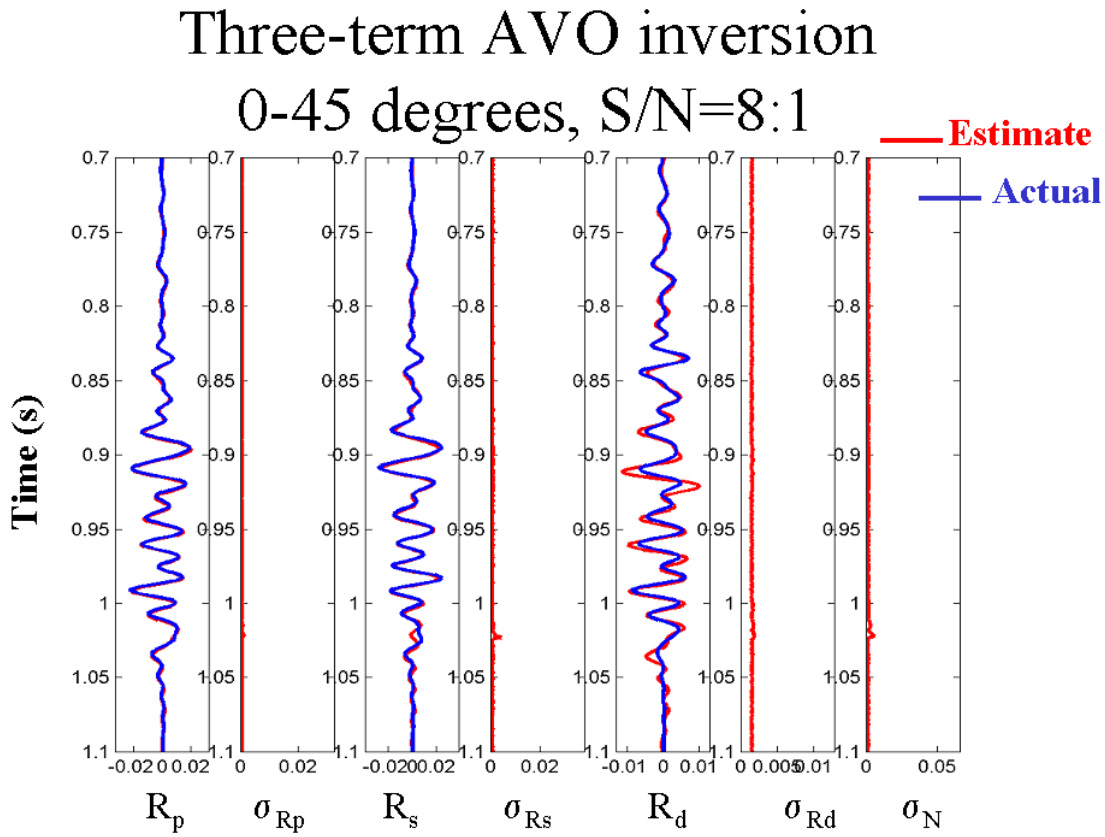

Figure 3.6: Results of AVO inversion from 0 to 45 degrees for P-impedance, S-impedance and density reflectivity attributes on a gather with a S/N=8. The estimate of the uncertainty for each reflectivity estimate along with noise is also shown.

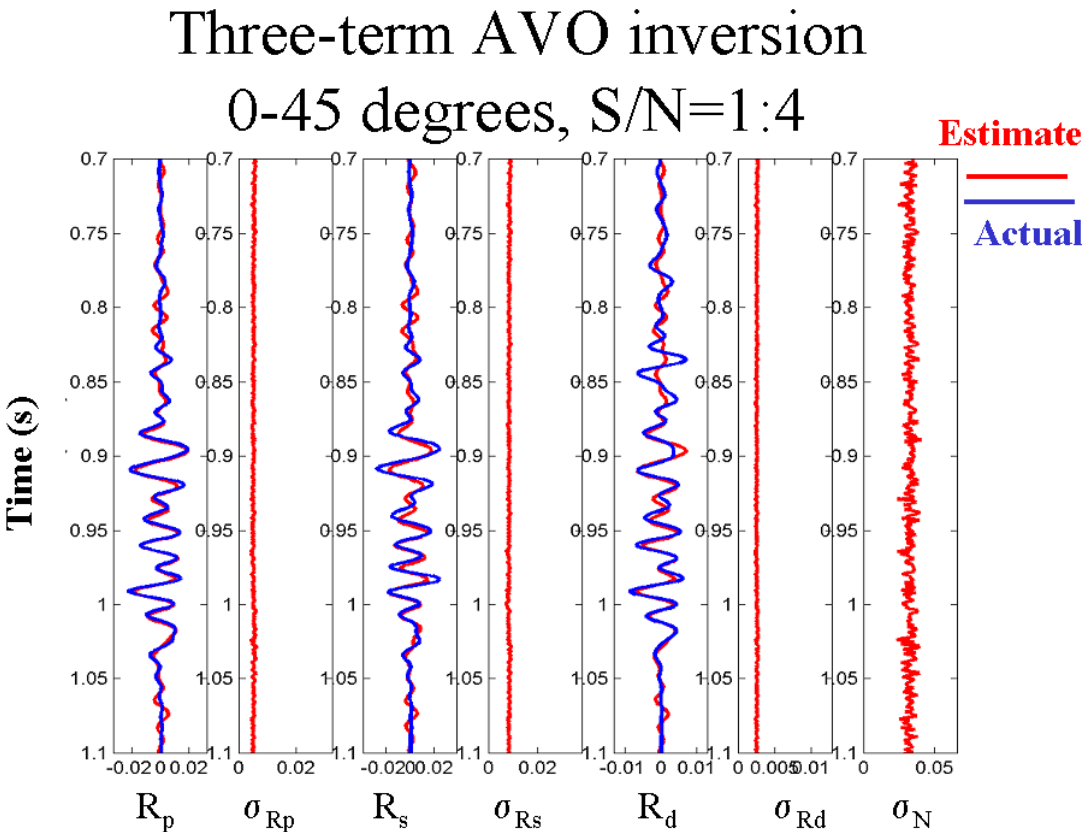

Figure 3.7: Results of AVO inversion from 0 to 45 degrees for P-impedance, S-impedance and density reflectivity attributes on a gather with a S/N=1/4. The estimate is in red and the actual reflectivity in blue. The estimate of the uncertainty for each reflectivity estimate along with noise is also shown.

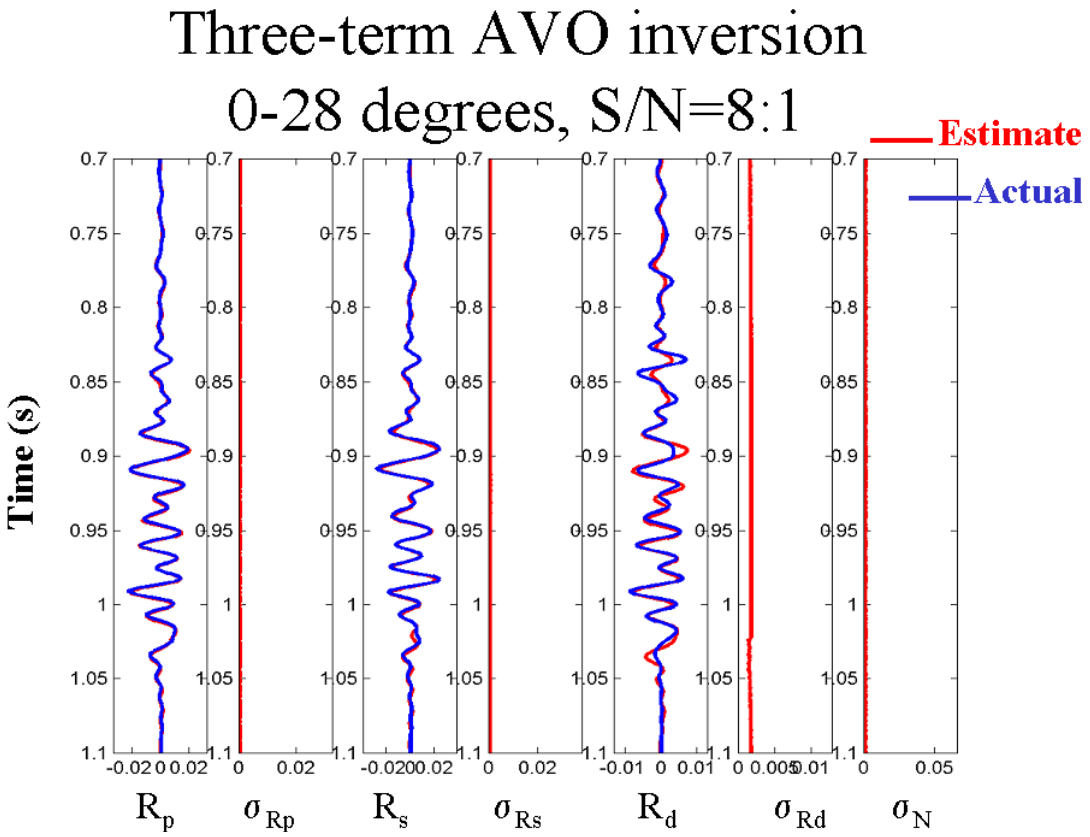

Figure 3.8: Results of AVO inversion from 0 to 28 degrees for P-impedance, S-impedance and density reflectivity attributes on a gather with a S/N=8. The estimate is in red and the actual reflectivity in blue. The estimate of the uncertainty for each reflectivity estimate along with noise is also shown.

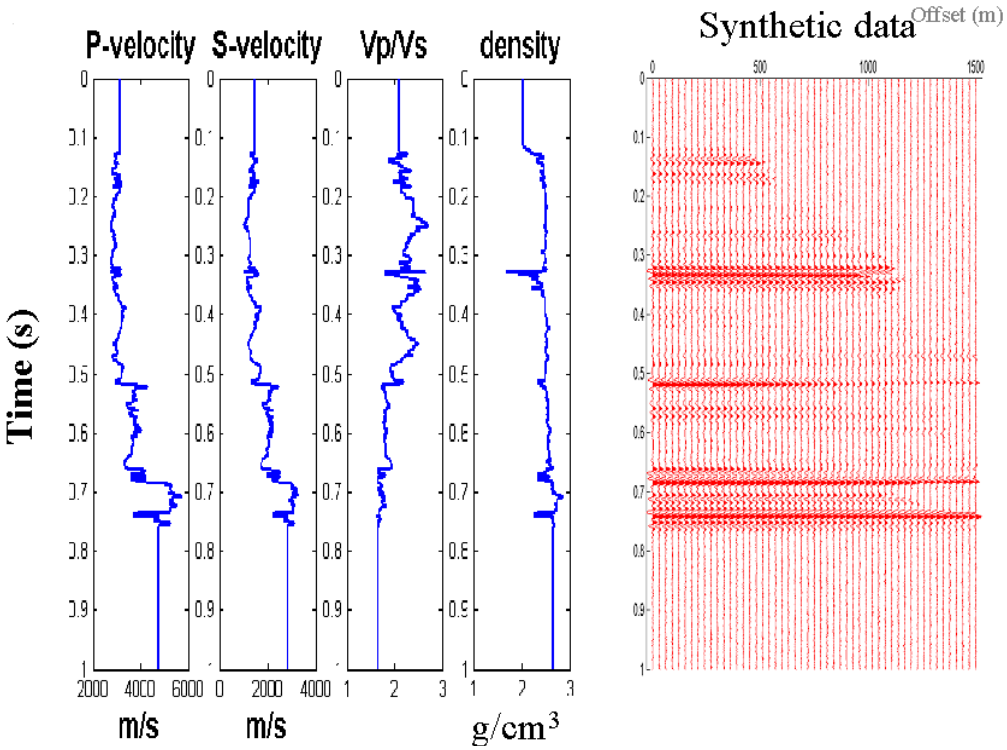

Figure 3.9: Wireline logs displayed in time for the Halfway well log. The synthetic gather was generated using reflectivity generated from the Zoeppritz equation. The reflectivity is shown without moveout, filtered and with noise added to give a S/N=4.

The synthetic model was inverted to estimate the P-wave and S-wave velocity, and density reflectivity. The parameters previously derived in Figure 3.3 were used to constrain the solution. In Figure 3.12, the reference zero offset reflectivity is shown in blue while the estimate from the constrained three-term AVO inversion is shown in red. There is little difference between the ideal and estimate even at 0.32 and 0.68 seconds where the P-wave velocity and density reflectivity are uncorrelated.

To test the influence of constraints, the inversion was redone with another set of constraints based on the following parameters $m = 1.14, g = 0.20, f = 0.10$ with the same correlation coefficients used before. The results shown in Figure 3.13 are almost identical to the previous results shown in Figure 3.12.

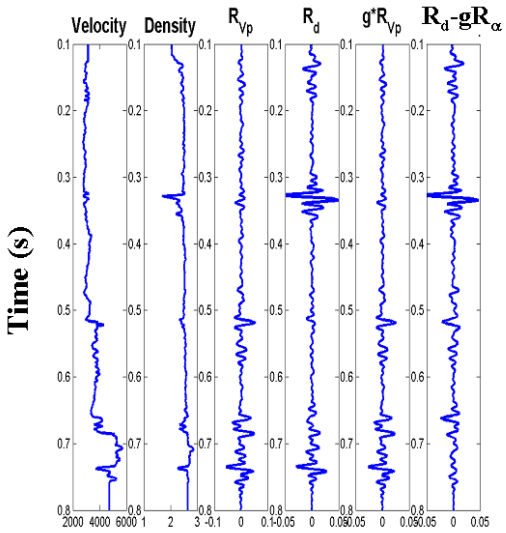

Figure 3.10: P-wave velocity and density of Halfway well log. Note that density and velocity are uncorrelated at 0.32 and 0.68 seconds. Note the corresponding reflectivity is as well. Density estimated using the Gardner equation $R_d = gR_\alpha$ poorly correlates with the actual density.

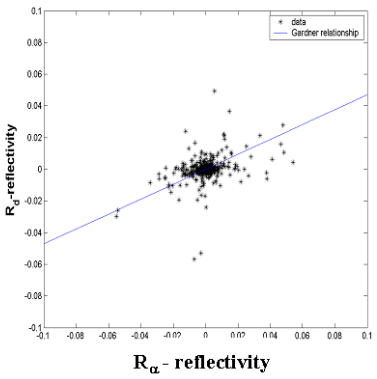

Figure 3.11: Cross-plot of P-wave velocity and density reflectivity. Note the large amount of scatter and the poor correlation.

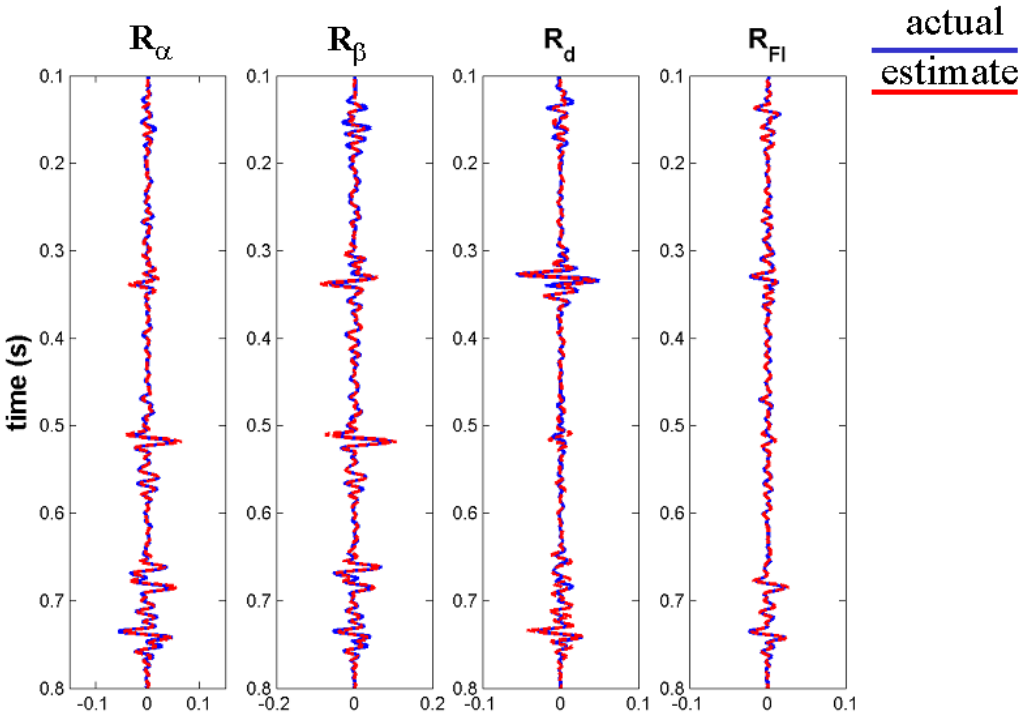

Figure 3.12: Comparision of three-term AVO inversion for the P- and S-wave velocity, density and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue).

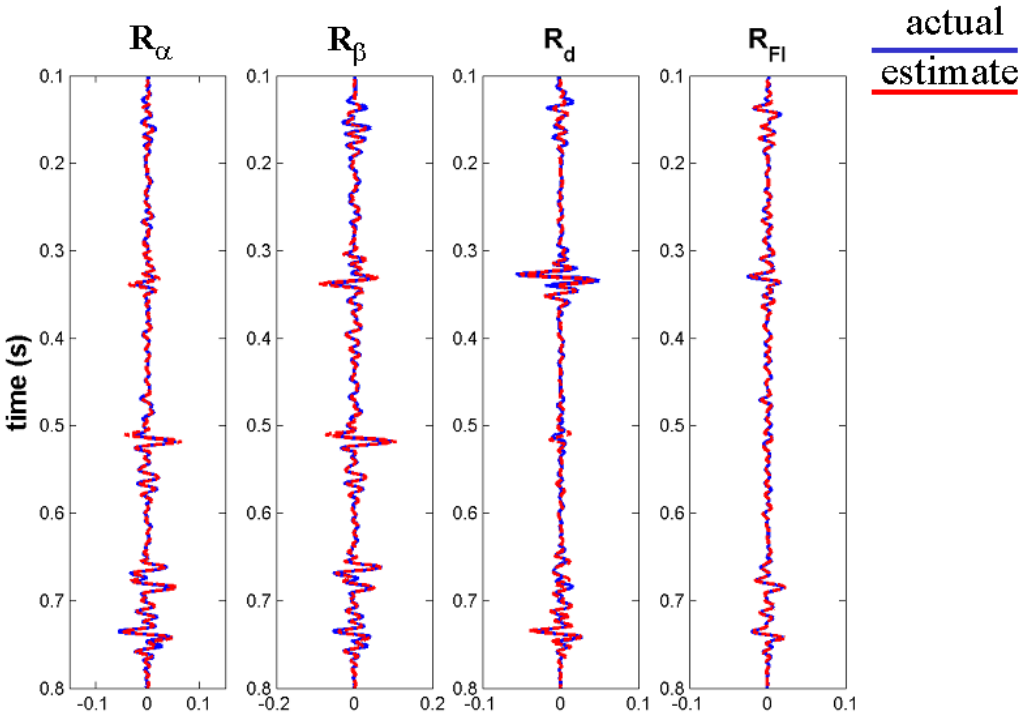

Figure 3.13: Comparision of three-term AVO inversion with another set of constraints for the P- and S-wave velocity, density and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue).

3.3.3 Halfway seismic data example

The inversion was run on a line which was part of a project acquired to explore for Halfway sand potential (Downton and Tonn, 1997). In general, the seismic data have a good signal-to-noise ratio. Offsets are long enough so the AVO inversion is performed on angles out to 45°. Forward modeling indicates that within this range supercritical reflections are not encountered. Figure 3.14 shows the P-wave and S-wave impedance, and density reflectivity estimates for three-term constrained inversion for the line. There are two bright spots on the line, evident on the P-wave impedance reflectivity section, at around 0.72 seconds The producing field (wells C and E) should have both a relatively low velocity and density response, hence large reflectivity responses. The uneconomic gas sand at well A has low gas saturation implying that density should not respond to the gas while still having a large velocity response due to the gas. This is evident on the reflectivity sections. The P-wave impedance reflectivity section shows bright spots at both locations. The density reflectivity section shows a weaker anomaly at the well with uneconomical gas saturations (well A) than the producing field.

Figure 3.15 shows the standard deviation and the influence of the constraints on the density reflectivity estimated. The standard deviation is shown with the same color scale as the reflectivity section so the two can easily be compared, though note the standard deviation is only positive. Around CMP 2000 the seismic line traversed muskeg resulting in poorer quality records. The standard deviation is much larger in this area. Because of the poorer signal-to-noise ratio, the constraints are weighted more in the solution. This influence shows in the ratio of constrained to unconstrained uncertainty (Figure 3.15, bottom panel). However, for most of the zone of interest the solution is largely coming from the data. It is also evident that in areas with more limited aperture, such as the end of the line or deeper in the section, the constraints play a larger role, as expected.

Figure 3.16 shows a comparison between the scaled P-wave impedance and density reflectivity sections. The P-wave impedance reflectivity section is scaled in a manner similar to equation (3.32) so as to give a crude estimate of density. Recall this is the assumption used in the Smith and Gidlow (1987) formulation to estimate density. The lower panel shows a difference between these two density estimates. It is evident by the large difference that the third term is contributing information independent from equation (3.32). This is also true for the S-wave impedance reflectivity. In

summary, it appears that the third term predicting density is providing independent information. From the uncertainty displays it appears that at the bright spots the prediction is largely being driven by the seismic data, that the uncertainty is tolerable. This combined with the good geologic correlation with the well control suggest the density section is reasonable.

Figure 3.17 shows how the methodology helped even improve on a traditional two-term inversion. This shows a comparison of the estimate of P-wave impedance reflectivity section generated by the constrained three-term AVO inversion and one generated by a two-term Gidlow equation. The two-term least squares inversion was performed using angles from 0° to 33°. The three-term inversion gives a smoother more geologically plausible result in the area contaminated by noise around CMP 2000. This is due to the inclusion of the constraints. In this area where the signal-to-noise ratio is weak, the *a priori* information dominates the solution.

3.4 Discussion

3.4.1 VTI anisotropy

In this chapter various wave propagation and processing complications have been ignored in order to explore and solve the simpler problem of how to solve the ill-conditioned problem of inverting for three reflectivity attributes. For example, the earth was assumed to be isotropic. If it is actually anisotropic, then the linearized approximation of the Zoeppritz equations needs to be modified to account for this. The linearized approximation for VTI media (Ruger 2002, equation 4.34)

$$R_p^{VTI}(i) = R_p + \left\{ R_\alpha - \left(2\frac{\bar{V}_{S0}}{\bar{V}_{P0}} \right)^2 R_\mu + \frac{\Delta\delta}{2} \right\} \sin^2\bar{\theta} + \left\{ R_\alpha + \frac{\Delta\epsilon}{2} \right\} \sin^2\bar{\theta}\tan^2\bar{\theta} \quad (3.66)$$

is underdetermined with only three of the five parameters being uniquely solved for. In equation (3.66) the reflectivity attributes R_p and R_α are modified so that they represent reflectivity for vertical propagation. The parameter R_μ is the rigidity reflectivity attribute for vertical wavefield propagation while $\Delta\delta = (\delta_2 - \delta_1)$ and $\Delta\epsilon = (\epsilon_2 - \epsilon_1)$ represent the difference in the Thomsen (1986) anisotropy parameters delta and epsilon across the interface. If anisotropy is present, then ray tracing and gain corrections also need to be modified from those of the isotropic assump-

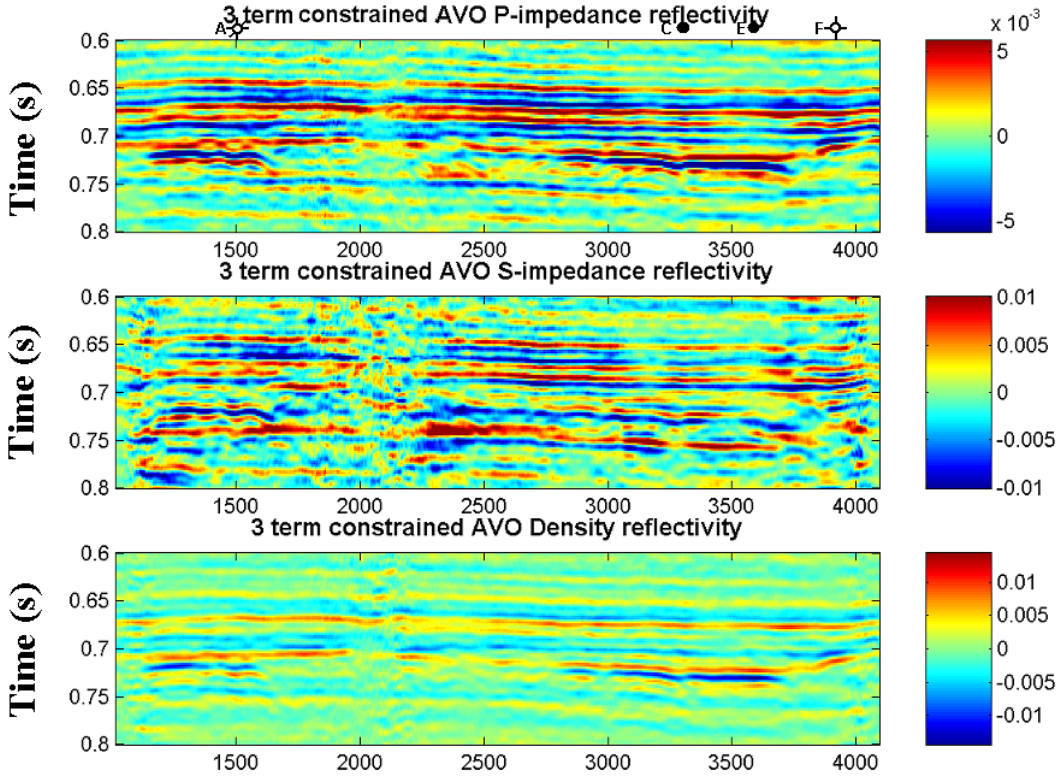

Figure 3.14: P-impedance, S-impedance and density reflectivity attribute inversions over producing and non-economic gas fields. Note that it is possible to differentiate on the density section the low gas saturation gas well (light blue at Well A at 0.72s) from the economic gas wells (dark blue Well C and E at 0.72s).

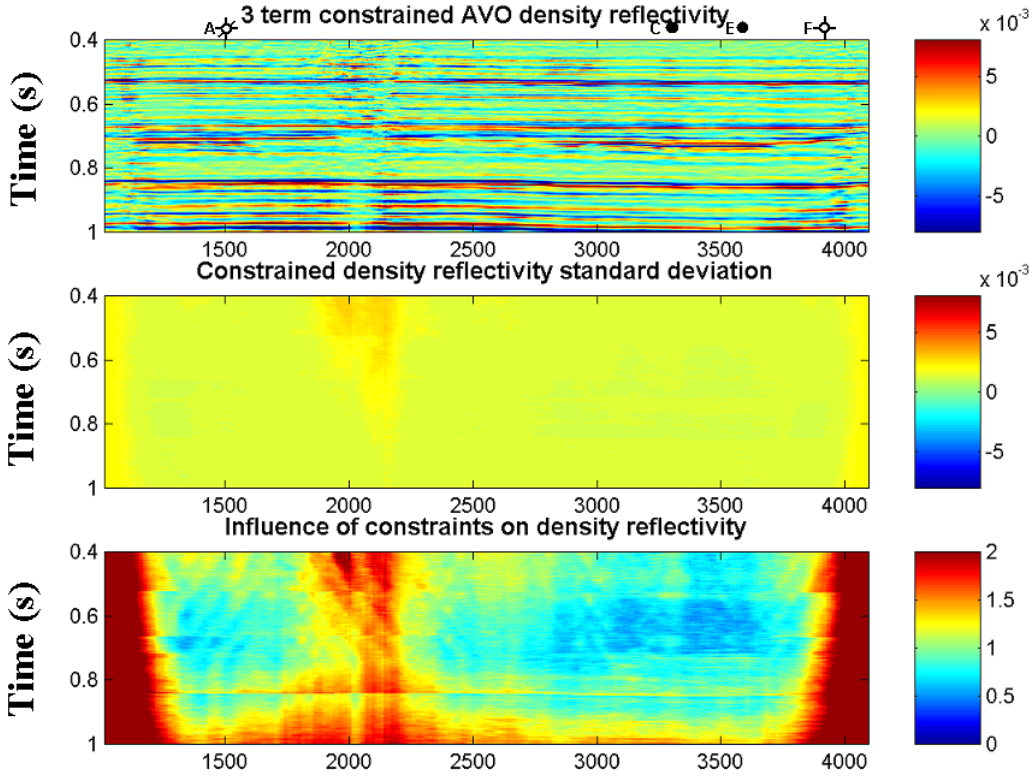

Figure 3.15: The density reflectivity and related quality control sections. The standard deviation of the density (middle panel) is considerably smaller than the density reflectivity at the zone of interest. The ratio of the unconstrained to constrained uncertainty (bottom panel) shows the influence of the constraints on the solution. Where this ratio is high, the constraints are dominating the solution. This occurs when the S/N is poor or the range of angles available for the inversion is limited.

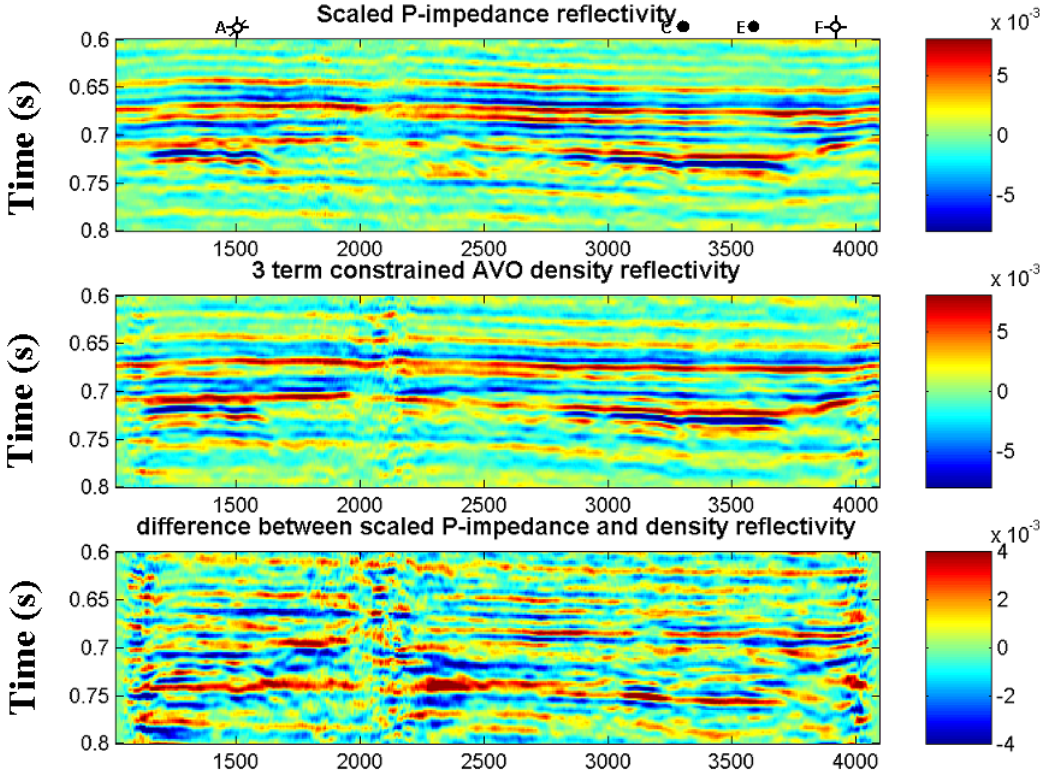

Figure 3.16: The scaled P-impedance and density reflectivity attributes shown. The P-impedance is scaled in a manner suggested by the Gardner constraint to try and estimate density. The bottom panel is a difference between the two.

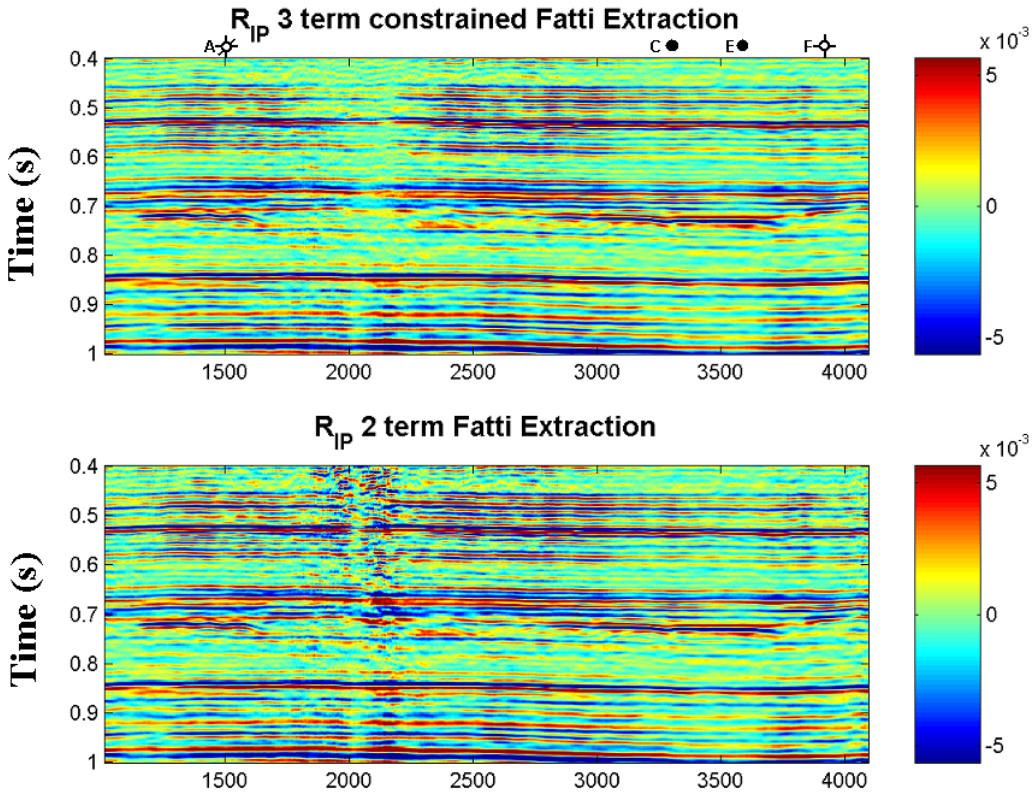

Figure 3.17: The P-impedance inversion from the constrained three-term AVO inversion (top) and from a least squares two-term AVO inversion (bottom). Note how the constraints improve the solution on the three-term inversion in the area with a poor signal-to-noise ratio around CMP 2000.

tion. If anisotropy is present and equation (3.1) is used to invert for density or any other reflectivity attribute the estimates at best will be biased and at worst seriously incorrect.

Even in an isotropic earth there are other complicating factors that can be significant. Offset dependent transmission losses are difficult to correct for and bias the amplitude as a function of offset leading to incorrect estimates. The linearized Zoeppritz equations do not model supercritical reflections well. The AVO inversion code has logic to ignore any supercritical reflections identified by the ray tracing. However, supercritical reflections arising due to large changes in the micro-velocity model are difficult to predict leading to biased estimates due to the theoretical error.

Lastly, offset dependent tuning and NMO stretch will bias the AVO estimates. Further, the band-limited nature of the seismic data are not considered in this inversion. The inversion is performed on a sample by sample basis assuming that each inversion is not coupled. In the following chapters the ideas in this chapter are further developed incorporating these additional factors into the inversion.

3.4.2 Nonuniform noise

Further, this algorithm was developed assuming uniform Gaussian random noise. Often for real data, the noise there will be not be uniform or Gaussian. The noise will have outliers implying long tailed distributions. In Chapter 8 it is shown how the likelihood function may be modified to incorporate these long tailed distributions so the inversion will behave in more robust fashion in the presence of this nonuniform noise with outliers.

3.5 Conclusions

This chapter demonstrated a three parameter AVO inversion incorporating probabilistic constraints from the available geologic control. The degree to which the constraints influence the solution is a function of the signal-to-noise ratio of the data and the acquisition geometry. The constraints preferably should be calculated from local well control. If local well control is not available, values from the literature may be used. The probabilistic constraints introduce less bias in the estimates than the hard constraints implicitly implemented in the two-term AVO inversions. The results of the three-term constrained AVO inversion are equivalent to the Smith and Gidlow

AVO inversion if the *a priori* constraints define the density reflectivity as a linear function of the P-wave velocity reflectivity. Similarly, the results of the inversion are equivalent to the two-term Gidlow equation if the *a priori* information specifies the density reflectivity is zero. Lastly the three-term AVO inversion is equivalent to the two-term Shuey equation with the *a priori* constraint that the P-wave velocity reflectivity is zero. By choosing constraints based on local well control, honoring known rock physical relationships, and weighting the constraints based on the needs of the data, the results of the constrained three parameter AVO inversion should be more accurate than the aforementioned methods. Impedance and density reflectivity are solved for, but can be transformed subsequently to virtually any other AVO attribute. Thus, the methodology is more general than any two-term AVO inversion and produces estimates having less bias.

Parameter uncertainty estimates are provided as part of the derivation and should be examined to determine the significance and reliability of a particular AVO attribute. This is particularly true for the density reflectivity since reliably estimating this attribute requires large incidence angles and good signal-to-noise. The variance of a particular parameter can be examined to determine its reliability while the ratio of the unconstrained to constrained variance conveys information about whether the seismic data or geologic control is dominating the solution. Although the constraints will create reflectivity sections that are geologically plausible, for an anomaly to be believable, solutions should be driven by the seismic data and not dominated by the constraints.

Chapter 4

Error in the AVO linear operator

4.1 Introduction

In Chapter 2 a method was developed to estimate uncertainty in the parameter estimates due to uncertainty in the data. This uncertainty estimate is overly optimistic for among other things it ignores uncertainties in the linear operator. Tarantola (1987) calls these types of errors theoretical errors, or errors arising from using inexact theory. In Chapter 1 it was shown that it is possible to linearize the AVO problem providing $\bar{\gamma}$ (the average S-wave to P-wave velocity ratio) and the average incidence offset-to-angle relationship is known. There is uncertainty associated with both these assumptions. In addition, instead of solving the exact Zoeppritz equations, some linearized, constrained approximation is solved for, again introducing error. This chapter examines and quantifies these errors.

Analytic expressions are developed to quantify the uncertainty associated with both $\bar{\gamma}$ and the ray tracing. Two-term rather than three-term AVO approximations are used since this simplifies the analysis. This introduces some error into the calculation but results in relatively simple expressions which can be intuitively understood. The two-term Gidlow equation (2.20) is used to quantify the uncertainty associated with $\bar{\gamma}$ while the two-term Shuey equation (2.14) is used to quantify the error associated with the ray tracing. In the ray tracing analysis, the Walden (1991) approach to ray tracing is used since it is only a function of the P-wave interval velocity and the stacking velocity at the time sample under consideration, again simplifying the analysis. Ignoring the C term in the two-term Shuey approximation leads to bias in the estimate of the gradient B. This bias is discussed in section 4.2. The two-term Shuey parameter estimates may be compared with the two-term Gidlow parameter

estimates using the two-term transform matrix developed in Appendix B.9 thus making it is possible to study the influence of both $\bar{\gamma}$ and the ray tracing uncertainty on the P-wave and S-wave impedance reflectivity estimates. It would have been preferable to study the uncertainty in ray tracing on the two-term Gidlow equation to keep things consistent and avoid biased estimates but this leads to intractable solutions.

For the two-term AVO inversions examined, the uncertainty lies in the second parameter estimated (the second term being the gradient B for Shuey, or, for Gidlow et al. (1992) the S-wave impedance reflectivity R_s). The intercept or P-wave impedance reflectivity is unaffected by these modeling errors. Each of these errors lead to both systematic and random errors. The random errors are an order of magnitude less than the reflectivity estimates themselves. For realistic noise found in real seismic data the influence of these errors is negligible compared to that of the noise. The systematic errors are more problematic. These tend to introduce a time variant scaling error in the gradient B or S-wave impedance reflectivity R_s.

Throughout this chapter the Blackfoot synthetic model, introduced in Chapter 2, is used to illustrate the influence of errors associated with the linear operator. The first section quantifies the impact of using two-term AVO approximations instead of the exact Zoeppritz equations for the inversion. Parameter estimates from the two-term Shuey and Gidlow inversions are compared with the ideal reflectivity attributes (which are generated by convolving the particular reflectivity series by the source wavelet to get the zero offset reflectivity attribute).

In the second section, analytic expressions predicting the uncertainty due to ray tracing errors are developed. These predictions are then verified by a series of modeling experiments. These results are compared to biased estimates obtained in the previous section, quantifying the error due to uncertainties in the ray tracing. In Section 4.4 analytic expressions are developed to quantify the impact of uncertainties in $\bar{\gamma}$. Once again these predictions are verified by a series of modeling experiments. Lastly the influence of errors in both $\bar{\alpha}$ and $\bar{\gamma}$ are studied.

4.2 Modeling error

Even if three-term linearized AVO inversion is performed using the correct $\bar{\gamma}$ and angles of incidence there will be error in the reflectivity attribute estimates. This is a result of using a linearized approximation rather than the actual nonlinear Zoeppritz equations in the inversion. In this chapter, instead of using a three-parameter

formulation, two-parameter formulations are used, further increasing the modeling error. Section 2.4.2 demonstrated that truncating the third term is equivalent to implementing a hard constraint. The solution must exist on a two-dimensional surface defined by that constraint. The best estimate occurs where the misfit probability function reaches a maximum on that surface. If the actual solution is not on this surface, the constraint has effectively introduced bias into the solution. The amount of bias is dependent on the geometry of the misfit function, the constraint used, and the location of the ideal solution. Figure 2.11 shows the bias introduced into the P-wave velocity reflectivity estimates by the different constraints. The geometry of the misfit function is relative to some parameterization. Thus, the amount of bias may be reduced by choosing certain parameterizations over others. For example, parameterizing the problem in terms of impedances results in the introduction of less relative bias than parameterizing in terms of velocity reflectivity (Section 2.4.1).

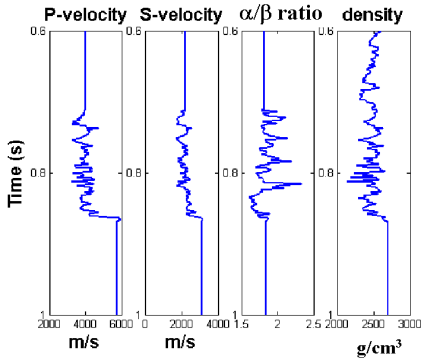

Figure 4.1: Blackfoot well logs used to generate synthetic model.

To quantify the error due to using a two-term AVO inversion, a synthetic modeling study was performed using the Blackfoot well logs (Figure 4.1) first introduced in Chapter 2. Synthetic gathers were generated using the Zoeppritz equations to generate the reflectivity. The angle of incidence was calculated using the classical ray tracing approach. The synthetic gather was generated without normal moveout so NMO stretch and offset dependent tuning would not introduce distortions. The reflectivity was band-pass filtered using a 10/15-90/100 Hz filter. Unlike Chapter 2, no random noise was introduced since the primary interest is modeling error. The AVO inversion was performed using angles from 0^0 to 30^0. The Walden approach

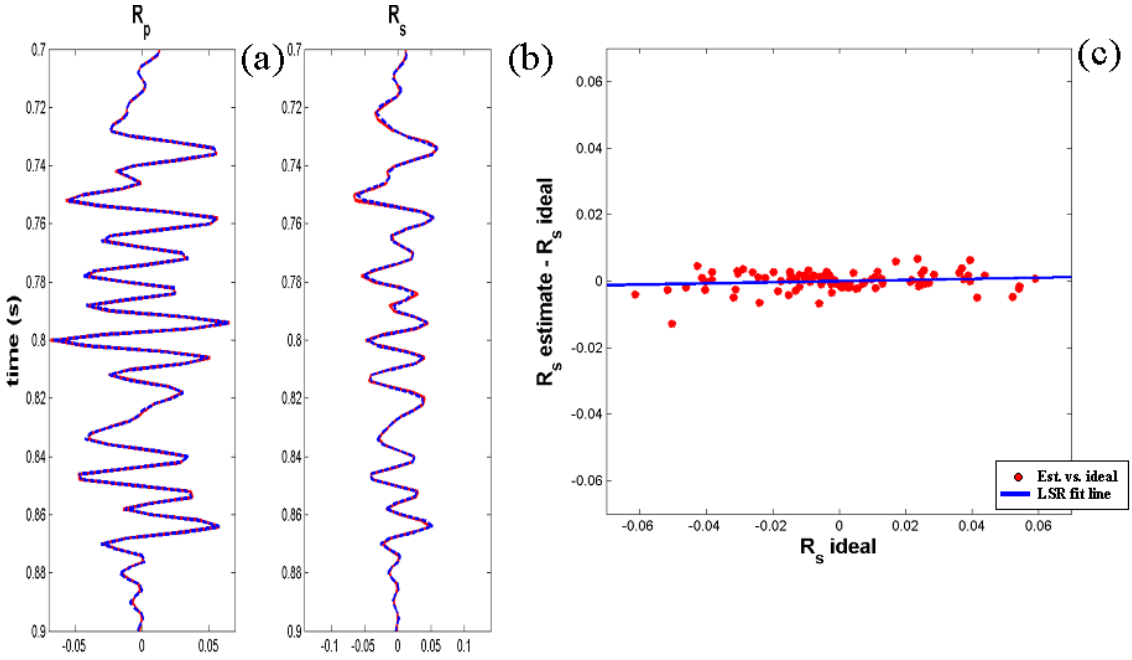

Figure 4.2: Comparison of ideal (red) and estimated (blue) reflectivity for the two-term Fatti inversion. The cross-plot shows the difference between the ideal and estimated S-impedance reflectivity.

(section 1.2.2) to ray tracing was used for the AVO inversion. In this case, the exact high frequency P-wave interval and stacking velocities were used. Figure 4.2 shows the results of the two-term Gidlow inversion. The estimated reflectivity is in red while the zero offset reflectivity is in blue. Upon close inspection it is evident that there is a small error between the estimated and ideal S-wave impedance reflectivity. To quantify this, the difference between the estimated and ideal S-wave impedance reflectivity is crossplotted versus the ideal S-wave impedance reflectivity. The line that best fits the estimated and ideal S-wave impedance reflectivities is also calculated along with its correlation coefficient. The slope of this line is identical to the scalar needed to be applied to make the estimate best fit the ideal reflectivity. The greater the departure of the scalar from unity the greater the systematic bias. The results of this section are tabulated in Table 4.1. In this case, the scalar is close to unity suggesting that the two-term Gidlow inversion has little bias. The correlation coefficient indicates the amount of scatter and may be used as a measure of random error in the solution.

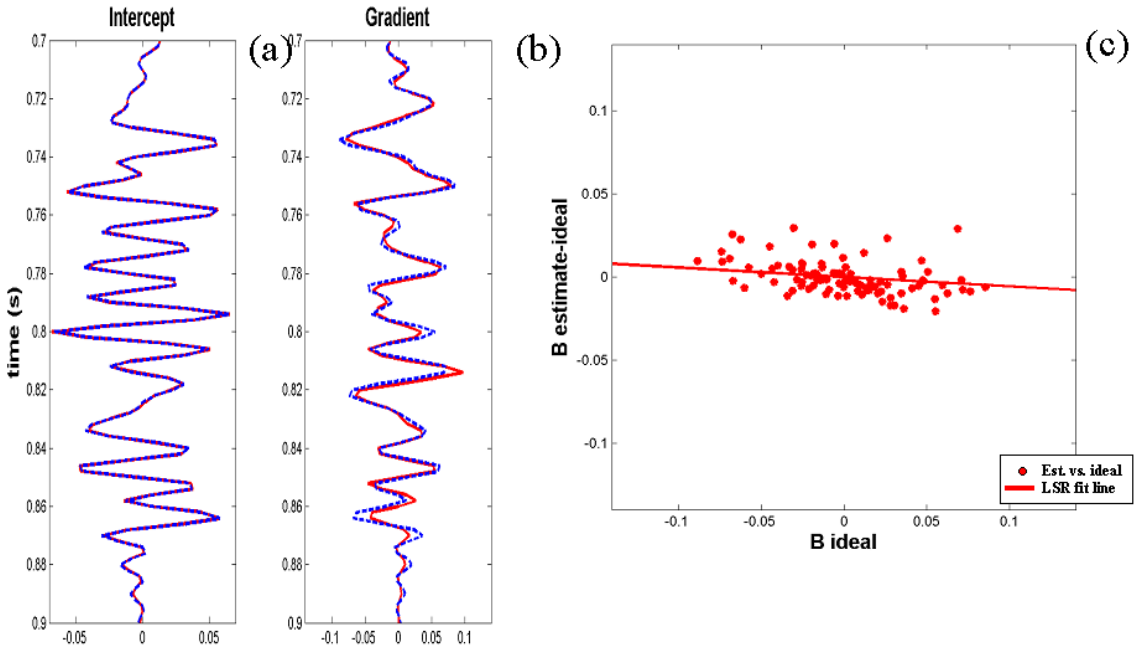

Figure 4.3: Comparison of ideal (red) and estimated (blue) reflectivity for two-term Shuey inversion. The cross-plot (c) shows the difference between the ideal and estimated gradient reflectivity. Note the bias in the gradient estimate.

	Shuey B	Shuey R_s	Gidlow R_s
scalar	0.9436	0.9300	1.0169
Correlation coefficient	0.9530	0.9834	0.9846

Table 4.1: Modeling error due to using a two-term linearized approximation instead of the Zoeppritz equation.

Figure 4.3 shows the results of the two-term Shuey inversion. It is evident that the estimated gradient is scaled lower than the ideal. This is confirmed by the least squares scalar 0.93 being smaller than unity. The gradient B is biased since the third term in C is ignored. If smaller angle ranges are used such as 0^0 to 22^0 the bias decreases, but it is still significant. For an inversion done with angles from 0^0 to 22^0 the scalar is 0.9653. In subsequent examples I felt it better to use parameter ranges (0^0 to 30^0) that are used by AVO practitioners rather than trying to decrease the bias by using arbitrarily small angle ranges. The two-term Shuey correlation coefficient 0.9530 for the gradient B is significantly worse than the correlation coefficient for the S-wave impedance reflectivity from two-term Gidlow equation.

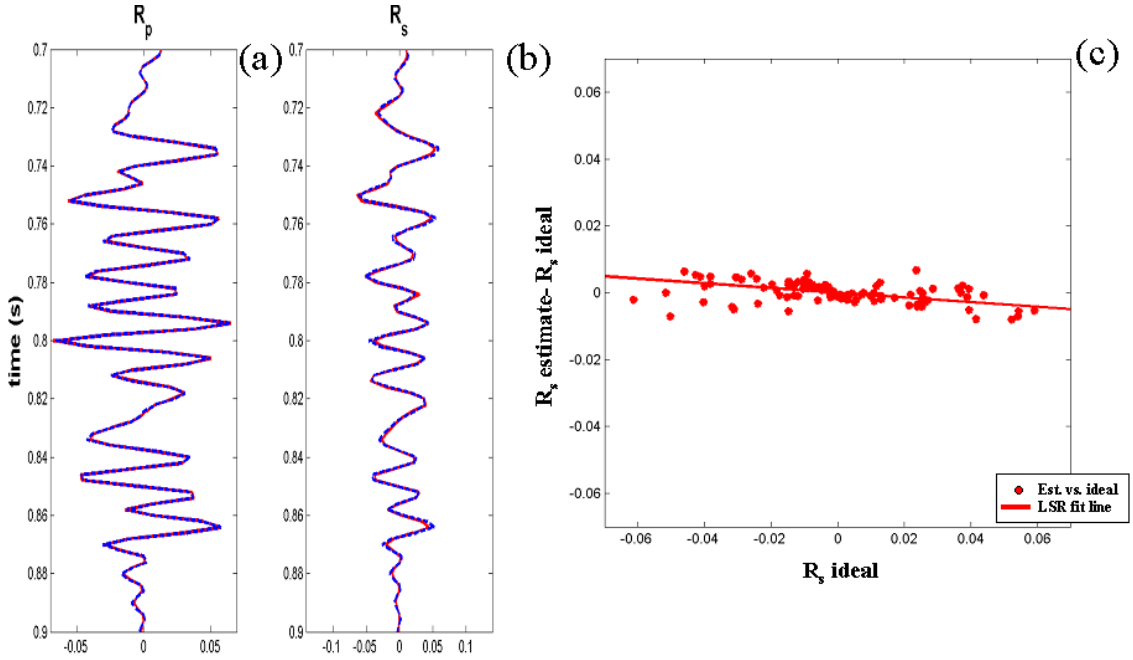

Figure 4.4: Comparison of ideal (red) and estimated (blue) reflectivity for two-term Shuey inversion converted to impedance. The cross-plot (c) shows the difference between the ideal and estimated S-impedance reflectivity.

However, this comparison is unfair since different reflectivity attributes are being compared. Equation (B.60)

$$R_s = \frac{1}{8\bar{\gamma}^2}\left[\left(\frac{g}{1+g}\left(4\bar{\gamma}^2-1\right)+1\right)A - B\right]. \tag{4.1}$$

may be used to convert the output of the two-term Shuey inversion to S-wave impedance reflectivity. Using $g = 0.3305$, which is calculated via least squares from the well control, the estimates were transformed to impedance reflectivities (Figure 4.4). The correlation coefficient of the transformed estimate R_s is 0.9834 virtually identical to the result from the two-term Gidlow inversion. The transform decreases the random error. However, the systematic error in the transformed reflectivity estimate remains. The transformed S-wave impedance reflectivity is a biased estimate since the gradient B is biased. It has a scalar of 0.930. Thus, the transformed two-term Shuey inversion gives results similar to the two-term Gidlow inversion, though the S-wave impedance reflectivity is slightly biased.

4.3 Ray tracing error

In this section the modeling error due to mapping the offset to angle of incidence is explored. An analytic expression describing the error is derived for the two-term Shuey formulation

$$R\left(\bar{\theta}_m\right) = A + B\sin^2\bar{\theta}_m. \tag{4.2}$$

where $\bar{\theta}_m$ is the average angle of incidence. The two-term Shuey formulation is chosen as the starting point since it leads to a relatively simple and intuitively understandable expression, the results of which may be transformed to S-wave impedance reflectivity using equation (4.1). In the modeling results section, the two-term Shuey and Gidlow S-wave impedance reflectivity are compared.

The Walden approach to ray tracing is used to map the offsets to angle-of-incidence. Recall from Section 1.2.2 that using this approach, the angle of incidence

$$\sin\bar{\theta}_m = \frac{\bar{\alpha}\left(t_0\right)h_m}{V_{stak}^2\left(t_0\right)\tau_m}, \tag{4.3}$$

is a function of the offset h_m, the two-way travel time τ_m to the interface at offset h_m, the average P-wave interval velocity across the interface $\bar{\alpha}\left(t_0\right)$, and stacking velocity at the interface $V_{stak}\left(t_0\right)$. In the majority of this thesis, the classical ray tracing approach is used as it is slightly more accurate. This accuracy comes at a cost of a much more complex error analysis, since the angle of incidence is a function of all the overburden layers' P-wave interval velocities. Figure 4.5 shows that for the Blackfoot synthetic example, the difference between the two approaches is negligible. Note the difference is shown with an altered color scale to highlight discrepancies. Otherwise they would appear to be identical. The reason the two approaches are similar is that for this particular case the P-wave velocity is constant down to 0.7 seconds. Since there is no velocity gradient, higher order terms are not needed to describe the kinematics, thus the stacking velocity describes the horizontal slowness p with a high degree of accuracy.

Substituting equation (4.3) into (4.2) results in the two-term Shuey equation written in terms of the P-wave interval and stacking velocity

$$R\left(h_m\right) = A + B\left(\frac{\bar{\alpha}h_m}{V_{stak}^2\tau_m}\right)^2. \tag{4.4}$$

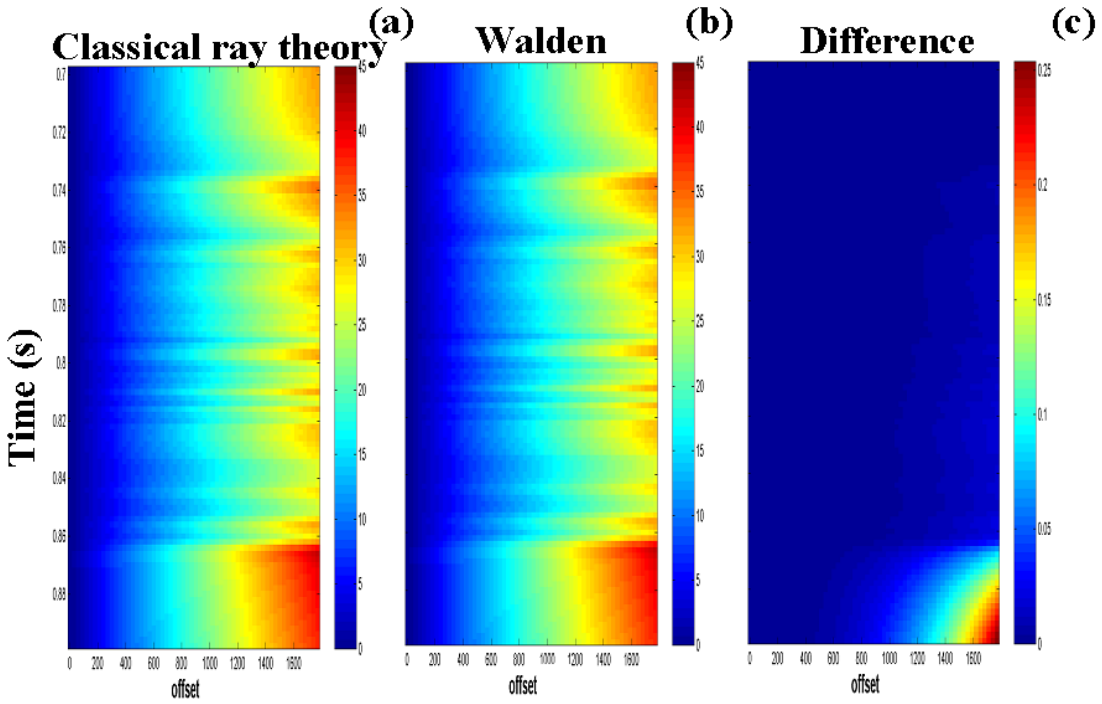

Figure 4.5: Angle of incidence generated by ray tracing Blackfoot model using Classical ray theory approach (a) and the Walden approach (b). The difference between the two (c) is neglible.

The error analysis is performed in two ways. First, equation (4.4) is expanded as a Taylor series expansion to linear order about $\bar{\alpha}_{back}$ and V_{back}, where V_{back} and $\bar{\alpha}_{back}$ are the smoothed stacking velocity and smoothed background average P-wave interval velocity used in the ray tracing respectively. The error due to uncertainty in $\bar{\alpha}$ and V_{stak} then appears as an extra term similar to noise. The linear model is thus modified

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}_d + \mathbf{n}_T, \tag{4.5}$$

to include an extra noise term due to theoretical error $\mathbf{n}_T$. The vector $\mathbf{n}_d$ represents noise in the data. Equation (4.5) may be solved by least squares. The advantage of this formulation is it leads to a convenient way to analyze errors coming from both the data and the theory (Section 4.6.3). The second method expands the least squares solution as a Taylor series expansion to linear order about $\bar{\alpha}_{back}$ and V_{back}. This is computationally simpler than the first method and is used to analyze errors in $\bar{\gamma}$.

Both methods give equivalent results for the ray tracing error in this section.

Method 1

An error analysis may be performed by expanding equation (4.4) as a Taylor series expansion about $\bar{\alpha}_{back}$ and $\bar{V}_{stak}$ to linear order

$$R = R_0 + \left.\frac{\partial R}{\partial \bar{\alpha}}\right|_{R_0} \Delta\alpha + \left.\frac{\partial R}{\partial V_{stak}}\right|_{R_0} \Delta V_{stak}, \tag{4.6}$$

where $R_0\left(h_m\right)$ is the solution at $\bar{\alpha}_{back}$ and V_{back}, $\Delta\alpha = \bar{\alpha} - \bar{\alpha}_{back}$, and $\Delta V_{stak} = V_{stak} - V_{back}$. Later in the section, it is shown that the $\bar{\alpha}_{back}$ and V_{back} should be running averages of α and V_{stak} to minimize systematic error.

For simplicity, only errors due to perturbations from the background interval velocity are initially considered, noting that

$$\left.\frac{\partial R}{\partial \bar{\alpha}}\right|_{R_0} = 2\frac{\bar{\alpha}_{back} h_m^2}{V_{back}^4 \tau_m^2}, \tag{4.7}$$

thus

$$R\left(h_m\right) = A + B\frac{\bar{\alpha}_{back}^2 h_m^2}{V_{back}^4 \tau_m^2}\left(1 + 2\frac{\Delta\alpha}{\bar{\alpha}_{back}}\right). \tag{4.8}$$

For M offsets equation (4.8) is written in matrix form as

$$\begin{bmatrix} 1 & \frac{\bar{\alpha}_{back}^2 h_1^2}{V_{back}^4 \tau_1^2}\left(1 + 2\frac{\Delta\alpha}{\bar{\alpha}_{back}}\right) \\ \vdots & \vdots \\ 1 & \frac{\bar{\alpha}_{back}^2 h_M^2}{V_{back}^4 \tau_M^2}\left(1 + 2\frac{\Delta\alpha}{\bar{\alpha}_{back}}\right) \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_M \end{bmatrix}, \tag{4.9}$$

where d_m is the reflectivity at offset h_m. To linear order in $\Delta\alpha$, the least squares solution of this is (Similar to Appendix D.1)

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{D}\begin{bmatrix} \left(\sum_{m=1}^M \frac{h_m^4}{\tau_m^4}\right)\left(\sum_{j=1}^M d_j\right) - \left(\sum_{m=1}^M \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^M d_j \frac{h_j^2}{\tau_j^2}\right) \\ \frac{V_{back}^4}{\bar{\alpha}_{back}^2}\left(1 - 2\frac{\Delta\alpha}{\alpha_{back}}\right)\left(M\sum_{m=1}^M d_m \frac{h_m^2}{\tau_m^2} - \left(\sum_{m=1}^M \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^M d_j\right)\right) \end{bmatrix}. \tag{4.10}$$

where

$$D = M\sum_{m=1}^M \frac{h_m^4}{\tau_m^4} - \left(\sum_m^M \frac{h_m^2}{\tau_m^2}\right)^2. \tag{4.11}$$

Equation (4.10) may be written as the sum of two parts

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A_0 \\ B_0 \end{bmatrix} + \begin{bmatrix} \Delta A \\ \Delta B \end{bmatrix}, \tag{4.12}$$

the part evaluated at $\bar{\alpha} = \bar{\alpha}_{back}$

$$\begin{bmatrix} A_0 \\ B_0 \end{bmatrix} = \frac{1}{D} \begin{bmatrix} \left(\sum_{m=1}^{M} \frac{h_m^4}{\tau_m^4}\right)\left(\sum_{j=1}^{M} d_j\right) - \left(\sum_{m=1}^{M} \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^{M} d_j \frac{h_j^2}{\tau_j^2}\right) \\ \frac{V_{back}^4}{\bar{\alpha}_{back}^2}\left(M\sum_{m=1}^{M} d_m \frac{h_m^2}{\tau_m^2} - \left(\sum_{m=1}^{M} \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^{M} d_j\right)\right) \end{bmatrix}, \tag{4.13}$$

and the deviation from this

$$\begin{bmatrix} \Delta A \\ \Delta B \end{bmatrix} = -2\frac{1}{D}\frac{\Delta\alpha}{\bar{\alpha}_{back}} \begin{bmatrix} 0 \\ \frac{V_{back}^4}{\bar{\alpha}_{back}^2}\left(M\sum_{m=1}^{M} d_m \frac{h_m^2}{\tau_m^2} - \left(\sum_{m=1}^{M} \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^{M} d_j\right)\right) \end{bmatrix}, \tag{4.14}$$

or

$$\begin{bmatrix} \Delta A \\ \Delta B \end{bmatrix} = -2\frac{\Delta\alpha}{\bar{\alpha}_{back}} \begin{bmatrix} 0 \\ B_0 \end{bmatrix}. \tag{4.15}$$

Method 2

A simpler way to derive this is to start with the least squares solution of the Shuey equation (Appendix D.1) written in terms of stacking velocity and interval velocity

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{D} \begin{bmatrix} \left(\sum_{m=1}^{M} \frac{h_m^4}{\tau_m^4} \sum_{m=1}^{M} d_m - \left(\sum_{m=1}^{M} \frac{h_m^2}{\tau_m^2}\right)\left(\sum_{j=1}^{M} d_j \frac{h_j^2}{\tau_j^2}\right)\right) \\ \frac{V_{stak}^4}{\bar{\alpha}^2}\left(M\sum_{m=1}^{M} d_m \frac{h_m^2}{\tau_m^2} - \left(\sum_{m=1}^{M} \frac{h_m^2}{\tau_m^2}\right)\sum_{j}^{M} d_j\right) \end{bmatrix}. \tag{4.16}$$

and perform a Taylor series expansion about the background P-wave velocity $\bar{\alpha}$

$$B = B_0 + \left.\frac{dB}{d\bar{\alpha}}\right|_{\bar{\alpha}=\bar{\alpha}_{back}} \Delta\alpha. \tag{4.17}$$

again with $\Delta\alpha = \bar{\alpha} - \bar{\alpha}_{back}$. Note the intercept A has no $\bar{\alpha}$ dependence while the gradient B is a function of the interval P-wave velocity. Thus, defining the error as

$\Delta B = B - B_0$,

$$\Delta B = \left.\frac{dB}{d\bar{\alpha}}\right|_{\bar{\alpha}=\bar{\alpha}_{back}} \Delta\alpha, \tag{4.18}$$

resulting in

$$\Delta B = -2\frac{\Delta\alpha}{\bar{\alpha}_{back}} B_0, \tag{4.19}$$

which is equivalent to equation (4.15).

4.3.1 Error due to uncertainty in interval P-wave velocity

Rewriting equation (4.15)

$$\begin{bmatrix} \Delta A \\ \Delta B \end{bmatrix} = -2\frac{\Delta\alpha}{\bar{\alpha}_{back}} \begin{bmatrix} 0 \\ B_0 \end{bmatrix}. \tag{4.20}$$

it is clear that the uncertainty in the interval P-wave velocity does not influence the intercept A estimate; however, the gradient B is influenced. The magnitude of the fractional error of the gradient is two times the velocity perturbation of the background P-wave velocity

$$\frac{\Delta B}{B_0} = -2\frac{\Delta\alpha}{\bar{\alpha}_{back}}. \tag{4.21}$$

This error may be either random and/or systematic. In Section 1.2.2 it was pointed out that a smooth background P-wave velocity model is used in ray tracing the model. A smooth model is used since high frequency *a priori* information about the velocity is not available prior to the inversion. The interval P-wave velocity model is constructed from either extrapolating velocity information from nearby well control or generating it from some sort of travel time inversion. Because of the scarcity of information, and the error associated with this information, significant error can exist in the background velocity field. This error will vary for each reflector or time sample. The error analysis is facilitated by breaking the error into a high and low frequency component. Ignoring the high frequency perturbations in the velocity field results in random errors. The low frequency error introduces systematic error.

The random error may be analyzed by noting that the random perturbations $\Delta\alpha$ from the average velocity $\bar{\alpha}_{back}$ (equation 4.21) are to a first order equivalent to the P-wave velocity reflectivity (Walden and Hosken, 1988)

$$R_\alpha = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}} \approx \frac{\Delta\alpha}{\bar{\alpha}_{back}},$$

thus

$$\frac{\Delta B}{B_0} = -4R_\alpha. \tag{4.22}$$

In Chapter 3, it is argued that P-wave velocity reflectivity is Gaussian, thus the random fractional error introduced by ignoring the high frequency information in the background velocity model in the ray tracing is Gaussian and much less than one.

More problematic are systematic errors. For the moment, we ignore high frequency errors. In this analysis $\bar{\alpha}_{back}$ is treated as a slowly varying perturbation from the actual average. If this is not the case, then $\Delta\alpha$ represents the systematic error. The fractional perturbation $\Delta\alpha/\bar{\alpha}_{back}$ from the correct average acts as a time variant scalar

$$B = B_0 + \Delta B = (1 - 4R_{\bar{\alpha}})\, B, \tag{4.23}$$

thus the gradient is either scaled too large or small. This may be intuitively understood by noting that the ray tracing maps the offsets to angle of incidence. If this mapping squeezes the x-axis too much, the slope (gradient B) will be too large. If the mapping stretches the x-axis too much, the slope (gradient B) will be too small.

4.3.2 Error in terms of S-wave impedance reflectivity

The output of the two-term Shuey inversion can be transformed to output S-wave impedance reflectivity using equation (4.1). The S-wave impedance reflectivity is a function of A, B, and $\bar{\gamma}$ so the uncertainty associated with this is governed by

$$dR_s\,(A, B, \bar{\gamma}) = \frac{\partial R_s}{\partial A} dA + \frac{\partial R_s}{\partial B} dB + \frac{\partial R_s}{\partial \bar{\gamma}} d\bar{\gamma}. \tag{4.24}$$

The differences dA and dB are calculated in equation (4.15). The partial derivatives $\frac{\partial R_s}{\partial B}$ and $\frac{\partial R_s}{\partial \bar{\gamma}}$ may be calculated from equation (4.1) resulting in

$$\frac{\partial R_s}{\partial B} = -\frac{1}{8\bar{\gamma}^2}, \tag{4.25}$$

$$\frac{\partial R_s}{\partial \bar{\gamma}} = \frac{1}{4\bar{\gamma}^3}\left[B - \frac{A}{1+g}\right], \tag{4.26}$$

thus

$$dR_s = \frac{1}{4\bar{\gamma}^2}\left(\frac{d\bar{\alpha}}{\bar{\alpha}} B + \left(B - \frac{A}{1+g}\right)\frac{d\bar{\gamma}}{\bar{\gamma}}\right). \tag{4.27}$$

This relation is discussed in greater detail in Section 4.4.

4.3.3 Errors in background interval velocity and stacking velocity

Previously, the analysis was restricted to understanding the uncertainty due to errors in the average P-wave interval velocity. In this section uncertainty due to errors in the stacking velocity is also considered. The least squares solution is given by equation (4.16). In order to calculate the uncertainty, the solution is expanded as a Taylor series about the average P-wave interval and stacking velocity. Once again, the intercept A is not a function of the stacking velocity or background velocity so it is not influenced by uncertainty in these parameters. The uncertainty is thus

$$B = B_0 + \left.\frac{\partial B}{\partial \bar{\alpha}}\right|_{B_0} \Delta\alpha + \left.\frac{\partial B}{\partial V_{stak}}\right|_{B_0} \Delta V_{stak}. \tag{4.28}$$

The partial derivative for $\left.\frac{\partial B}{\partial \bar{\alpha}}\right|_{B_0}$ may be calculated from equation (4.21). The partial derivative $\left.\frac{\partial B}{\partial V_{stak}}\right|_{B_0}$ may be calculated from equation (4.16) and is

$$\left.\frac{\partial B}{\partial V_{stak}}\right|_{B_0} = \frac{4}{V_{back}} B_0, \tag{4.29}$$

so therefore

$$\Delta B = \left(-2\frac{\Delta\alpha}{\bar{\alpha}_{back}} + 4\frac{\Delta V_{stak}}{V_{back}}\right) B_0. \tag{4.30}$$

Provided that the stacking velocity is not used to estimate the interval velocity, the error due to the background interval velocity should be uncorrelated to that of the stacking velocity. Similar to the preceding analysis there can be both random and systematic error. Depending on the quality of the data, the stacking velocity should be picked within an accuracy of 1% to 3% of the actual stacking velocity and the error should be random. Systematic error might be introduced if the picks are biased, for example by a multiple. Further, if the stacking velocity is poorly picked or preliminary stacking velocities are used, the stacking velocity could also be biased.

Note, this analysis is only considering errors in the linear operator. If there are stacking velocity errors there could also be residual moveout and systematic errors in the data. This will lead to errors in the gradient (Spratt, 1987).

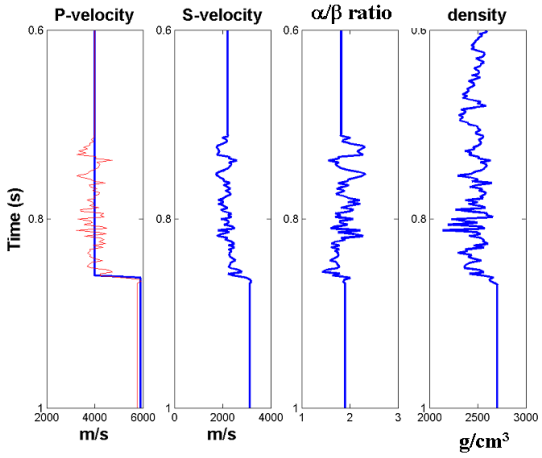

Figure 4.6: Well logs used to construct Blackfoot model. Blocky P-velocity (blue) used for ray tracing in AVO inversion.

4.3.4 Modeling results

To test the predictions made in this section a series of experiments were performed on the Blackfoot synthetic model. The experiments were focused on understanding how changing the background P-wave velocity used in the ray tracing would influence the estimates. The Walden approach to ray tracing was used with exact stacking velocities input. However, instead of the exact P-wave velocities a blocky velocity model was used to perform the ray tracing (Figure 4.6). In the first set of tests the blocky velocity model corresponded to the average velocity over specific geologic intervals. This tested how the AVO inversion responds to random errors in P-wave interval velocity. Secondly the blocky P-wave interval velocity model was systematically distorted by multiplying it by 0.9. This test was designed to show how the AVO inversion responds to systematic error in the P-wave interval velocity. No tests were performed on perturbing the stacking velocity. It is expected that the AVO inversion responds to errors in stacking velocity in a similar, but opposite fashion as errors in the P-wave interval velocity as governed by equation (4.30).

For each of these experiments the AVO inversion was performed first using the two-term Shuey inversion. The results are compared to the predictions in a qualitative fashion. Then the results are transformed to P-wave and S-wave impedance

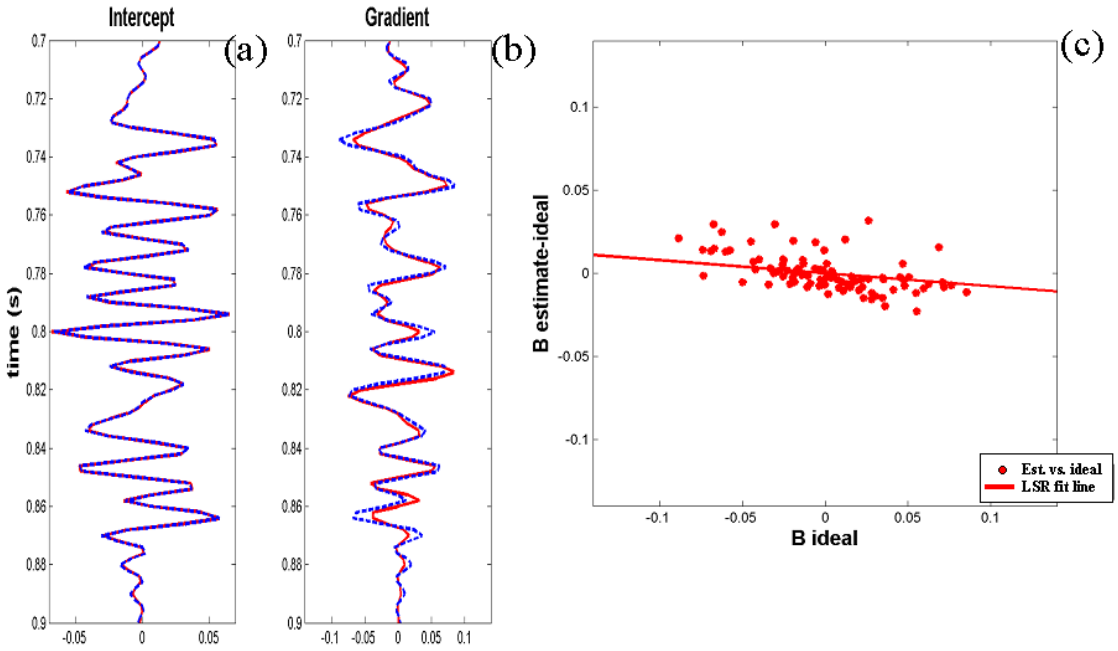

Figure 4.7: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Shuey inversion using blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated gradient reflectivity.

reflectivity. These results are then compared to the results of the two-term Gidlow inversion performed in a similar fashion. For each of these experiments the scalar and correlation coefficient are calculated for the least squares fit between the estimated reflectivity (B or R_s as the case may be) and the ideal and tabulated in Tables 4.2 and 4.3. The scalar is used to judge whether systematic error has been introduced and the correlation coefficient is used to judge the goodness of fit. In these series of tests the background S-wave velocity used in the AVO inversion or transform (equation 4.1) was modified so that $\bar{\gamma}$ remained unchanged from the original model, even though the P-wave velocity changed. This was done so that changes in $\bar{\gamma}$ would not influence the error as specified by equation (4.27). Changes in $\bar{\gamma}$ are tested in Section 4.4 and 4.5.

Figure 4.7 shows the results of performing the AVO inversion using the blocky P-wave interval velocity model. There is slightly more scatter than when the exact P-wave velocity model is used to do the ray tracing. The systematic error is slightly greater, as evidenced by the scalar. This might be due to the fact that a blocky model

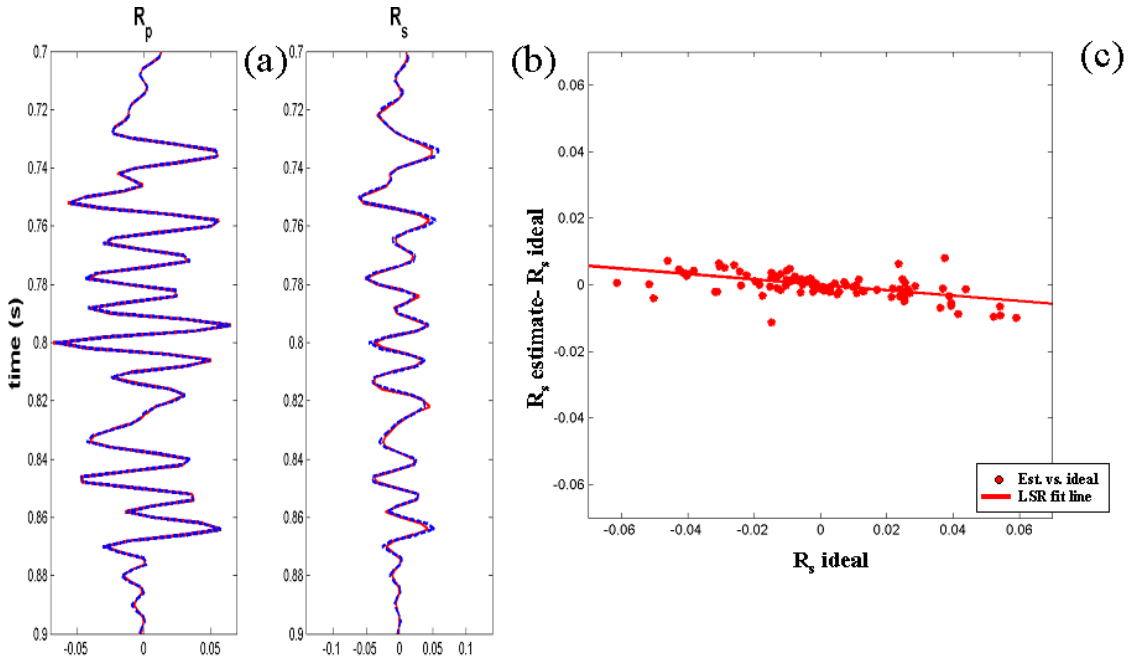

Figure 4.8: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Shuey inversion transformed to impedance reflectivity using blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated S-impedance reflectivity.

was used rather than a running average creating local systematic errors. Figure 4.8 shows the results of the Shuey inversion transformed to impedance reflectivity. The misfit of these results are similar to those of the two-term Gidlow inversion (Figure 4.9). This suggests that error predictions generated for the two-term Shuey inversion may be transferable to the two-term Gidlow inversion, where I was not able to derive an analytic expression to predict the error. It is preferable to use the two-term Gidlow inversion over the Shuey inversion because of the larger systematic error evident in the Shuey solution, as evidenced by their respective scalars. The scatter between the estimate and the ideal in each one of these estimates is not significantly greater than that arising due to using a two-term linear approximation of the Zoeppritz equations. The error seems to be no more and in fact less, than that predicted by equation (4.30). Generalizing these results, suggests a smooth average velocity model may be used to do the ray tracing without introducing significant error.

More problematic are systematic distortions in the background velocity model

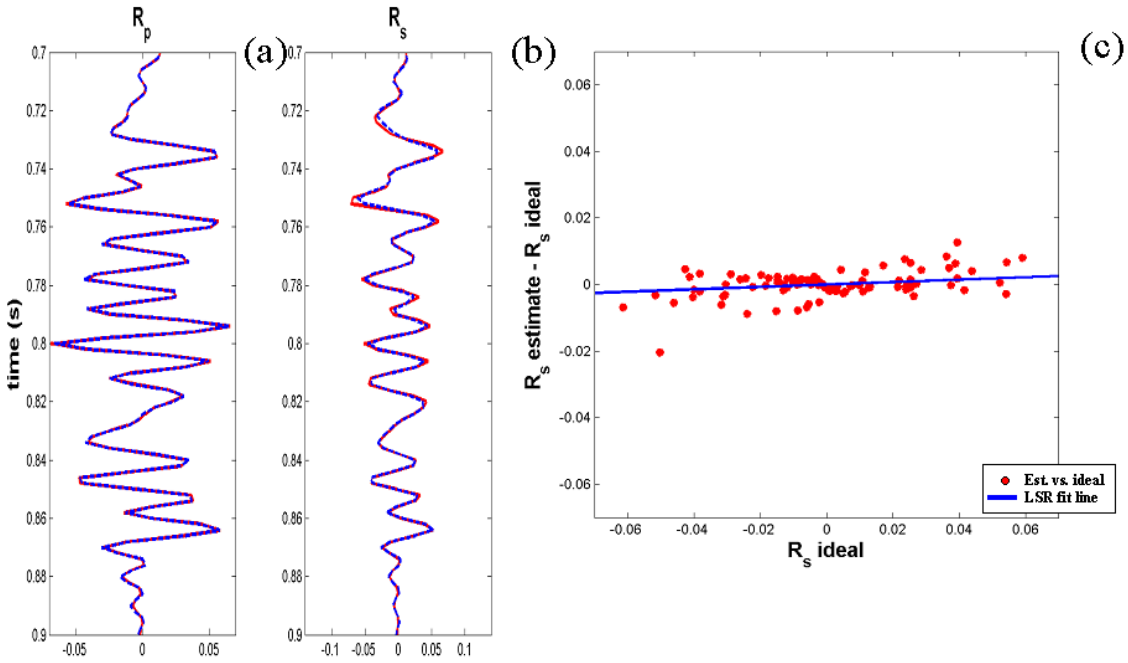

Figure 4.9: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Fatti inversion using blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated S-impedance reflectivity.

used for the ray tracing. Figure 4.10 shows the results of the two-term Shuey inversion performed with the background $\bar{\alpha}$ model systematically distorted. There is no discernible error in the intercept A but there is a systematic error in the gradient B. Table 4.2 shows the estimated gradient is now 1.1341 time greater than the ideal. For the blocky velocity model the scalar was 0.9218, a change of 0.2123. This is close

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9436	0.9300	1.0169
B	blocky background P-wave velocity	0.9218	0.9195	1.0355
C	0.9× blocky background P-wave velocity	1.1341	1.0331	1.1445
	difference between C&B	0.2123	0.1136	0.1090

Table 4.2: Systematic error (scalar) as a result of using a two-term linearized approximation and approximate P-wave velocity instead of the Zoeppritz equation.

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9530	0.9834	0.9846
B	blocky background P-wave velocity	0.9519	0.9821	0.9815
C	0.9× blocky background P-wave velocity	0.9496	0.9776	0.9815
	difference between C&B	−0.0023	−0.0045	0.0000

Table 4.3: Random error (correlation coefficient) as a result of using a two-term linearized approximation and approximate P-wave velocity instead of the Zoeppritz equation.

to what equation (4.15) predicts

$$\frac{\Delta B}{B_0} = -2\left(\frac{.9-1}{1}\right) = 0.2. \tag{4.31}$$

Note that equation (4.15) is a linear approximation. Higher order terms have be ignored. Figure 4.11 shows these results transformed to impedance reflectivity. This may be compared to the two-term Gidlow inversion (Figure 4.12). The scalar error between the estimated and ideal value is half that for the gradient. This is once again due to the transform of variables. It is interesting to note that in Table 4.3 the two-term Shuey inversion shows slightly greater scatter in this case while the two-term Gidlow inversion does not. The Shuey result is somewhat surprising as this is not predicted from the preceding analysis, though the effect is small.

4.4 β/α ratio error

In this section the impact of errors due to our limited *a priori* knowledge of $\bar{\gamma}(t)$ are considered. In this section errors associated with the P-wave velocity field are ignored. One of the advantages of the Shuey formulation is that no explicit knowledge of $\bar{\gamma}$ is required for the inversion. This means it is insensitive to these errors. Advocates of this equation point to this fact in championing its use. However, transforming the intercept and gradient to other parameterizations such as impedance requires knowledge of $\bar{\gamma}$. Equation (4.27) may be modified

$$dR_s = \frac{1}{4\bar{\gamma}^2}\left(B - \frac{A}{1+g}\right)\frac{d\bar{\gamma}}{\bar{\gamma}}, \tag{4.32}$$

to predict the uncertainty associated with this transformation.

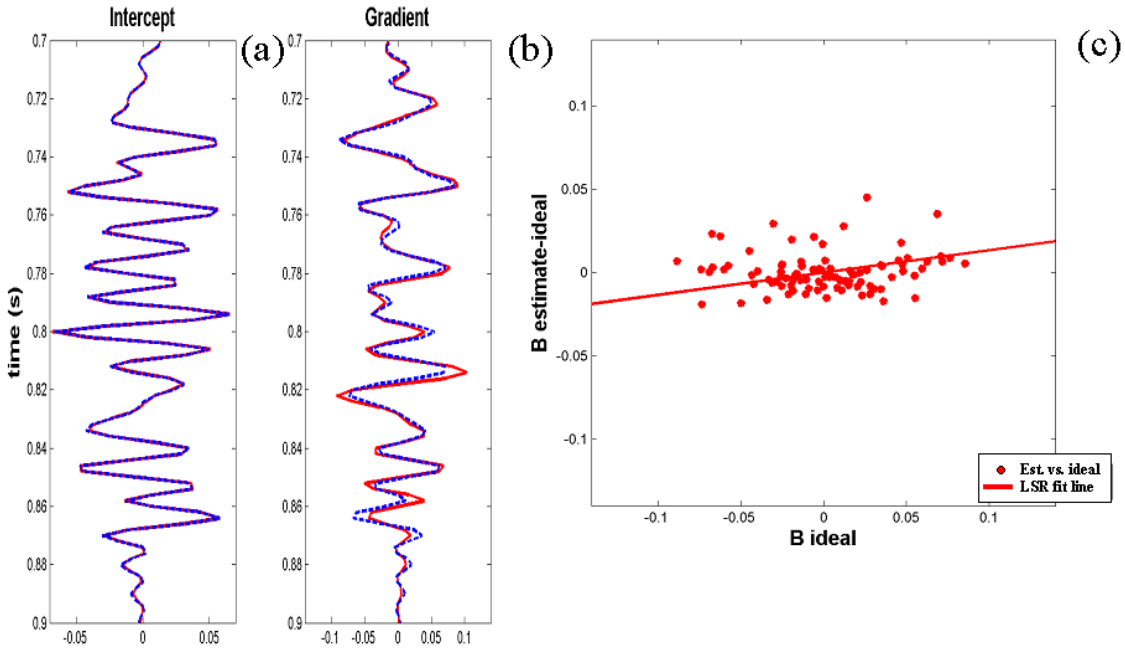

Figure 4.10: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Shuey inversion using distorted blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated gradient reflectivity.

Unlike the previous section it is relatively easy to understand how the Gidlow equation responds to errors in $\bar{\gamma}$. The least squares estimate for the Gidlow equation are (Appendix D.2)

$$R_p = \frac{\left(\sum_{m=1}^{M} \sin^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j}{\left(\sum_{j=1}^{M} \sec^4 \bar{\theta}_j\right) \sum_{m=1}^{M} \sin^4 \bar{\theta}_m - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right)^2}, \qquad (4.33)$$

and

$$R_s = \frac{1}{8\bar{\gamma}^2} \frac{\left(\sum_{m=1}^{M} \sec^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j}{\left(\sum_{j=1}^{M} \sec^4 \bar{\theta}_j\right) \sum_{m=1}^{M} \sin^4 \bar{\theta}_m - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right)^2}. \qquad (4.34)$$

Note that there is no dependence on the $\bar{\gamma}$ ratio in R_p so this parameter is insensitive to errors in $\bar{\gamma}$. The R_s reflectivity is scaled proportionally to the inverse of square of the $\bar{\gamma}$ ratio. If the wrong $\bar{\gamma}$ ratio is used in AVO inversion the R_s reflectivity will be

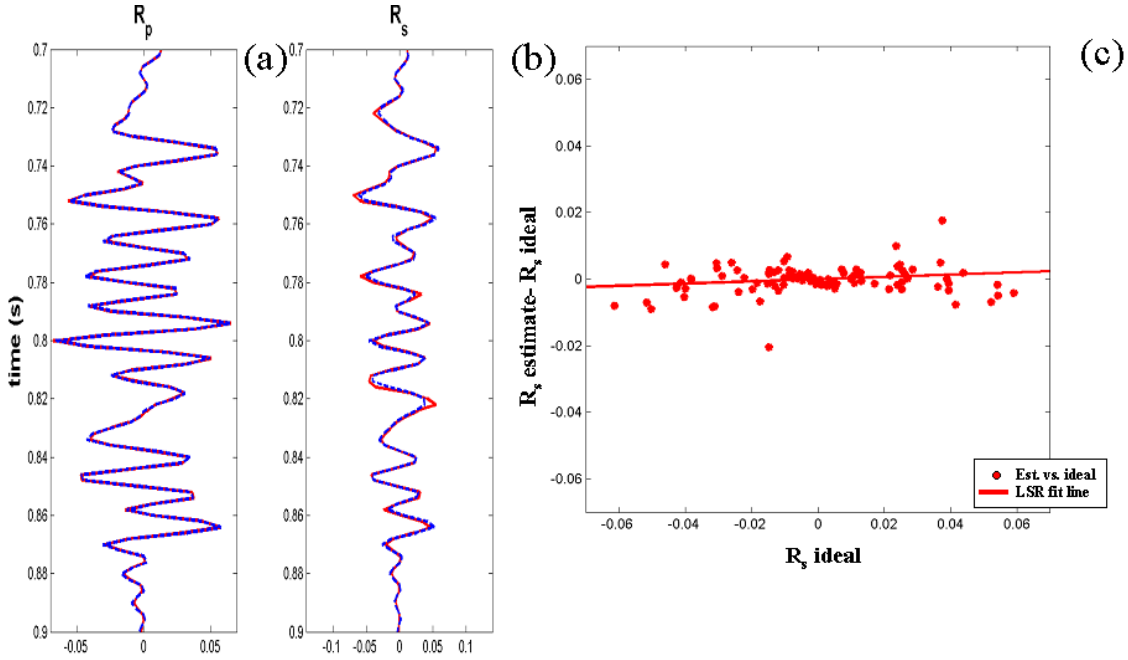

Figure 4.11: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Shuey inversion converted to impdedance using distorted blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated S-impedance reflectivity.

either too large or small. This provides one means of performing an error analysis. The S-wave impedance reflectivity obtained by performing an AVO inversion with the correct $\bar{\gamma}_{cor}$ is

$$R_s^{cor} = \frac{1}{8\bar{\gamma}_{cor}^2} \frac{\left(\sum_{m=1}^{M} \sec^4\bar{\theta}_m\right)\sum_{j=1}^{M} d_j \sin^2\bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2\bar{\theta}_m\right)\sum_{j=1}^{M} d_j \sec^2\bar{\theta}_j}{\left(\sum_{j=1}^{M} \sec^4\bar{\theta}_j\right)\sum_{m=1}^{M} \sin^4\bar{\theta}_m - \left(\sum_{m=1}^{M} \tan^2\bar{\theta}_m\right)^2}, \tag{4.35}$$

while the S-wave impedance reflectivity obtained by performing an AVO inversion with a biased $\bar{\gamma}_{biased}$ is

$$R_s^{biased} = \frac{1}{8\bar{\gamma}_{biased}^2} \frac{\left(\sum_{m=1}^{M} \sec^4\bar{\theta}_m\right)\sum_{j=1}^{M} d_j \sin^2\bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2\bar{\theta}_m\right)\sum_{j=1}^{M} d_j \sec^2\bar{\theta}_j}{\left(\sum_{j=1}^{M} \sec^4\bar{\theta}_j\right)\sum_{m=1}^{M} \sin^4\bar{\theta}_m - \left(\sum_{m=1}^{M} \tan^2\bar{\theta}_m\right)^2}. \tag{4.36}$$

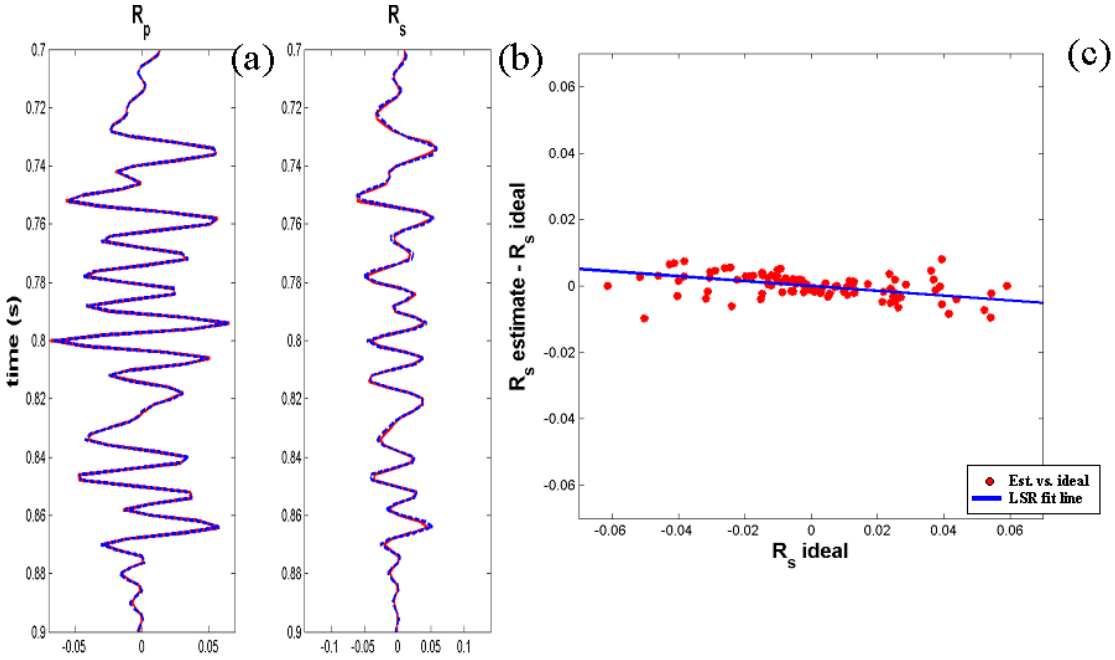

Figure 4.12: Comparison of ideal (red) and estimated (blue) reflectivities for two-term Fatti inversion using distorted blocky velocity model for ray tracing. The cross-plot (c) shows the difference between the ideal and estimated S-impedance reflectivity.

Thus, the biased estimate is a scaled version of the correct one

$$R_s^{biased} = \frac{\bar{\gamma}_{cor}^2}{\bar{\gamma}_{biased}^2} R_s^{cor}, \tag{4.37}$$

with the scalar proportional to $\bar{\gamma}_{cor}^2/\bar{\gamma}_{biased}^2$. This suggests that if the correct $\bar{\gamma}_{cor} = \frac{1}{4}$ and the biased gamma used in the AVO inversion is $\bar{\gamma}_{biased} = \frac{1}{2}$ then the estimated S-wave impedance reflectivity will be

$$R_s^{biased} = \frac{1}{4} R_s^{cor}. \tag{4.38}$$

This situation can easily happen for shallow unconsolidated sands. In the shallow low velocity section the mudrock relationship suggests that the $\bar{\gamma}$ should be small, with $\bar{\gamma} = \frac{1}{4}$ not being atypical. Often for convenience people use $\bar{\gamma} = \frac{1}{2}$ leading to errors similar to that predicted by equation (4.38).

Similar to the last section, the uncertainty in the S-wave impedance reflectivity

due to $\bar{\gamma}$ can also be estimated by expanding the solution as a Taylor series expansion about the average background S-wave to P-wave velocity ratio $\bar{\gamma}_{back}$.

$$\Delta R_s = \left.\frac{dR_s}{d\bar{\gamma}}\right|_{\bar{\gamma}=\bar{\gamma}_{back}} \Delta\gamma. \tag{4.39}$$

where $\Delta\gamma = \bar{\gamma} - \bar{\gamma}_{back}$. This leads to a linear approximation

$$\Delta R_s = -\frac{\Delta\gamma}{\bar{\gamma}_{back}} 2R_s, \tag{4.40}$$

which is not as accurate as equation (4.37), but illustrative none the less. For small perturbations equation (4.37) and (4.40) give approximately the same results. Further, equation (4.40) is approximately equal to equation (4.32) derived in a completely different fashion. To see this, note that for the simple case where the background $\bar{\gamma} = 1/2$, $R_s = \frac{A-B}{2}$ and equation (4.40) becomes

$$\Delta R_s = [B - A] \frac{\Delta\gamma}{\bar{\gamma}_{back}}. \tag{4.41}$$

This compares favorably with equation (4.32) which under these circumstances becomes

$$dR_s = \left(B - \frac{A}{1+g}\right) \frac{d\bar{\gamma}}{\bar{\gamma}}. \tag{4.42}$$

The differences may be explained by the different approximations used to derive these results. The derivation from the Gidlow equation should be the more accurate of the two. The derivation based on the Shuey equation makes use of the Gardner equation to replace the C term and also ignores the bias introduced into the B term in the two-term Shuey AVO inversion.

Returning to equation (4.40), the $\bar{\gamma}$ error may once again be broken into two parts, one due to random errors and the other due to systematic errors. The random errors come from the high frequency perturbations from the background running average and the systematic errors are a result of low frequency perturbations from the background running average $\bar{\gamma}$. Rearranging equation (4.40) results in

$$\frac{\Delta R_s}{R_s} = -2\frac{\Delta\gamma}{\bar{\gamma}_{back}}, \tag{4.43}$$

suggesting the fractional uncertainty in the S-wave impedance reflectivity due to $\bar{\gamma}$

uncertainty is proportional to the γ ratio reflectivity, R_γ. Further, R_γ may be written in terms of the P-wave and S-wave velocity reflectivity

$$R_\gamma = R_\beta - R_\alpha. \tag{4.44}$$

Since R_α and R_β are positively correlated, their difference is less than the individual parts. Thus the error should be again an order of magnitude less than the S-wave impedance reflectivity and its effect relatively minor.

Systematic errors occur if the $\bar{\gamma}$ ratio used for the AVO inversion is not really the average. For example, if the value used is 0.9 times the actual value then $\frac{\Delta\gamma}{\bar{\gamma}} = -0.1$ and by equation (4.43)

$$\Delta R_s = 0.2 R_{s_0}, \tag{4.45}$$

thus the estimated S-wave impedance reflectivity

$$\hat{R}_s = R_{s_0} + \Delta R_s = 1.2 R_s. \tag{4.46}$$

Alternatively, equation (4.37) may be used to calculate this error as well. In this case, $\bar{\gamma}_{biased} = 0.9\bar{\gamma}_{cor}$ implying $\bar{\gamma}^2_{biased} = 0.81\bar{\gamma}^2_{cor}$ so equation (4.37) becomes

$$R_s^{biased} = \frac{1}{0.81} R_s^{cor} = 1.234\,6 R_s^{cor}, \tag{4.47}$$

which is approximately equivalent to equation (4.46) within the error of the linear approximation.

4.4.1 Modeling results

Once again the Blackfoot synthetic model was used to test how changing $\bar{\gamma}_{back}$ used in the inversion would influence estimates. Similar to the preceding section the Walden approach to ray tracing was used with exact interval and stacking velocities so there would be no error due to these factors. The initial test was run using a constant $\bar{\gamma}_{back} = 0.545$ for all time samples. This was the average $\bar{\gamma}$ for the interval so the first test was designed to understand the influence of random errors from the background trend. In the second test a biased version of $\bar{\gamma}_{back} = 0.490\,5$ was used to test the effect of systematic errors. This is 0.9 times $\bar{\gamma}$.

In a similar manner to the preceding section, a two-term Shuey inversion was considered. The least squares scalar and correlation coefficient were calculated between

the estimated and ideal reflectivity. As expected, there was no variation in any of the tests for the gradient B as it is not a function of $\bar{\gamma}$. The Shuey estimates were then transformed to S-wave impedance reflectivity and then compared to the estimates from the two-term Gidlow inversion. This was done for both $\bar{\gamma} = 0.545$ and $\bar{\gamma}_{biased} = 0.490\,5$. The least squares scalars and correlation coefficients are tabulated in Tables 4.4 and 4.5. The correlation coefficient and the scalar virtually did not change between the reference (row A in Table 4.4 and 4.5) and the random error test (row D in Table 4.4 and 4.5). Interestingly, the correlation coefficient for the S-wave impedance reflectivity predicted by the Shuey inversion increased, indicating there is less error than the reference. This is probably due to the empirical nature of equation (4.1). Perturbing the background velocity in a systematic fashion results in a systematic distortion of the parameters as predicted by equation (4.47).

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9436	0.9300	1.0169
D	$\bar{\gamma} = 0.545$ (average gamma)	0.9436	0.9337	1.0281
E	$\bar{\gamma} = 0.490\,5$	0.9436	1.1243	1.2692
	difference between E&D	0.0000	0.1906	0.2411

Table 4.4: Systematic error (scalar) as a result of using a two-term linearized approximation and approximate S-wave to P-wave velocity ratio instead of the Zoeppritz equation.

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9530	0.9834	0.9846
D	$\bar{\gamma} = 0.545$ (average gamma)	0.9530	0.9836	0.9838
E	$\bar{\gamma} = 0.490\,5$	0.9530	0.9834	0.9838
	difference between E&D	0.0000	-0.0002	0.0000

Table 4.5: Random error (correlation coefficient) as a result of using a two-term linearized approximation and approximate S-wave to P-wave velocity ratio instead of the Zoeppritz equation.

4.5 Combined effect of modeling errors

In the previous two sections it was shown that the error due to high frequency perturbations from the running average for both $\bar{\alpha}$ and $\bar{\gamma}$ leads to random Gaussian

error. The two effects have been examined individually but not together. This needs to be done since the two effects are correlated through the mudrock relationship.

The Blackfoot synthetic was inverted using the two-term Shuey and Gidlow equations using average background P-wave velocities and $\bar{\gamma}$. These results are compared to inversions where only one of the variables had random error. The least squares scalar and correlation coefficient were calculated between the estimated and ideal reflectivity and tabulated in Tables 4.6 and 4.7. The correlation coefficient of the gradient is the same as the test when only the P-wave velocity is averaged as input into the AVO inversion. This is expected since the $\bar{\gamma}$ error does not influence the Shuey inversion. However, upon converting the estimates to S-wave impedance reflectivity, the correlation coefficient increases, indicating the error decreases. This suggests that the two errors are negatively correlated. The S-wave impedance estimate correlation coefficient from the Gidlow inversion increases as well. Interestingly the correlation coefficient decreases from the case when only $\bar{\gamma}$ is smoothed (row D) indicating that there is greater error when there is random error in $\bar{\gamma}$ and $\bar{\alpha}$ rather than only $\bar{\gamma}$. This effect is opposite to when there is only random error in $\bar{\alpha}$. The S-wave impedance has a higher correlation when there are random errors in $\bar{\gamma}$ (row D) than when there are random errors in $\bar{\alpha}$ (row B). To summarize the results of this example, the errors due to random errors in $\bar{\alpha}$ seem to be greater than those due to random errors in $\bar{\gamma}$. Also, the errors due to random errors in $\bar{\gamma}$ and $\bar{\alpha}$ seem to be negatively correlated so that the total error is less than the individual parts.

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9436	0.9300	1.0169
B	blocky background P-wave velocity	0.9218	0.9195	1.0355
C	0.9× blocky background P-wave velocity	1.1341	1.0331	1.1445
D	$\bar{\gamma} = 0.545$ (average gamma)	0.9436	0.9337	1.0281
E	$\bar{\gamma} = 0.490\,5$	0.9436	1.1243	1.2692
F	blocky α and $\bar{\gamma} = 0.545$ (average gamma)	0.9218	0.9230	1.0198
G	blocky $0.9 \times \alpha$ and 1.053×gamma	1.1341	0.9474	1.1075

Table 4.6: Systematic error (scalar) as a result of using a two-term linearized approximation, approximate P-wave velocity, and S-wave to P-wave velocity ratio instead of the Zoeppritz equation.

The last experiment introduced systematic distortions in $\bar{\alpha}$ and $\bar{\gamma}$. These results are tabulated in row G in Tables 4.6 and 4.7 The experiment was designed so that the two systematic distortions would cancel out and that the correct estimate of S-wave

		Shuey B	Shuey R_s	Gidlow R_s
A	two-term approximation	0.9530	0.9834	0.9846
B	blocky background P-wave velocity	0.9519	0.9821	0.9815
C	$0.9\times$ blocky background P-wave velocity	0.9496	0.9776	0.9815
D	$\bar{\gamma} = 0.545$ (average gamma)	0.9530	0.9836	0.9838
E	$\bar{\gamma} = 0.490\,5$	0.9530	0.9834	0.9838
F	blocky α and $\bar{\gamma} = 0.545$ (average gamma)	0.9519	0.9837	0.9834
G	blocky $0.9 \times \alpha$ and $\bar{\gamma} = 0.490\,5$	0.9496	0.9821	0.9840

Table 4.7: Random error (correlation coefficient) as a result of using a two-term linearized approximation, approximate P-wave velocity, and S-wave to P-wave velocity ratio instead of the Zoeppritz equation.

impedance reflectivity would be obtained.

4.6 Discussion

For lack of knowledge we supply a smooth $\bar{\gamma}$ a priori model to the AVO inversion. This leads to both random and systematic error. From a theoretical and experimental point of view the random error seems not to be significant. The systematic error is more problematic as it leads to biased estimates of the S-wave impedance reflectivity.

4.6.1 Crossplotting

Systematic error potentially might be identified by crossplotting the secondary reflectivity attribute with the primary reflectivity attribute. For brine filled clastics the mudrock relationship

$$R_p = m_I \bar{\gamma} R_s, \tag{4.48}$$

should be observed. Rearranging the mudrock relationship (section 1.4.1)

$$\bar{\gamma} = \frac{1}{m_I}\left(1 - \frac{b}{\bar{\alpha}}\right), \tag{4.49}$$

results in the alternate relationship

$$R_p = \left(1 - \frac{b}{\bar{\alpha}}\right) R_s. \tag{4.50}$$

If the P- and S- wave impedance reflectivity are crossplotted, then the slope observed should be $\left(1-\frac{b}{\bar{\alpha}}\right)$. If the observed slope does not match the predicted slope this might indicate there is an issue with systematic error due to bias in the background $\bar{\alpha}$ or $\bar{\gamma}$. Further complicating this analysis is the fact systematic error can also be present in the data due to incorrect scaling or residual NMO. Swan (2001) corrects for residual normal moveout by examining the slope in the crossplot domain. The systematic modeling errors identified in this chapter will have a negative impact on this analysis and procedure.

The crossplot analysis may also be performed in the intercept-gradient domain by transforming equation (4.48) using equation (4.1) and noting $R_p = A$. This results in

$$B = \left(\left(\frac{g}{1+g}\left(4\bar{\gamma}^2-1\right)+1\right)-\frac{8\bar{\gamma}}{m_I}\right)A. \tag{4.51}$$

The apparent advantage of having no $\bar{\gamma}$ dependence using Shuey estimates is mitigated by the fact both $\bar{\gamma}$ and m_I appear in the slope of equation (4.51).

4.6.2 Bias in the two-term Shuey inversion estimates

Throughout this chapter, the two-term Shuey inversion performed on the Blackfoot synthetic example has produced biased estimates of the gradient B. The third term becomes more significant as the range of angles used increases, though it is still significant at relatively small angle ranges such as 20°. Since the AVO inversion problem becomes more ill-conditioned as the range of angles used decreases, it is not possible to select a range of angles that produces no bias and yet gives stable estimates in the presence of noise. One possible solution to address this bias is to invert the three-term Shuey inversion, but retaining only the first two terms. However, uncertainty analysis indicates this greatly increases the uncertainty in the gradient over the two term solution. This again is due to the ill-conditioned nature of the three-term AVO inversion.

Another approach is to change the constraint used to reduce the problem to two terms. Instead of truncating the third term, which is effectively defining the P-wave velocity reflectivity to be zero, the more geologically plausible Gardner constraint (2.16) is used. This leads to the relationship

$$\frac{A}{(1+g)} = C, \tag{4.52}$$

which upon substitution into the three-term Shuey equation (2.4)

$$R(\bar{\theta}) = A + B\sin^2\bar{\theta} + C\sin^2\bar{\theta}\tan^2\bar{\theta}, \tag{4.53}$$

leads to

$$R(\bar{\theta}) = A\left(1 + \frac{1}{(1+g)}\sin^2\bar{\theta}\tan^2\bar{\theta}\right) + B\sin^2\bar{\theta}. \tag{4.54}$$

The resulting equation (4.54) behaves in the same fashion as the Smith and Gidlow equation though parameterized differently. On the negative side, under this new parameterization A and B no longer have the physical significance of slope and intercept. The parameters are more abstract and the equation loses some of its appeal as a result. Thus, I believe if bias is a concern and one insists on using a two-term inversion, one is better off using the two-term Gidlow equation which is naturally parameterized in terms of the physically significant P-wave and S-wave impedance reflectivity.

4.6.3 Quantifying data and theoretical uncertainty

In Chapter 2, the parameter uncertainty was estimated by performing a linear transform on the data uncertainty which was quantified by the data covariance matrix. Tarantola (1987) introduces random theoretical or modeling error by incorporating it into a modified data covariance matrix

$$\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_T, \tag{4.55}$$

where $\mathbf{C}_T$ is covariance matrix describing the uncertainty due to modeling error. This modified covariance matrix may be used to analyze the uncertainty in fashion similar to Chapter 2. For typical noise levels the uncertainty due to the data will be dominant over the uncertainty due to theoretical error. This is evidenced by Figure 2.17. In that figure, only when the signal-to-noise ratio is greater than 16:1 does the theoretical error become dominant. This is evidenced by the normalized error increasing as the range of angles used in the inversion increases. In this case, the theoretical error is a result of using a two-term approximation rather than the exact Zoeppritz equations.

4.7 Conclusions

In this chapter the uncertainty due to theoretical error arising from the linear approximation of the Zoeppritz equations was studied. Through a synthetic modeling experiment it was shown that this approximation introduces what appears to be random error into the reflectivity estimates, the size of which is a function of how well the linear approximation is met.

Further, in order to linearize the problem it is assumed that $\bar{\gamma}$ and the offset to angle-of-incidence mapping is known *a priori*. This is never the case and leads to random and systematic errors. The errors due to the angle-of-incidence to offset mapping were studied, following the Walden approach to ray tracing. In this approach, deviations in $\bar{\alpha}$ from the ideal average velocity $\bar{\alpha}$ introduce random errors in the gradient and S-wave impedance reflectivity. These random errors are an order of magnitude less than estimates themselves and in practice are not significant when compared to the uncertainty due to the random noise. The implication of this is that the exact P-wave velocity is not required to do the ray tracing, only a smooth background model which follows the average $\bar{\alpha}$. Systematic errors in the estimate of the gradient and S-wave impedance reflectivity arise if the smooth background model departs from the average $\bar{\alpha}$. These manifest themselves as a time variant scalar that appears to be applied to the gradient or the S-wave impedance reflectivity.

As well, uncertainty in $\bar{\gamma}$ leads to both random and systematic error in the S-wave impedance reflectivity. Since the two-term Shuey inversion is not a function of $\bar{\gamma}$, it is not influenced by it. However, if the estimates of the Shuey inversion are transformed to impedance reflectivity, the transformed results behave in a similar fashion to uncertainties in $\bar{\gamma}$ as the Gidlow results. Once again, if deviations in $\bar{\gamma}$ exist from the ideal average, this introduces random error into the S-wave impedance reflectivity, which is an order of magnitude less than the estimate itself. In practice, only a smooth background $\bar{\gamma}$ model is needed for the inversion, provided it follows the average $\bar{\gamma}$. If the background model departs from the average, systematic error is introduced.

Thus far, the conclusions about errors due to $\bar{\alpha}$ and $\bar{\gamma}$ are both supported by analytic expressions and experimental results from a synthetic modeling study. When combining the random errors due to $\bar{\alpha}$ and $\bar{\gamma}$ they appear to be negatively correlated, thus potentially reducing their additive effect. This conclusion is only empirical being based on the modeling study. I was unable to come up with an analytic

expression to support this. The systematic errors due to these sources may reinforce one another or negatively combine so as to eliminate one another. The effect of the systematic error might be seen in the crossplot space. Isolating these errors due to these sources is difficult since systematic errors in the data manifest themselves in a similar fashion.

In a pragmatic sense, systematic errors manifest themselves as time variant scalars on the gradient and S-wave impedance reflectivity. For situations where these errors are important, such as in creating the fluid stack, a post-AVO scalar may be applied to the affected attribute. This scalar may be based on information from local well control or statistical relationships.

Lastly, the random errors due to $\bar{\alpha}$ and $\bar{\gamma}$ are smaller than those due to the linear approximation made in simplifying the Zoeppritz equations. If one wanted to reduce the theoretical error in the problem, one might examine using nonlinear approximations of the Zoeppritz equations. However, in this chapter we have only looked at theoretical approximations associated with the linearization of the Zoeppritz equations. Much larger errors exist that have been ignored, such as the influence of multiples, the laterally heterogenous nature of the inverse problem, NMO stretch and offset-dependent tuning. It is the last two issues that are examined next.

Chapter 5

NMO stretch and tuning artifacts

5.1 Introduction

In Chapter 3 it was shown that to successfully estimate three independent reflectivity attributes using AVO inversion large offsets and angles are needed. Large offsets are also desirable in areas with strong random or coherent noise. The large angles improve the condition number of the inverse problem. However, including these large offsets brings on other problems. NMO stretch and offset dependent tuning both become problematic. This chapter explores the nature and magnitude of these distortions and their inter-relationships. In subsequent chapters these distortions are addressed either through modified ways of pre-conditioning the data prior to AVO inversion or incorporating the corrections into the AVO inversion itself.

NMO stretch is described by Dunkin and Levin (1973) and its effect on AVO inversion is summarized in the first section of this Chapter. The removal of NMO stretch is complicated by offset dependent tuning (Lin and Phair, 1993). The second section describes offset dependent tuning and its inter-relationship to NMO stretch. Dong, in a series of papers (Dong, 1996; Dong, 1998; Dong, 1999), quantified these errors and suggested a correction. This chapter builds on his results, rearranging his equations to arrive at expressions for fractional error for the intercept A and gradient B. The predictions from these relationships are then tested and verified on several synthetic seismic data sets.

The intercept is not influenced by these distortions, only the gradient is effected. Both the analytic expressions and the modeling studies show that the different classes of gas sands (Rutherford and Williams, 1989) react to NMO stretch and offset dependent tuning differently. Class I and II anomalies hardly affect the gradient at

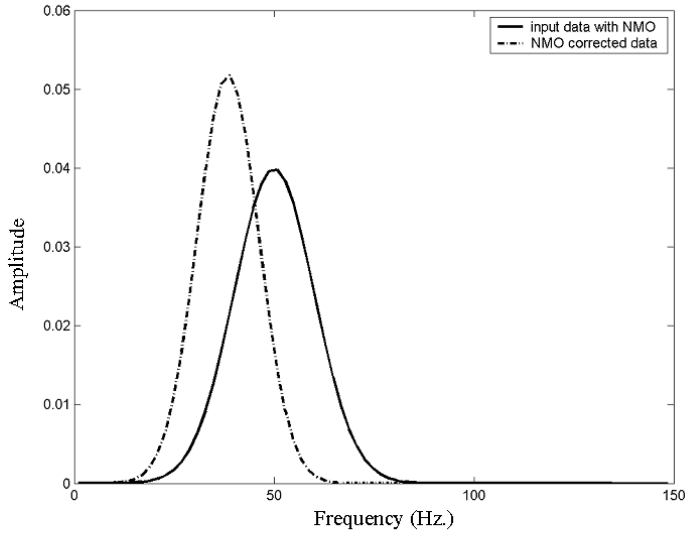

Figure 5.1: Amplitude spectrum before and after NMO correction (equation 5.1). The NMO correction (NMO stretch) shifts the spectrum to lower frequencies and amplifies the values.

all. However, the gradient is significantly distorted for Class III and IV gas sands. Lastly, these distortions are only significant when large angles are being used in the AVO inversion (such as 45°). If the maximum angle used in the AVO inversion is less than 30° these distortions are not significant.

5.2 NMO stretch artifacts

NMO correction is a kinematic operation. It is more concerned with the positioning rather than the amplitudes of the events. Claerbout (1992) points out that NMO correction is a conjugate operation, not an inverse operation. As such it introduces amplitude and character distortions. NMO stretch is one of the most familiar forms of this. For two isolated reflectors, Dunkin and Levin (1973) describe NMO stretch analytically with the expression

$$\tilde{S}_x(f) = \frac{1}{\alpha_x} S_x\left(\frac{f}{\alpha_x}\right), \tag{5.1}$$

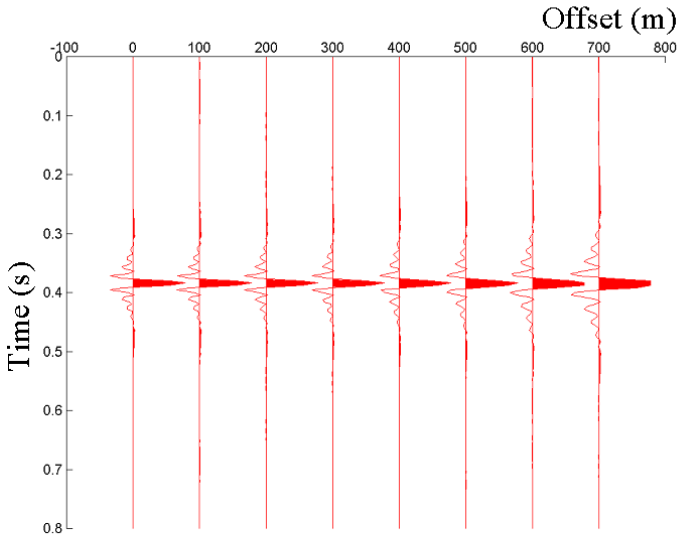

Figure 5.2: Synthetic gather of a single spike after NMO correction and 10/14-60/70 Hz band-pass filter for incident angles from 0 to 45 degrees. Note how NMO stretch lowers the frequency on the far offsets and changes the wavelet character.

where S_x is the spectrum before NMO correction, $\tilde{S}_x$ is the spectrum after NMO correction, f is frequency and α_x is the compression factor or the ratio of the time difference between the two events after and before NMO. The compression factor is always less than one, so the frequency spectrum will be shifted to lower frequencies and amplified as shown in Figure 5.1.

The compression factor, α_x, becomes smaller for larger offsets thus the shape of the wavelet changes in an offset dependent fashion. For example, Figure 5.2 shows a gather after NMO correction for incident angles from $0\,^\circ$ to $45\,^\circ$. The model generating this is a single reflector or spike that is convolved with a 5/10-60/70 Hz band-pass filter. For this to match the assumptions of the traditional methodology, the reflector after NMO correction must have constant waveform and amplitude. It does not. The band-width of the far offset data are noticeably lower frequency than the near offset data and the overall character changes as a function of offset.

This biases the AVO inversion and introduces error. This can be understood intuitively, for this example, by calculating the intercept and gradient. The intercept of the zero crossing at 0.39 seconds is zero. The gradient at this same time is positive since the wavelet broadens as a function of offset due to NMO stretch. However, if

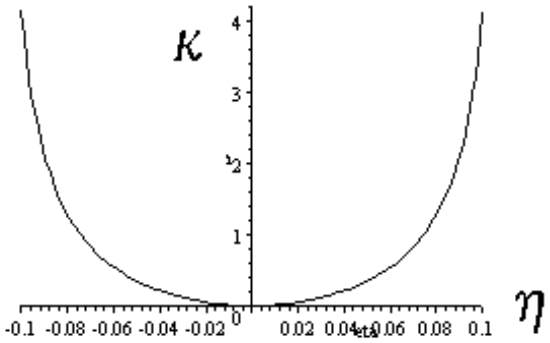

Figure 5.3: Scaling κ distortion on the gradient estimate as a function of η. Note the distortion gets larger as η gets larger.

there was no NMO stretch both the intercept and gradient would be zero.

For a Ricker wavelet, Dong (1996) equation (9), approximates the biases to the AVO parameter estimates intercept A' and gradient term B' due to NMO stretch as

$$A' = e^{-4\pi^2\eta^2}\left(1 - 8\pi^2\eta^2\right)A, \tag{5.2}$$

and

$$B' = e^{-4\pi^2\eta^2}\left[\left(1 - 8\pi^2\eta^2\right)B + 4\pi^2\eta^2\left(3 - 8\pi^2\eta^2\right)A\right], \tag{5.3}$$

where A and B are the true intercept and gradient. The parameter

$$\eta = f_0 dt, \tag{5.4}$$

is defined in terms of the dominant frequency f_0 and the time dt of how far the time sample under investigation is from center of the wavelet. The Ricker wavelet written in terms of η is

$$w(\eta) = \left(1 - 8\pi^2\eta^2\right)e^{-4\pi^2\eta^2}, \tag{5.5}$$

which upon comparison with equation (5.2) implies that there is no distortion introduced to the intercept other than the wavelet itself. The biased gradient can be written in terms of two parts $B' = B + dB$, that due to the wavelet and that due to an error term. Dividing the biased gradient (equation 5.3) by the Ricker wavelet

(5.5)

$$B' = B + \frac{4\pi^2\eta^2\left(3 - 8\pi^2\eta^2\right)}{\left(1 - 8\pi^2\eta^2\right)}A,$$

the error term is

$$dB = \frac{4\pi^2\eta^2\left(3 - 8\pi^2\eta^2\right)}{\left(1 - 8\pi^2\eta^2\right)}A,$$

or the fractional error is

$$\frac{dB}{B} = \kappa\frac{A}{B}, \tag{5.6}$$

where

$$\kappa = \frac{4\pi^2\eta^2\left(3 - 8\pi^2\eta^2\right)}{\left(1 - 8\pi^2\eta^2\right)}. \tag{5.7}$$

Thus, the fractional error of the gradient is a function of $\kappa\left(\eta\right)$ and the ratio of the intercept to the gradient is A/B. If the analysis is performed on the center of the wavelet, then $\eta = 0$ and $dB/B = 0$. Figure 5.3 shows that as η increases, κ increases. Thus, the size of the gradient error increases. The other factor which controls the size of the error is the ratio A/B. Thus, it is possible to predict the size of the error for different classes of AVO anomalies (Section 1.4.2). For Class I anomalies where $A << B$ and Class II anomalies where $A \approx 0$ the expected error will be small. This is also true for most reflectors which fall along the mudrock trend, since typically, $A << B$. However, for Class III and IV anomalies where $A > B$ the error is potentially large.

To test these predictions, a synthetic seismic data model was constructed. Four isolated reflectors were generated corresponding to the four classes outlined above. The synthetic data was generated using a convolutional model with a Ricker wavelet with a 32.5 Hz dominant frequency. Preliminary testing suggested that large offsets and angles were needed to make the NMO stretch artifacts apparent. To avoid theoretical error being introduced due to these large offsets the Aki and Richards (1980) linearized approximation of the Zoeppritz equations was used to generate the reflectivity. The third parameter and term was defined using the Gardner density approximation $R_d = 0.25R_\alpha$. Further, to keep the relationship between offset and angle of incidence simple, and to avoid supercritical reflections, a constant background velocity was used to generate the model. With the maximum offset about four times the target depth, angles out to 65° were generated.

Figure 5.4 shows both the Shuey (1985) two term and three term response for reflectors generated with no moveout. Note the far offset reflectivity behavior is

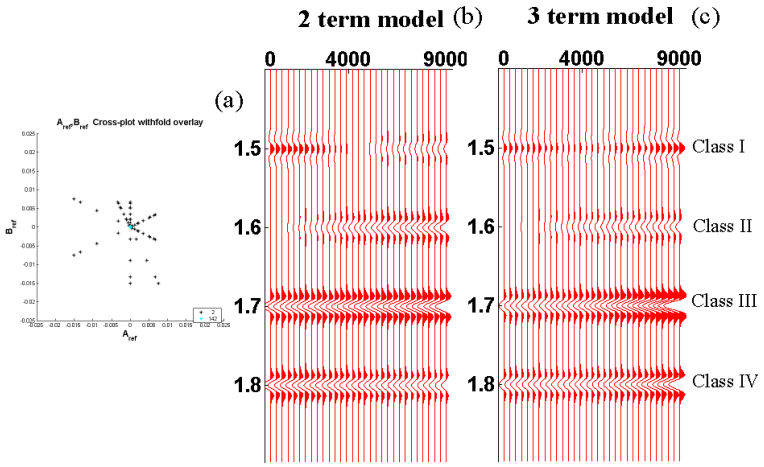

Figure 5.4: Cross-plot of ideal reflectivity (a) used to generate the synthetic gather generated without NMO using two term Shuey approximation (b) and the synthetic gather generated without NMO using three term Shuey approximation (c). Note the two-term model clearly shows the Class I - IV behavior expected while the three term model behavior is more complex.

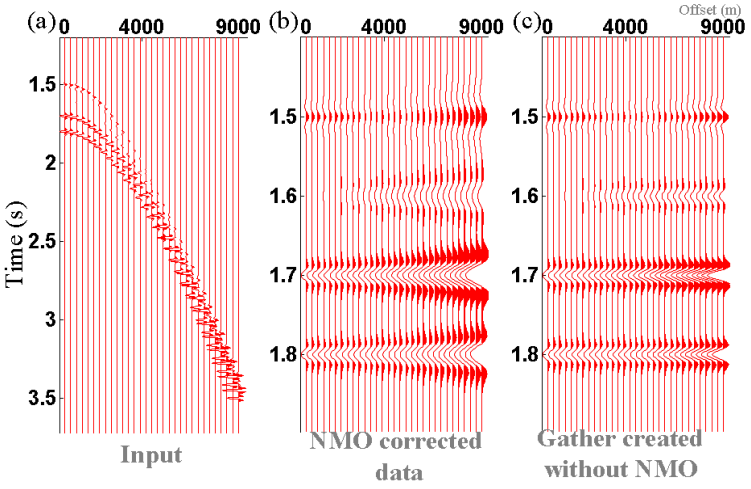

Figure 5.5: The input model prior to NMO (a), after NMO correction (b) and compared to the synthetic gather generated without NMO (c). Note on the NMO corrected gather the introduction of low frequencies at large offsets due to NMO stretch.

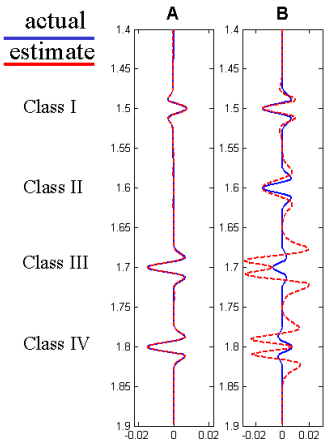

Figure 5.6: The estimate (red) of the AVO intercept A and gradient B compared to the ideal (blue). Note the gradient estimate is distorted for both the Class III and IV anomalies as predicted.

dramatically different due to the inclusion of the 3rd term. Figure 5.5 shows the synthetic gather before and after NMO correction. For comparison, the gather generated without NMO (Figure 5.5c) is shown next to the NMO corrected gather (Figure 5.5b). On the far offsets of the NMO corrected gather, it is possible to see the character change and frequency shift of the wavelet. This character change with offset due to NMO stretch will bias the AVO inversion.

The AVO inversion was performed using angles up to 45°. The Smith and Gidlow AVO inversion was done to avoid theoretical error due to the inclusion of the large angles used in this model and inversion. The parameters were then transformed for display purposes to intercept and gradient. These are shown in Figure 5.6 compared to the ideal intercept and gradient reflectivity. The estimated intercept is almost a perfect match to the ideal. The estimate of the gradient is close to the ideal for both the Class I and II anomalies. For the Class III and IV anomalies the estimate of the gradient shows a large error for $\eta > 0$. For $\eta = 0$ the gradient error is zero as expected. These results are consistent with the predictions based on equation (5.6). One factor that equation (5.6) ignores is the aperture or range of angles used for the AVO inversion. Originally, the AVO inversion was done only to 30°. The

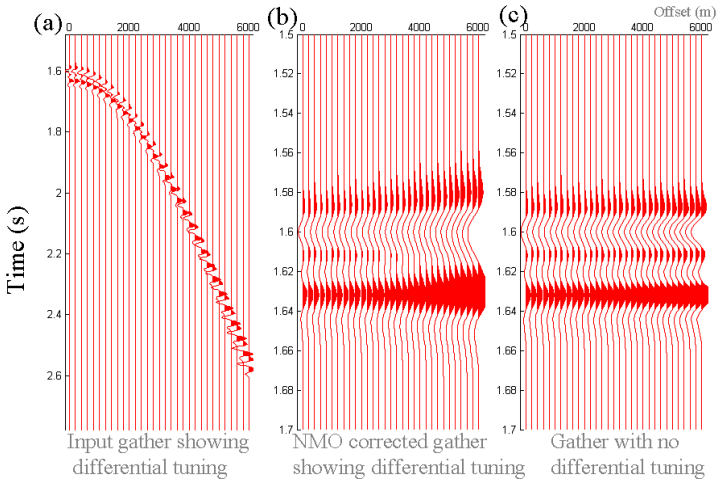

Figure 5.7: The input model prior to NMO with tuning (a), with NMO (b) and compared to the synthetic gather generated without NMO (c).

size of the gradient error was much smaller for that angle range suggesting that NMO stretch error is only significant if one wants to do AVO inversions including large angle ranges.

5.3 Offset dependent tuning

The fact NMO stretch introduces this error raises the question: "Why don't we use equation (5.1) to remove the effects of NMO stretch?" However as equation (5.1) is an approximation, the better question is: "Why don't we avoid introducing NMO stretch by using the exact NMO inverse operator (rather than the conjugate operator)?" Unfortunately, offset-dependent tuning introduces notches into the frequency spectrum or unrecoverable information so that the inversion in both cases is underdetermined and ill-conditioned. Conceptually, this can be understood by noting that the reflectivity of the "far-offset trace" is time delayed and squeezed into a smaller time window relative to that of the "zero-offset trace". If both the zero-offset and the far-offset reflectivity data are high-cut filtered, the far-offset data contains less information after filtering than the near-offset data. It is the combination of the high-cut filter and the NMO operator that introduces the null space.

In the time domain this effect manifests itself as offset dependent tuning. Figure 5.7 illustrates offset dependent tuning for a model using the same acquisition geometry

and wavelet as the previous example. At zero offset, the reflectivity is described by a dipole whose separation is one wavelength apart. Because of the moveout, the events tune differently as a function of offset. This tuning again biases the AVO estimates.

For a Ricker wavelet, Dong (1999) equation (18), approximates the AVO parameter estimates intercept, A',and gradient term, B',in the presence of NMO stretch and offset dependent tuning as

$$A' = e^{-\eta^2} 2\xi\eta \left(2\eta^2 - 3\right) A, \tag{5.8}$$

and

$$B' = e^{-\eta^2} 2\xi\eta \left[\left(2\eta^2 - 3\right) B + \left(2\eta^2 - 1\right)\left(\eta^2 - 3\right) A\right]. \tag{5.9}$$

The parameter ξ is defined by

$$\xi = \frac{2\pi\Delta z}{\lambda_0}, \tag{5.10}$$

where Δz is the distance between the two interfaces and λ_0 is the dominant wavelength. The tuned zero offset wavelet response is

$$w'(t) = e^{-\eta^2} 2\xi\eta \left(2\eta^2 - 3\right). \tag{5.11}$$

Upon comparing equation (5.11) with equation (5.8) it is evident that there is no distortion to the intercept. The biased gradient can be written in terms of two parts $B' = B + dB$, that due to the wavelet and that due to an error term. Dividing the biased gradient (equation 5.9) by the Ricker wavelet (5.11)

$$B' = B + \frac{\left(2\eta^2 - 1\right)\left(\eta^2 - 3\right)}{\left(2\eta^2 - 3\right)} A,$$

so the fractional error is

$$\frac{dB}{B} = \zeta \frac{A}{B}, \tag{5.12}$$

where

$$\zeta = \frac{\left(2\eta^2 - 1\right)\left(\eta^2 - 3\right)}{\left(2\eta^2 - 3\right)}. \tag{5.13}$$

It has the same basic functional form as that for NMO stretch with the exception that scalar ζ has a different functional relationship with respect to η. There was zero error for NMO stretch for $\eta = 0$. Because of the tuning, this behavior occurs at

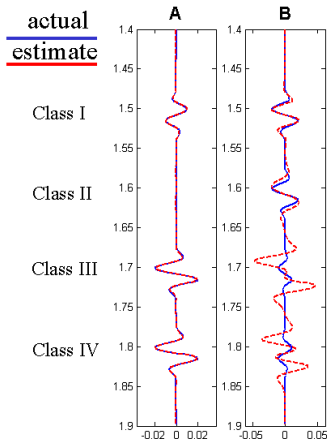

Figure 5.8: The estimate (red) of the AVO intercept A and gradient B compared to the ideal (blue). Note the gradient estimate is distorted for both the Class III and IV anomalies as predicted.

$\eta = \sqrt[\pm]{3}$ and $\eta = \sqrt[\pm]{\frac{1}{2}}$. Once again we might expect Class III and IV anomalies to show large errors while Class I, II and regional reflectivity show little error. To test this prediction, the previous model was modified so that instead of single reflectors at the zero offset, dipole reflectors are modeled. Figure 5.8 shows the AVO inversion results when the dipole was 1/2 the dominant wavelength of the source wavelet. As expected the estimate of intercept is almost a perfect match to the ideal. The estimate of the gradient shows significant error particularly for the Class III and IV anomalies. The error is several times larger than the gradient itself. To obtain a rough understanding of how the tuning layer thickness influences the error the model was run for a series of different tuning layer thicknesses. Figure 5.9 demonstrates that the error changes as a function of layer thickness.

Other reflectivity attributes behave differently to the error. Figure 5.10 shows how the S-wave impedance and fluid stack reflectivity behave to the distortion. Interestingly the fluid stack shows little distortion other than a phase delay. It is quite a robust AVO attribute even in the presence of NMO stretch and tuning.

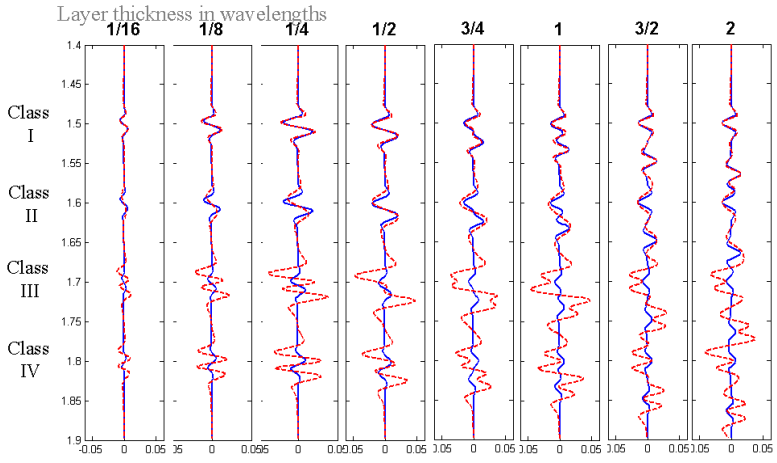

Figure 5.9: The estimate (red) of the AVO gradient B compared to the ideal (blue) for various layer thicknesses. Note that distortion changes as function of thickness.

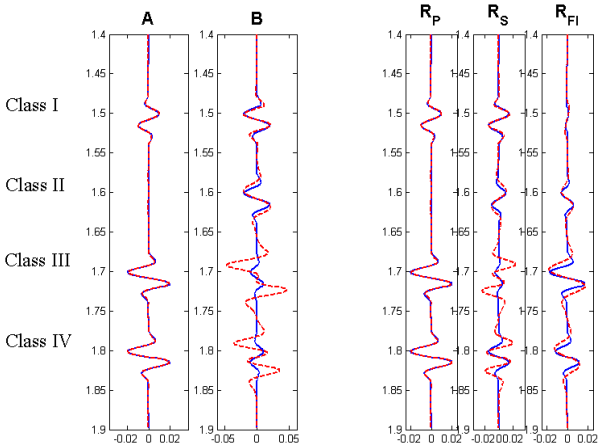

Figure 5.10: Intercept A and Gradient B converted to fluid stack, P-wave and S-wave impedance reflectivity. Note that the fluid stack shows little distortion due to NMO stretch and offset dependent tuning for all classes.

5.4 Discussion

The analytic expressions for fractional error derived in this chapter do not show a dependence on aperture. This is due to the fact that the starting point of their derivation was equations from Dong (1996), (1999), which are empirical relations based on least square fitting in which aperture was not considered. It would be interesting to rederive these equations incorporating aperture effects to better understand the influence of aperture.

5.5 Conclusion

This chapter has demonstrated that NMO stretch and offset dependent tuning potentially introduce large distortions into the AVO reflectivity estimates. The two distortions are interrelated since the offset dependent tuning introduces a null space into the NMO inversion problem. It is difficult to remove the NMO stretch without addressing this null space.

Having said this, there are surprisingly few situations for which these distortions introduce error into the AVO reflectivity estimates. The error due to these distortions primarily shows up on the secondary AVO reflectivity attribute, such as the gradient or S-wave impedance reflectivity. The attribute associated with the first term, the intercept or P-wave impedance reflectivity, is unaffected. Further, there is a surprisingly large group of geologic interfaces for which NMO stretch and offset dependent tuning do not distort the reflectivity estimates. Regional reflectors that arise due to interfaces between clastics following the mudrock trend are not distorted. Class I and II gas sands are not distorted. Only, reflectivity estimates for Class III and IV gas sand anomalies are distorted. Further, these distortions are only significant when large angles are used. For the synthetic data sets shown in this chapter, the AVO inversions, when done with angles less than 30° reflectivity estimates had insignificant error. Only when larger angles were used (for example, greater than 45°) were the errors significant.

The fact that different classes have different sensitivities to NMO stretch is somewhat counterintuitive. The synthetic gathers shown in Figure 5.5 show that all the classes experience NMO stretch at the far offsets. One might expect that all of them should be showing distortions to the gradient. If a two-term Shuey AVO inversion was done, this probably would be the case. However, the two-term Smith and Gid-

low inversion used in this chapter is effectively a three-term inversion, because of the Gardner constraint. This makes for a more complex fitting than just intercept and gradient. I believe this fitting is more appropriate since ultimately we are interested in using these large offsets to do three-term AVO inversion. Further, Class I and II far offset reflection strength is less than that of the Class III and IV anomalies. Using the viewpoint that least squares AVO inversion may be viewed as a weighted stack (Smith and Gidlow, 1987), the resulting distortion in the gradient due to the Class I and II anomalies is less because of these weaker far offset amplitudes.

Chapter 6

Two-term AVO waveform inversion

6.1 Introduction

In Chapter 5 it was shown that NMO stretch and offset dependent tuning can introduce distortions into the AVO inversion estimates, particularly for Class III and IV gas sands. This chapter develops an AVO waveform inversion that addresses these issues. It is necessary to incorporate the waveform and suitable constraints to address the band-limited nature of the seismic data and the resultant tuning issues. The AVO waveform inversion is performed on gathers that still have moveout. NMO corrections are never applied, thus avoiding the associated distortions. This chapter develops the AVO waveform inversion based on a two-term linear approximation of the Zoeppritz equations. Later in Chapter 8, this is extended to three terms. The method produces sparse spike reflectivity series similar to that done by poststack sparse deconvolution. (Levy and Fullagar, 1981; Sacchi, 1999; and Trad, 2002). Provided with suitable additional constraints, the output of the AVO waveform inversion may be transformed to P-wave and S-wave impedances, in a manner similar to Oldenburg et al. (1983).

Simmons and Backus (1996) also developed an AVO waveform inversion that incorporates traveltime information, so it also had the potential to address NMO stretch and offset dependent tuning distortions. However, they addressed the underdetermined nature of the problem by parameterizing the model with a fixed number of equally spaced layers, for example, one every 8 milliseconds. This parameterization is ill-suited to deal with the offset dependent tuning distortions. Buland and Omre (2003) also perform AVO waveform inversion using a linearized approximation, but on NMO corrected gathers, ignoring the distortions due to NMO stretch and offset dependent tuning. Both of these methods assume Gaussian priors.

The theory for this new algorithm is developed using a Bayesian formalism. Instead of the typical assumption of a Gaussian probability distribution, various long-tailed distributions are used for the prior distribution. The implementation of these constraints is analogous to the high-resolution Radon transform presented by Sacchi and Ulrych (1995). The likelihood model is based on the AVO NMO formalism (Downton and Lines, 2002) where the AVO and NMO inversions are performed simultaneously assuming some input wavelet. The inversion is nonlinear and must be solved using a bootstrap procedure. The conjugate gradient method is used to solve the inverse problem. Typically only a few iterations are needed to solve the problem so the algorithm is relatively fast.

The algorithm is first demonstrated on a simple synthetic created to illustrate how the new algorithm better estimates AVO reflectivity on events undergoing NMO stretch and tuning. This example is then used to demonstrate that the AVO waveform inversion is relatively insensitive to the wavelet used for the inversion. Subsequent to this, the algorithm is demonstrated on a more realistic synthetic gather based on well data. Several real data examples are shown, including one from the same area as the well data. Lastly, it is shown how to transform the reflectivity data to impedances on synthetic data.

Compared to traditional AVO performed on a sample-by-sample basis on NMO corrected gathers, the new approach is better able to estimate reflectors undergoing NMO stretch and differential tuning. In addition, the resulting sparse spike reflectivity is better able to resolve thin layers. The approach is more reliable than the estimates provided by the traditional AVO analysis performed on a sample-by-sample basis on NMO corrected gathers. This greater reliability is due to the classic trade-off between resolution and stability. With the new method, a few sparse reflectivity values are estimated with greater certainty than the dense reflectivity at every time sample as in the traditional AVO analysis.

6.2 Theory

6.2.1 Convolutional model

The convolutional model is used as the basis for the likelihood model. This model assumes the earth is composed of a series of flat, homogeneous, isotropic layers. Transmission losses, converted waves, and multiples are not incorporated in this model and

so must be addressed through prior processing. In theory, gain corrections such as spherical divergence, absorption, directivity, and array corrections can be incorporated into this model, but are not considered in this chapter for brevity and simplicity. Consequently they must be applied prior to the inversion. Any linear approximation of the Zoeppritz equations may be used as the starting point for this derivation. This chapter uses a modified version of the two-term Gidlow et al. (1992).

In order to make all the variables stationary, the scaled S-wave impedance reflectivity $\tilde{R}_s = \gamma R_s$ is solved for rather than S-wave impedance reflectivity (Section 3.2.4). Equation (2.20) parameterized in terms of the P-wave impedance reflectivity R_p and scaled S-wave impedance reflectivity $\tilde{R}_s$ is

$$y(\bar{\theta}) = \begin{bmatrix} \sec^2\bar{\theta} & -8\bar{\gamma}\sin^2\bar{\theta} \end{bmatrix} \begin{bmatrix} R_p \\ \tilde{R}_s \end{bmatrix}, \tag{6.1}$$

where $y(\bar{\theta})$ is the offset dependent reflectivity, $\bar{\theta}$ is the average angle of incidence and, $\bar{\gamma}$ the average S-wave to P-wave velocity across the interface. If it is assumed that $\bar{\gamma}$ and the relationship between the offset and angle of incidence is known, then equation (6.1) is linear. As an example, for the case of two offsets, a near-offset y_1 and a far-offset y_2, written in matrix form is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} e_1 & f_1 \\ e_2 & f_2 \end{bmatrix} \begin{bmatrix} R_p \\ \tilde{R}_s \end{bmatrix}, \tag{6.2}$$

where $e_j = \left[1 + \tan^2\bar{\theta}_j\right]$, $f_j = 8\bar{\gamma}\sin^2\bar{\theta}_j$ and the subscript indicates the offset.

Typically equation (6.2) is solved on an interface-by-interface basis, where each interface corresponds to a time sample. This ignores the band-limited nature of the seismic data. To address this, equation (6.2) can be modified to solve for multiple time samples simultaneously. To illustrate this, consider the case with two interfaces: equation (6.2) becomes

$$\begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \\ y_1^{(2)} \\ y_2^{(2)} \end{bmatrix} = \begin{bmatrix} e_1^{(1)} & f_1^{(1)} & 0 & 0 \\ e_2^{(1)} & f_2^{(1)} & 0 & 0 \\ 0 & 0 & e_1^{(2)} & f_1^{(2)} \\ 0 & 0 & e_2^{(2)} & f_2^{(2)} \end{bmatrix} \begin{bmatrix} R_p^{(1)} \\ \tilde{R}_s^{(1)} \\ R_p^{(2)} \\ \tilde{R}_s^{(2)} \end{bmatrix}, \tag{6.3}$$

where the superscript has been introduced to indicate the interface number. Rear-

ranging equation (6.3) so that the vectors are ordered along common offsets rather than common time samples results in

$$\begin{bmatrix} y_1^{(1)} \\ y_1^{(2)} \\ y_2^{(1)} \\ y_2^{(2)} \end{bmatrix} = \begin{bmatrix} e_1^{(1)} & 0 & f_1^{(1)} & 0 \\ 0 & e_1^{(2)} & 0 & f_1^{(2)} \\ e_2^{(1)} & 0 & f_2^{(1)} & 0 \\ 0 & e_2^{(2)} & 0 & f_2^{(2)} \end{bmatrix} \begin{bmatrix} R_p^{(1)} \\ R_p^{(2)} \\ \tilde{R}_s^{(1)} \\ \tilde{R}_s^{(2)} \end{bmatrix}, \tag{6.4}$$

which is further simplified by writing equation (6.4) as the partitioned matrix

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{E}_1 & \mathbf{F}_1 \\ \mathbf{E}_2 & \mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}. \tag{6.5}$$

In equation (6.5) the vector $\mathbf{y}_k$ is the data for offset k, $\mathbf{r}_p$ the P-wave impedance reflectivity vector, $\tilde{\mathbf{r}}_s$ the scaled S-wave impedance reflectivity vector, and the block matrices $\mathbf{E}_k$ and $\mathbf{F}_k$ are diagonal matrices.

AVO NMO model

NMO can be written as a linear operator (Claerbout, 1992). The reflectivity sequence $\mathbf{y}_k$ referenced to zero-offset time can be transformed to offset-dependent traveltime $\mathbf{d}_k$ for the k^{th} offset by the linear operator $\mathbf{N}_k$ so that

$$\mathbf{d}_k = \mathbf{N}_k \mathbf{y}_k. \tag{6.6}$$

The matrix $\mathbf{N}_k$ can be constructed using whatever offset-traveltime relationship one desires. In order to invert data at large angles of incidence, it is important to correctly position the event without introducing residual NMO. In this thesis, I use the higher order correction (Section 1.2.2) following Castle (1994). This has the advantage of introducing high-order terms without introducing the theoretical complications of intrinsic anisotropy. Implicit in this derivation is that the NMO velocity is known *a priori* and that static corrections have been applied.

Combining equations (6.5) and (6.6) results in the set of linear equations that may be used to solve NMO and AVO simultaneously.

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1\mathbf{E}_1 & \mathbf{N}_1\mathbf{F}_1 \\ \mathbf{N}_2\mathbf{E}_2 & \mathbf{N}_2\mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}. \tag{6.7}$$

To adequately describe the band-limited nature of the data and the offset dependent tuning the wavelet must be included as part of the formulation. This may be done by pre-multiplying equation (6.7) by the convolution matrix $\mathbf{W}$ to obtain

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{W}\mathbf{N}_1\mathbf{E}_1 & \mathbf{W}\mathbf{N}_1\mathbf{F}_1 \\ \mathbf{W}\mathbf{N}_2\mathbf{E}_2 & \mathbf{W}\mathbf{N}_2\mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}. \tag{6.8}$$

This is the key equation describing the physics of the problem. This formulation can be modified to include more offsets by just adding additional rows. By incorporating NMO into the inverse problem, the NMO correction is never done as a processing step thus avoiding introducing NMO stretch distortions and biases. Differential tuning is incorporated as part of the forward model and so is also dealt with.

Data mutes and masks Equation (6.1) is a truncated polynomial of a linear approximation of the Zoeppritz equations. It is only valid for subcritical angles. Thus, the elements of the $\mathbf{E}_k$ and $\mathbf{F}_k$ operators must be zeroed for supercritical angles. This is accomplished with the masking operator $\mathbf{M}_k^f$, where the superscript f indicates it is designed on flat data (data without NMO). In addition, zeros may have been introduced into the seismic data through the data processing, for example first break muting. Further, because of noise and theoretical considerations one might only want to only perform the AVO inversion over some predefined angle range, from $\bar{\theta}_{\min}$ to $\bar{\theta}_{\max}$. Both these operations may be accounted for by the masking operator $\mathbf{M}_k^n$, where the superscript n indicates it gets applied to data with NMO. In practice, because of the band-limited nature of the seismic data, I applied the $\mathbf{M}_k^n$ operator with a taper resulting in

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1^n\mathbf{W}\mathbf{N}_1\mathbf{M}_1^f\mathbf{E}_1 & \mathbf{M}_1^n\mathbf{W}\mathbf{N}_1\mathbf{M}_1^f\mathbf{F}_1 \\ \mathbf{M}_2^n\mathbf{W}\mathbf{N}_2\mathbf{M}_2^f\mathbf{E}_2 & \mathbf{M}_2^n\mathbf{W}\mathbf{N}_2\mathbf{M}_2^f\mathbf{F}_2 \end{bmatrix} \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}. \tag{6.9}$$

Note the data must be suitably muted as well.

Linear operator For future reference and simplicity, this linear model (equation 6.9) is written as

$$\mathbf{d} = \mathbf{G}\mathbf{x}, \tag{6.10}$$

where $\mathbf{G}$ is the linear AVO waveform operator

$$\mathbf{G} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 \\ \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K \end{bmatrix}, \tag{6.11}$$

$\mathbf{d}$ is the data vector

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \end{bmatrix}, \tag{6.12}$$

and $\mathbf{x}$ is the unknown parameter vector describing the reflectivity

$$\mathbf{x} = \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}. \tag{6.13}$$

The data vector contains L time samples by K offsets resulting in a total of $M = L \times K$ data. There is a total of $N = 2L$ unknown parameters in equation (6.13)

Transform of variables In the next section it is necessary to work with independent variables. However, Chapter 3 demonstrated that the P-wave and S-wave velocity reflectivity are statistically correlated. Likewise the P-wave and S-wave impedance are statistically correlated. In a similar fashion to Chapter 3 a covariance matrix may be constructed that describes the statistical relationship between the P-wave and the scaled S-wave impedance reflectivity (Appendix F, equation F.8)

$$\mathbf{C}_{\tilde{\mathbf{x}}} = \begin{bmatrix} \sigma_{R_p}^2 & \sigma_{R_p \tilde{R}_s} \\ \sigma_{R_p \tilde{R}_s} & \sigma_{\tilde{R}_s}^2 \end{bmatrix} = \sigma_{R_p}^2 \begin{bmatrix} 1 & \frac{r_{R_p \tilde{R}_s}^2}{m_I} \\ \frac{r_{R_p \tilde{R}_s}^2}{m_I} & \frac{r_{R_p \tilde{R}_s}^2}{m_I^2} \end{bmatrix}, \tag{6.14}$$

where $\sigma_{R_p}^2$ is the variance of the P-wave impedance reflectivity, m_I is the mudrock slope, and $r_{R_p \tilde{R}_s}$ is the correlation coefficient between R_p and $\tilde{R}_s$. If the two variables to be solved are independent, then the off-diagonal terms of the correlation matrix should be zero. This may be accomplished by performing a change of variables. This change of variables is facilitated by determining the eigenvectors $\mathbf{V}$ and eigenvalues $\mathbf{\Sigma}$ of the covariance matrix

$$\mathbf{C}_{\mathbf{x}} = \mathbf{V} \mathbf{\Sigma} \mathbf{V}^T. \tag{6.15}$$

The eigenvectors of this simple 2×2 symmetric matrix are

$$\mathbf{V} = \left[\begin{array}{cc} \mathbf{v}_1 & \mathbf{v}_2 \end{array} \right], \tag{6.16}$$

where

$$\mathbf{v}_1 = \left[\begin{array}{c} \frac{m_I^2 - r_{R_p\tilde{R}_s}^2 + \sqrt{r_{R_p\tilde{R}_s}^4 - 2m_I^2 r_{R_p\tilde{R}_s}^2 + m_I^4 + 4m_I^2 r_{R_p\tilde{R}_s}^4}}{2m_I r_{R_p\tilde{R}_s}^2} \\ 1 \end{array} \right], \tag{6.17}$$

and

$$\mathbf{v}_2 = \left[\begin{array}{c} \frac{m_I^2 - r_{R_p\tilde{R}_s}^2 - \sqrt{r_{R_p\tilde{R}_s}^4 - 2m_I^2 r_{R_p\tilde{R}_s}^2 + m_I^4 + 4m_I^2 r_{R_p\tilde{R}_s}^4}}{2m_I r_{R_p\tilde{R}_s}^2} \\ 1 \end{array} \right], \tag{6.18}$$

while the eigenvalues are

$$\mathbf{\Sigma} = \left[\begin{array}{cc} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{array} \right] = \sigma_{r_p}^2 \left[\begin{array}{cc} \tilde{\sigma}_1^2 & 0 \\ 0 & \tilde{\sigma}_2^2 \end{array} \right], \tag{6.19}$$

where

$$\tilde{\sigma}_1^2 = \frac{r_{R_p\tilde{R}_s}^2 + m_I^2 + \sqrt{r_{R_p\tilde{R}_s}^4 - 2m_I^2 r_{R_p\tilde{R}_s}^2 + m_I^4 + 4m_I^2 r_{R_p\tilde{R}_s}^4}}{2m_I^2}, \tag{6.20}$$

and

$$\tilde{\sigma}_2^2 = \frac{r_{R_p\tilde{R}_s}^2 + m_I^2 - \sqrt{r_{R_p\tilde{R}_s}^4 - 2m_I^2 r_{R_p\tilde{R}_s}^2 + m_I^4 + 4m_I^2 r_{R_p\tilde{R}_s}^4}}{2m_I^2}. \tag{6.21}$$

Using the eigenvectors as the transform matrix

$$\mathbf{x} = \mathbf{V}\mathbf{x}', \tag{6.22}$$

the transformed covariance matrix becomes (Appendix I)

$$\mathbf{C}_{\mathbf{x}'} = \mathbf{\Sigma}, \tag{6.23}$$

which is the desired result in that all the off-diagonal elements are now zero. Recall, that the diagonal of the covariance corresponds to the variance of each of the variables so the first eigenvalue corresponds to the variance of variable 1 and the second

eigenvalue corresponds to the variance of variable 2, thus for a single interface

$$\mathbf{C}_{\mathbf{x}'} = \mathbf{\Sigma} = \sigma^2_{r_p} \begin{bmatrix} \tilde{\sigma}^2_1 & 0 \\ 0 & \tilde{\sigma}^2_2 \end{bmatrix}. \tag{6.24}$$

The covariance matrix shown in equation (6.24) is for a single interface. If it is assumed, as is typically done in deconvolution, that the reflectivity is stationary and independent then it is straightforward to extend the covariance matrix (6.24) to L time samples resulting in the $N \times N$ sparse covariance matrix $\mathbf{C}_{\mathbf{x}}$. The eigenvector analysis can also be easily extended to the N variable case.

This transformation is similar to doing a principal component analysis on the well control. The major difference is that the covariance matrix is constructed from parameters rather than the actual well log reflectivity statistics These parameters, however, are determined from the statistics of the well log reflectivity series. The transformed variables can be thought of as an average of the P-wave and S-wave impedance reflectivities and the fluid stack. This parameterization is similar to Simmons and Backus (1996).

Transforming the linear operator (equation 6.9) using the transform matrix (6.22) results in

$$\mathbf{d} = \mathbf{G}'\mathbf{x}' \tag{6.25}$$

where $\mathbf{G}'$ is the transformed linear AVO NMO operator

$$\mathbf{G}' = \begin{bmatrix} \mathbf{M}^n_1\mathbf{W}\mathbf{N}_1\mathbf{M}^f_1\mathbf{E}_1 & \mathbf{M}^n_1\mathbf{W}\mathbf{N}_1\mathbf{M}^f_1\mathbf{F}_1 \\ \vdots & \vdots \\ \mathbf{M}^n_K\mathbf{W}\mathbf{N}_K\mathbf{M}^f_K\mathbf{E}_K & \mathbf{M}^n_K\mathbf{W}\mathbf{N}_K\mathbf{M}^f_K\mathbf{F}_K \end{bmatrix} \mathbf{V}. \tag{6.26}$$

6.2.2 Prior model

Both the matrices $\mathbf{W}$ and $\mathbf{N}$ are typically underdetermined or ill-conditioned. This is due to the fact the data are band-limited and the differential tuning introduces null spaces into the NMO operator $\mathbf{N}$. Because of this, the problem needs to be regularized. This can be done by choosing a weighting function that treats certain reflection coefficients as being more reliable than others. Choosing a long tailed *a priori* distribution leads to such a weighting function.

A long tailed distribution or sparse reflectivity may be justified on the basis of

experimental evidence (Walden and Hosken, 1986) and physical arguments. The fluid factor reflectivity is sparse by its nature since it only responds to anomalous fluids or large changes in lithology. The P-wave impedance reflectivity may also be modeled as a long tailed distribution, such as the exponential distribution (Levy and Fullagar, 1981; Shapiro and Hubral, 1999). Thus, the two reflectivity series can be modeled by a variety of distributions including the Huber, Cauchy or ℓ^p norm. In this section, these three distributions are introduced. In addition, the gradient of these distributions are calculated for future use in the optimization problem. These derivations follow Sacchi (1999) with modifications made to account for the fact that although the variables are independent, they have different variances.

ℓ^p Norm

The ℓ^p norm metric J_p is defined as

$$J_p = \frac{1}{p} \sum_i |x_i'|^p \,, \tag{6.27}$$

where p is defined as the norm. Both the ℓ^1 and ℓ^2 norms are members of this. The ℓ^1 norm comes from the exponential distribution while the ℓ^2 norm comes from the Gaussian distribution. The derivative of this function with respect to the unknown variable x' is required in the Bayesian formulation. For $p = 1$ the derivative is discontinuous so the limit as $p \longrightarrow 1$ must be used to approximate the ℓ^1 norm. There are several advantages in using the ℓ^p norm, namely that the norm can be adjusted to better fit the statistics of the known geology and that large problems can be efficiently solved using methods such as conjugate gradient.

In order that no one large variable dominates the calculation of the norm, the definition (equation 6.27) is modified

$$J_p = \frac{1}{p} \sum_i \left| \frac{x_i'}{\sigma_{x_i}} \right|^p \,, \tag{6.28}$$

so that each variable is normalized by its standard deviation σ_{x_i}. Taking the gradient of J_p results in (equation G.5)

$$\frac{\partial J_p}{\partial x_n} = \frac{1}{\sigma_{x_n}^2} \left| \frac{x_n'}{\sigma_{x_n}} \right|^{p-2} x_n'. \tag{6.29}$$

This simplifies somewhat since each variable is stationary. Upon substitution of $\sigma_1^2 = \sigma_{r_p}^2 \tilde{\sigma}_1^2$ and $\sigma_2^2 = \sigma_{r_p}^2 \tilde{\sigma}_2^2$ equation (6.29) becomes

$$\frac{\partial J_p}{\partial x_n} = \frac{x_n'}{\sigma_{r_p}^2} \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right|^{p-2} & n \leq L \\ \frac{1}{\sigma_2^2} \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right|^{p-2} & n > L \end{cases}, \tag{6.30}$$

where L is the number of time samples in the analysis window. For $p < 2$ a singularity exists if $x_n = 0$. To avoid this, a lower threshold ϵ is implemented

$$\frac{\partial J_p}{\partial x_n} = \frac{x_n'}{\sigma_{r_p}^2} \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \begin{cases} \epsilon^{p-2} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right| < \epsilon \\ \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right|^{p-2} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right| > \epsilon \end{cases} & n \leq L \\ \frac{1}{\tilde{\sigma}_2^2} \begin{cases} \epsilon^{p-2} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right| < \epsilon \\ \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right|^{p-2} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right| > \epsilon \end{cases} & n > L \end{cases}, \tag{6.31}$$

where $\epsilon > 0$ and typically $\epsilon << 1$. Specifically for $p = 1$

$$\frac{\partial J_1}{\partial x_n} = \frac{x_n'}{\sigma_{r_p}^2} \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \begin{cases} \frac{1}{\epsilon} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right| < \epsilon \\ \frac{1}{\left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right|} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_1} \right| > \epsilon \end{cases} & n \leq L \\ \frac{1}{\tilde{\sigma}_2^2} \begin{cases} \frac{1}{\epsilon} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right| < \epsilon \\ \frac{1}{\left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right|} & \left| \frac{x_n'}{\sigma_{r_p}\tilde{\sigma}_2} \right| > \epsilon \end{cases} & n > L \end{cases}. \tag{6.32}$$

Equation (6.32) may be written in matrix form

$$\frac{\partial J_1}{\partial x_n} = \frac{1}{\sigma_{r_p}^2} \mathbf{Q}_1 \mathbf{x}', \tag{6.33}$$

where the diagonal elements of the matrix $\mathbf{Q}_1$ are defined as

$$Q_{nn} = \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \begin{cases} \frac{1}{\epsilon} & \left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| < \epsilon \\ \frac{1}{\left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_1}\right|} & \left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| > \epsilon \end{cases} & n \leq L \\ \frac{1}{\tilde{\sigma}_2^2} \begin{cases} \frac{1}{\epsilon} & \left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| < \epsilon \\ \frac{1}{\left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_2}\right|} & \left|\frac{x'_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| > \epsilon \end{cases} & n > L \end{cases} . \tag{6.34}$$

The hyperparameter ϵ must be set prior to the inversion, while the parameters $\tilde{\sigma}_1^2$ and $\tilde{\sigma}_2^2$ may be calculated using equations (6.20) and (6.21). The parameter ϵ controls the spareness of the solution. Good results were obtained by setting ϵ quite small, for example $\epsilon = 0.001$. This results in reflectivities lower than ϵ being weighted by a large weights $(1/\epsilon)$, while large reflectivities are weighted by small weights.

Cauchy distribution

The Cauchy distribution is another long tailed distribution which leads to a sparse reflectivity series. Appendix G.2 describes the physical basis for using such a distribution and the derivation of the following results. Similar to the preceding derivation the prior distribution results in a influence function of the form

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\lambda_1}{\sigma_{r_p}^2}\mathbf{Q}_c\mathbf{x}' \tag{6.35}$$

with a diagonal matrix $\mathbf{Q}_c$ whose elements are defined by

$$Q_{nn} = \begin{cases} \frac{1}{\tilde{\sigma}_1^2}\frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2\tilde{\sigma}_1^2}+1\right)} & n \leq L \\ \frac{1}{\tilde{\sigma}_2^2}\frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2\tilde{\sigma}_2^2}+1\right)} & n > L \end{cases} . \tag{6.36}$$

In this case, the hyperparameter λ_1 controls the sparseness of the model. Large values lead to sharp distributions while very small values lead to uniform distributions. In Section 6.2.3 it is shown that λ_1 controls the relative weighting of the prior term relative to the misfit term. The overall weights generated by equation (6.36) are

similar in a qualitative sense to that of the ℓ^1 case. Once again, large reflectivities results in small weights and small reflectivities results in large weights. The relative weighting of variable 1 and 2 are controlled by $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$.

Huber norm

Another possibility is to use the Huber Norm (Huber, 1977). The normal distribution may be modified so that at large deviations it becomes Laplacian. This leads to the Huber loss function which is a mixture of the Gaussian and Laplacian loss functions. The following probability function

$$\Pr(\mathbf{x}'|\sigma_1, I) \propto \prod_{n=1}^{N} \exp\left(-\ell\left(u_n\right)\right), \tag{6.37}$$

where $\ell\left(\cdot\right)$ is the loss function and

$$u_n = \begin{cases} \frac{x'_n}{\sigma_1} & n \leq L \\ \frac{x'_n}{\sigma_2} & n > L \end{cases}, \tag{6.38}$$

leads to the Huber norm. The Huber loss function is defined as

$$\ell\left(u_n\right) = \begin{cases} \frac{1}{2}u_n^2 & |u_n| \leq \epsilon \\ \epsilon\left|u_n\right| - \frac{\epsilon^2}{2} & |u_n| > \epsilon \end{cases} \tag{6.39}$$

where $\epsilon > 0$. For example, Figure (6.1) shows the ℓ^2, ℓ^1 and Huber loss functions for $\epsilon = 3$. For standard deviations greater than $\epsilon = 3$ the loss function grows linearly rather than quadratically as in the case of the ℓ^2 loss function.

In Appendix G.3 the a priori constraints for the Huber norm are developed. In matrix form this is expressed as

$$\frac{\partial J_q}{\partial x_n} = \frac{1}{\sigma_{r_p}^2}\mathbf{Q}_h\mathbf{x}' \tag{6.40}$$

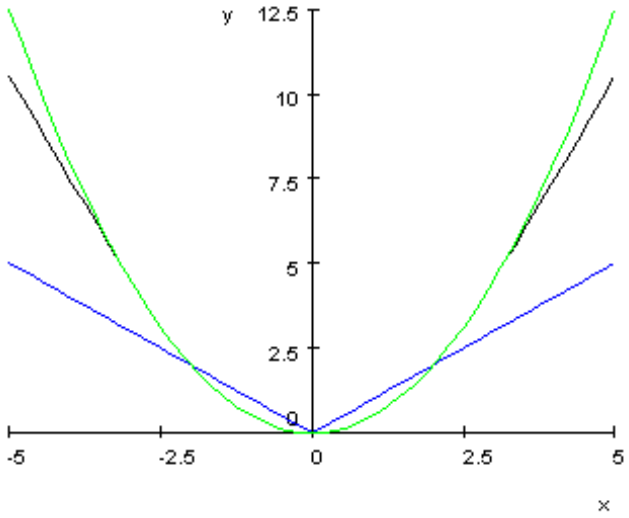

Figure 6.1: L2 (green), L1 (blue) and Huber (black) loss functions.

where $\mathbf{Q}_h$ is a diagonal matrix whose elements are

$$Q_{nn}=\begin{cases} \frac{1}{\tilde{\sigma}_1^2}\begin{cases} 1 & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| \le \epsilon \\ \epsilon\,\frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_1}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| > \epsilon \end{cases} & n \le L \\ \frac{1}{\tilde{\sigma}_2^2}\begin{cases} 1 & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| \le \epsilon \\ \epsilon\,\frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_2}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| > \epsilon \end{cases} & n > L \end{cases} . \tag{6.41}$$

For this norm ϵ is the number of standard deviations at which the function becomes Laplacian rather than Gaussian so it has a different meaning than in ℓ^p norm case. Once again, for large reflectivity small weights are calculated. For small reflectivity the weight of 1 is applied. Note, that if equation (6.41) is rewritten in terms of the

common scalar ϵ

$$Q_{nn}=\begin{cases} \frac{\epsilon}{\tilde{\sigma}_1^2} \begin{cases} \frac{1}{\epsilon} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| \leq \epsilon \\ \frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_1}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| > \epsilon \end{cases} & n \leq L \\ \frac{\epsilon}{\tilde{\sigma}_2^2} \begin{cases} \frac{1}{\epsilon} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| \leq \epsilon \\ \frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_2}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| > \epsilon \end{cases} & n > L \end{cases}, \tag{6.42}$$

the weights are identical to the ℓ^1 norm (equation 6.34) with the exception of the scalar.

6.2.3 Nonlinear inversion

The optimal solution for the AVO waveform inversion problem is found with the aid of Bayes' theorem. Assuming uniform uncorrelated Gaussian noise, the likelihood function is Gaussian with the misfit defined by equation (6.10). In Chapter 8 alternative misfit functions are explored.

The previous section developed a variety of *a priori* probability functions with long tailed distributions that lead to sparse reflectivity. The choice of one of these over the others should be guided by geologic concerns. The *a priori* probability function is combined with the likelihood function using Bayes' theorem (equation 3.4) following a derivation similar to Chapter 3. The case of the Cauchy *a priori* probability distribution, which is derived in Appendix H, leads to the nonlinear equation

$$\left[\mathbf{G}'^T\mathbf{G}' + \Theta\mathbf{Q}\right]\mathbf{x}' = \mathbf{G}'^T\mathbf{d} \tag{6.43}$$

where

$$\Theta = \lambda_1 \frac{\hat{\sigma}_N^2}{\sigma_{r_p}^2}, \tag{6.44}$$

$$\hat{\sigma}_N^2 = \frac{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{(M-1)}, \tag{6.45}$$

$\boldsymbol{\varepsilon} = \mathbf{G}'\mathbf{x}' - \mathbf{d}$ and $\mathbf{Q}$ is defined by equation (6.36).

There are two sources of nonlinearity in equation (6.43), the estimate of the reg-

ularization parameter Θ and the calculation of $\mathbf{Q}$. The matrix $\mathbf{Q}$ (equation 6.36) requires a previous estimates of $\mathbf{x}'$ to construct the diagonal weighting terms. This must be done in a bootstrap fashion. Since the actual inverse problem being solved is quite large, it is most efficiently solved using iterative techniques such as conjugate gradient (Shewchuck, 1994). Solving the inverse problem requires two nested loops. In the inner loop the conjugate gradient algorithm is used to solve equation (6.43) using the previously calculated values of $\mathbf{x}'$ and $\mathbf{Q}$. The maximum number of conjugate gradient iterations is used as a parameter to help stabilize the solution. By limiting the number of conjugate gradient iterations, the smaller eigenvalues are excluded from the solution (Shewchuck, 1994). After solving the reflectivity estimate $\mathbf{x}'$, the constraining matrix $\mathbf{Q}$ and the penalty term Θ are updated. Iteratively updating these parameters and re-estimating the reflectivity parameters constitute the outer loop. Generally, a satisfactory sparse solution is obtained after 3 to 5 iterations of the outer loops. For the first outer loop iteration, the inversion is run as an unconstrained inversion by setting $\mathbf{x}' = \mathbf{0}$. Care must be taken in the first outer loop not to put too much detail into the solution or the problem will not converge. This can be controlled by carefully setting the maximum number of conjugate gradient iterations (inner loop) to a value that limits resolution. After the final iteration, the parameters can be transformed to a more natural parameterization using equation (6.22). In fact, the parameters may be transformed to any convenient parameterization using the two-term transform matrix (equation B.57).

The penalty term Θ (equation 6.44) is the inverse ratio of the signal-to-noise multiplied by the hyperparameter λ_1. It controls the relative weighting of the constraints (defined by $\mathbf{Q}$) versus the data constraints. If Θ is zero, then equation (6.45) is equivalent to least squares. If Θ is large, then the $\mathbf{Q}$ matrix dominates the solution. In this case, it biases the solution so as to create a sparse reflectivity series. The nonuniqueness introduced by the band-limited nature of the seismic data is addressed by the weighting function. Reflectivities are preferentially located where the weights are small.

In the calculation of equation (6.44) there is a scaling problem. The seismic data are often arbitrarily scaled, while $\sigma^2_{r_p}$ is defined from the well control. Similar to Chapter 3, $\sigma^2_{r_p}$ may be alternatively estimated from the initial estimate of $\mathbf{r}_p$. Thus, both the denominator and numerator have the same unknown scalar, which cancels out.

Even after addressing the scaling issue, there is some subjectivity about the value Θ. For the Cauchy distribution, there is the question about what value to use for the hyperparameter λ_1. Strictly speaking, $\lambda_1 = 1$ for the Cauchy distribution, but other choices may be made. Note that the difference between the ℓ^1 and Huber weights is a scaling factor similar to λ_1. To optimize its value, I experimented with different values of λ_1 constructing an L-curve (Hansen, 1998).

The analysis is complicated by the fact that the number of conjugate gradient iterations also acts as a penalty term. This necessitated constructing a grid of values to construct the L curve. Lower maximum iteration values lead to faster solutions, but the reflectivity estimates tend to be band-limited. More iterations provide more eigenvalues from which to construct a full spectrum reflectivity series. Too many iterations lead to the solution oscillating rapidly and not converging.

Two methods were experimented with to numerically solve the inverse problem. The first method required that the linear operator be symmetric. The matrix $\left[\mathbf{G}'^T\mathbf{G}' + \Theta\mathbf{Q}\right]$ was explicitly calculated using sparse matrices. For less than 1000 time samples and 50 offsets this lead to a computationally efficient algorithm. For larger problems, I ran into memory problems. I was running on a machine with one gigabyte of memory. The second method solved equation (6.43) in standard form

$$\begin{bmatrix} \mathbf{G}' \\ \sqrt{\Theta}\mathbf{Q}^{1/2} \end{bmatrix} \mathbf{x}' \approx \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}. \tag{6.46}$$

The conjugate gradient algorithm for this form is not as efficient, but the symmetric matrix $\left[\mathbf{G}'^T\mathbf{G}' + \Theta\mathbf{Q}\right]$ is not required. There is a substantial savings in time and memory by avoiding the calculation of this matrix. Further, the linear operator $\mathbf{G}'\mathbf{x}'$ need never be explicitly calculated. It may be calculated on the fly, in stages (Claerbout, 1992). Ultimately, I was able to construct an algorithm which was as fast as the first, but more memory efficient.

6.3 Synthetic example demonstrating AVO waveform inversion in presence of NMO stretch and offset dependent tuning

Recall from Chapter 5, that in the presence of NMO stretch and offset dependent tuning, the AVO inversion estimates of Class III and IV gas sands are distorted.

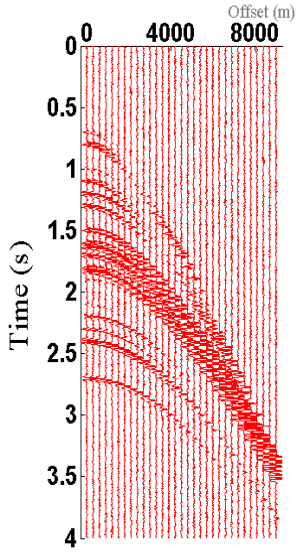

Figure 6.2: Synthetic seismic gather. The Class III and IV isolated reflectors are at 1.5 s and 1.7 s zero offset time respectively. The tuned Class III and IV reflectors are at 1.6 s and 1.8 s zero offset time respectively.

In contrast, Class I and II gas sands, along with regional reflectors that follow the mudrock trend are not distorted. To demonstrate that the new approach performs better in the presence of NMO stretch and offset dependent tuning a simple synthetic model (Figure 6.2) with sparse reflectivity was generated. The AVO behavior of most of the reflectors in the synthetic gather follow the mudrock trend, but several Class III and IV anomalies are also present. Both isolated and tuned anomalies were created. The synthetic data was generated using a convolutional model with a Ricker wavelet with a dominant frequency of 32.5 Hz. In order to isolate the effects of NMO stretch and differential tuning on the AVO inversion, the reflectivity was generated using the three-term Shuey equation (1985) using the Gardner relationship $R_d = 0.25R_\alpha$ to calculate the density reflectivity. Random noise was added to give a signal-to-noise ratio of 4:1. A constant background velocity was used so there would be a simple angle-to-offset mapping. The synthetic gather was generated with the maximum offset equal to four times the target depth, so angles out to 65° were available for the inversion, though only angles to 45° were actually used. These large angles were included to highlight the distortions. Figure 6.3 shows the prestack synthetic gather

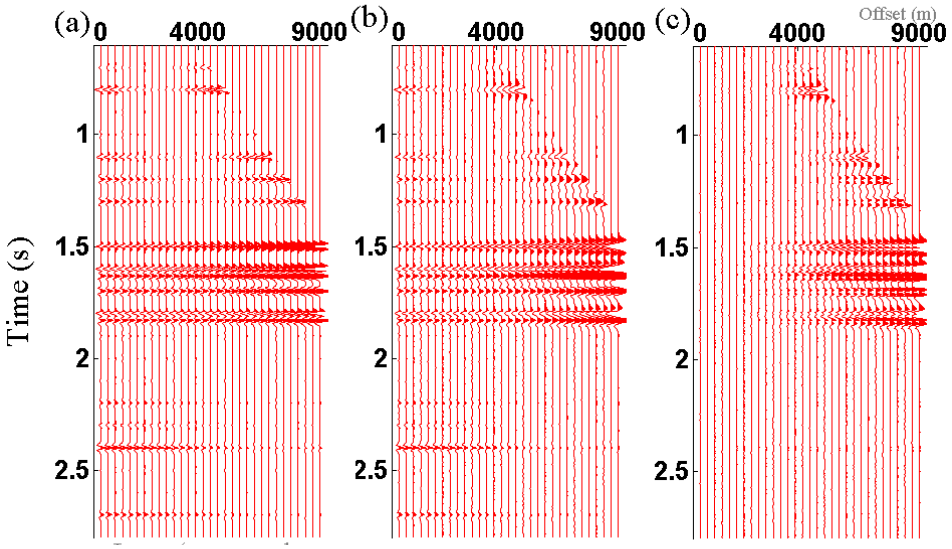

Figure 6.3: The synthetic gather (a) generated without moveout is compared to the NMO corrected gather (b) while (c) shows the difference between the two.

after NMO correction. This is compared to the same gather but generated without NMO. The difference (Figure 6.3c) highlights the theoretical error introduced by the NMO correction.

The AVO inversion was performed using the Smith and Gidlow formulation using angles from 0° to 45°. The Smith and Gidlow formulation was used rather than the two-term Shuey approximation since the former is exact under the assumptions made in the model design, while the latter is not. The reflectivity estimates were then transformed to intercept and gradient for display purposes as shown in Figure 6.4. As expected, there is no error in the intercept term while the gradient term only shows error for both the Class III and Class IV anomalies. Reflectivity of the interfaces whose α/β relationship fall along the mudrock trend are predicted well even though the events have undergone NMO stretch. Figure 6.4 also shows the reflectivity in the crossplot domain. The Class III and IV anomalies show significant scatter.

The same data was used to perform the AVO waveform inversion. The results of this are shown in Figure 6.5 and 6.6. Figure 6.5 shows the input to the inversion, the estimated model reconstructed from the parameter estimates and the difference. The model matches the data closely with the difference showing only random noise. The estimated reflectivity is shown in Figure 6.6. The estimate is excellent for all reflectors including the Class III and IV anomalies that were undergoing NMO stretch and had

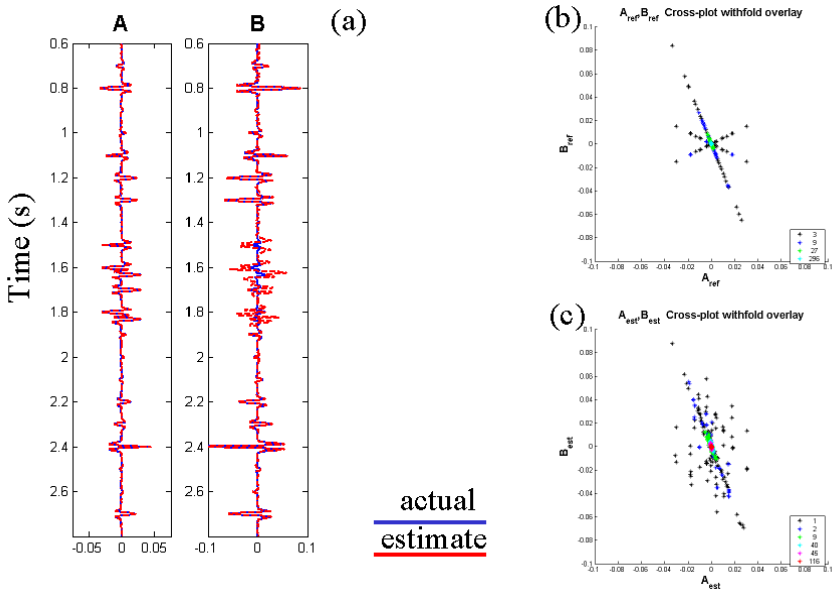

Figure 6.4: The traditional AVO estimates (a) for intercept A and gradient B (red) are compared to the ideal results (blue). Also, the ideal data (b) is compared to the estimated data (c) in the cross-plot domain. Note the estimated Class III and IV anomalies are scattered in cross-plot space.

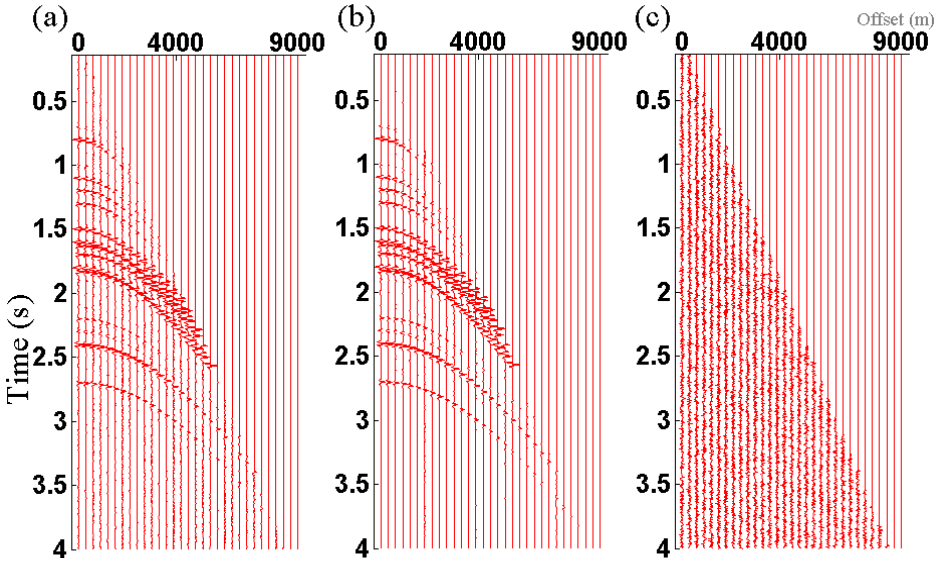

Figure 6.5: Comparison of the model estimated from the two-term AVO waveform inversion (b) and the input (a). Note the good match between the two with only random noise evident in the scaled (5X) difference (c).

differential tuning. This is confirmed by crossplotting the reflectivity attributes.

The AVO inversion estimates shown in Figure 6.6 are filtered with the source wavelet. Figure 6.7 shows the unfiltered reflectivity estimates shown in terms of the fluid factor, P-wave and S-wave impedance reflectivity. The match for the large reflectivity is excellent. With suitable modifications to the constraints to account for the background trend, these sparse spike reflectivity estimates may be used to estimate the P-wave and S-wave impedances as shown later in the chapter.

6.4 Wavelet

A practical implementation issue is that of obtaining the wavelet for the inversion. Referring to Chapter 1, the data have been processed through deconvolution so that ideally the seismic wavelet should have a uniform amplitude spectrum and a consistent phase across the data set. To correct for residual amplitude and phase distortions, a seismic wavelet may be designed using the well control (White, 1980; Lines and Treitel, 1985). Generally, each well tie will have a slightly different wavelet for a variety of reasons such as spatially changing noise and well log quality issues.

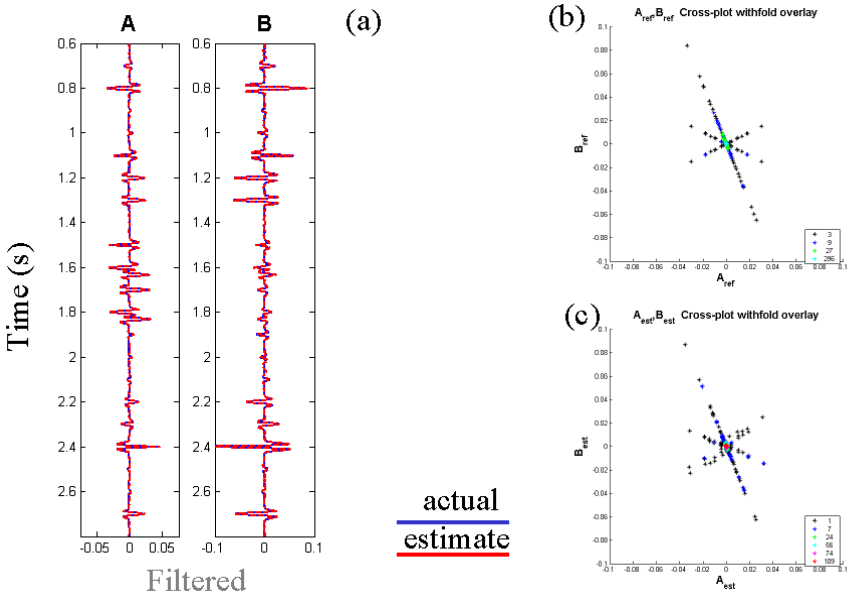

Figure 6.6: The estimated two-term AVO waveform results (a) for intercept A and gradient B (red) are compared to the ideal results (blue). Also, the ideal data (b) is compared to the estimated data (c) in the cross-plot domain. Note the good agreement for the Class III and IV anomalies.

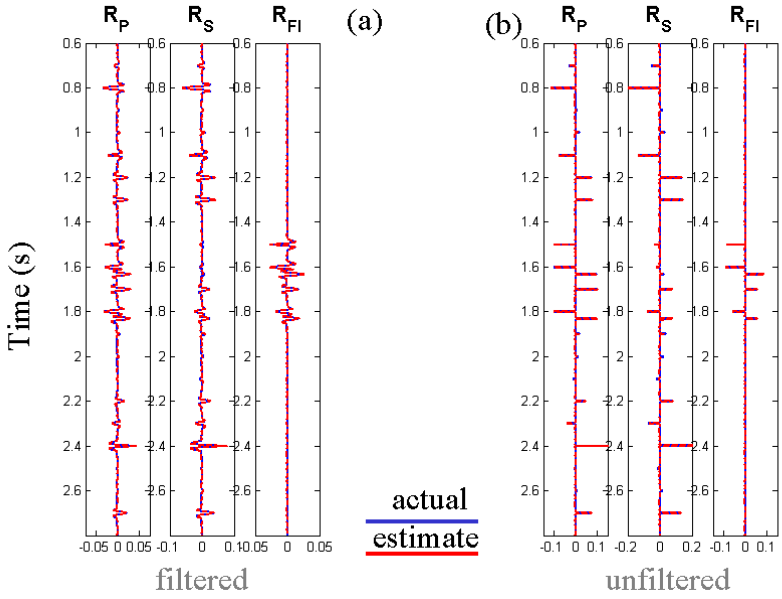

Figure 6.7: The estimated (red) Fluid factor, P- and S-wave impedance reflectivities based on the AVO waveform inversion compared to the ideal (blue). Panel (b) shows the full spectrum estimate while (a) shows the result filtered by the estimated source wavelet.

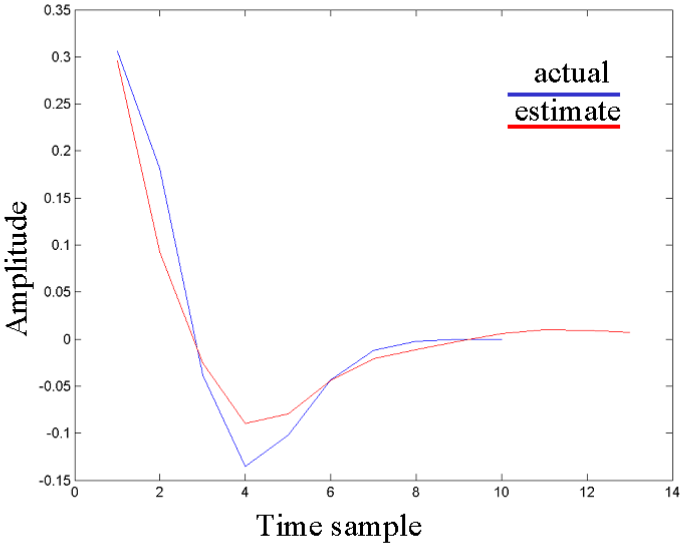

Figure 6.8: Extracted wavelet (red) versus ideal wavelet (blue). Note only the non-negative time samples are shown.

For this reason, often a constant phase correction is applied to the seismic data rather than applying the frequency dependent phase correction as suggested above. In order to allow for spatial variations in the wavelet, the amplitude spectrum is estimated based on an ensemble average of the autocorrelations within each CDP gather. If it is assumed that the reflectivity has a white amplitude spectrum, then the wavelet spectrum is equivalent to the estimated spectrum. Further, if it is assumed that the wavelet is zero phase, then the wavelet may simply be calculated from the estimated amplitude spectrum. Figure 6.8 shows the estimate of the wavelet following this method for the previous synthetic example compared with the ideal wavelet used to generate the synthetic. Only the non-negative time samples are shown for the wavelet. There are some discrepancies between the estimate and the ideal, partly due to the limited number of samples used to estimate the wavelet. Figure 6.9 shows the estimate of the AVO waveform inversion using the estimated wavelet. At this scale it is difficult to discern any error. However, upon closer inspection there are small errors with the estimate. The four reflectors undergoing NMO stretch and offset dependent tuning (Figure 6.10) display differences between the estimate and ideal. However, when the result of the AVO waveform inversion using the estimated wavelet (Figure 6.9) is compared to that done with the ideal wavelet (Figure 6.6) over the whole interval there is only a small degradation.

In actual practice there will also be residual phase errors in the seismic data. Phase errors were introduced into the synthetic data to understand the impact of these errors. First, a 180° rotation was applied to the seismic data. The wavelet was estimated and then the AVO waveform inversion was performed on the distorted seismic data. For display purposes the reference (ideal) reflectivity is shown with a 180° rotation as well. For this case the estimate is almost identical to the ideal (Figure 6.11). This is due to the fact that a 180° rotation is just a scalar, and as such does not influence the inversion.

Things become more complicated with a 90° phase rotation. For zero offset reflectivity, there is an inherent ambiguity between two tuned reflectors (dipole) and a single reflector that has been rotated 90°. For non-zero offset data, the offset dependent tuning should allow one to potentially distinguish between the two situations. Winslow (2000) uses this as the basis of a wavelet estimation technique. Thus, having an uncorrected 90° rotation should introduce theoretical error into the model. Figure 6.12 and 6.13 show the results of the inversion based on the seismic data with

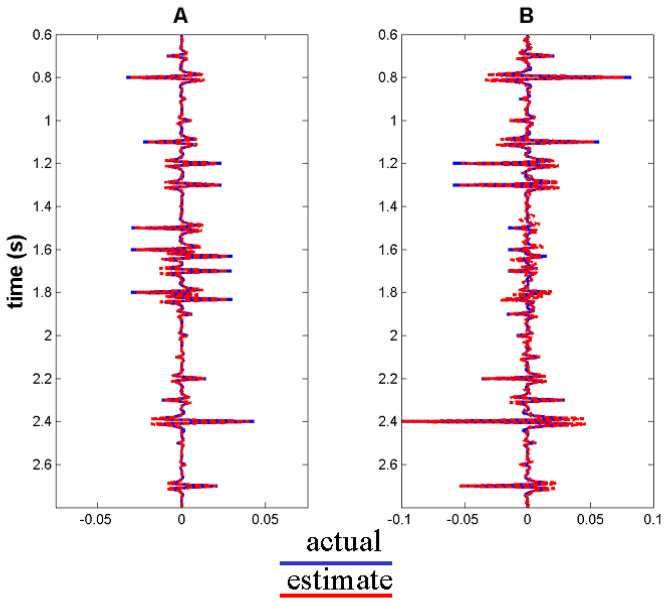

Figure 6.9: Comparison of estimated (red) versus ideal (blue) intercept and gradient using estimated wavelet for AVO waveform inversion.

the input rotated 90°. For comparison, the reference reflectivity is rotated 90° as well. The error is somewhat worse than using the ideal wavelet but substantially better than following the traditional AVO methodology.

The last two models have explored the impact of phase error on the problem. The last model in this section explores the impact of using a wavelet with the wrong spectrum. The synthetic data was created using a Ricker wavelet with a 35 Hz dominant frequency. Instead of using this wavelet an Ormsby filter with 10/20−60/90 Hz corner frequencies was used in the inversion. Figures 6.14 and 6.15 show the results of this inversion. The results of this inversion are similar to those of doing the AVO inversion in a traditional fashion (Figure 6.4). This is a worst case scenario, provided some effort is put into estimating the wavelet, the results of the AVO waveform inversion are better than the traditional AVO methodology.

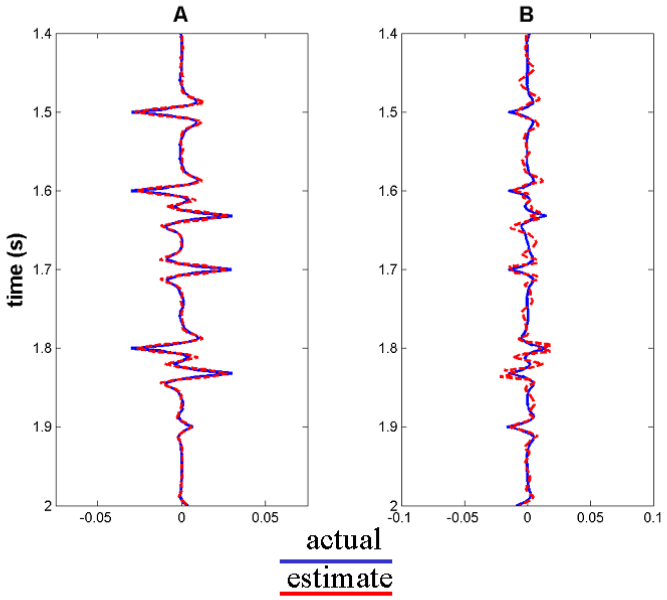

Figure 6.10: Zoom of comparison of estimated (red) versus ideal (blue) intercept and gradient using estimated wavelet for AVO waveform inversion.

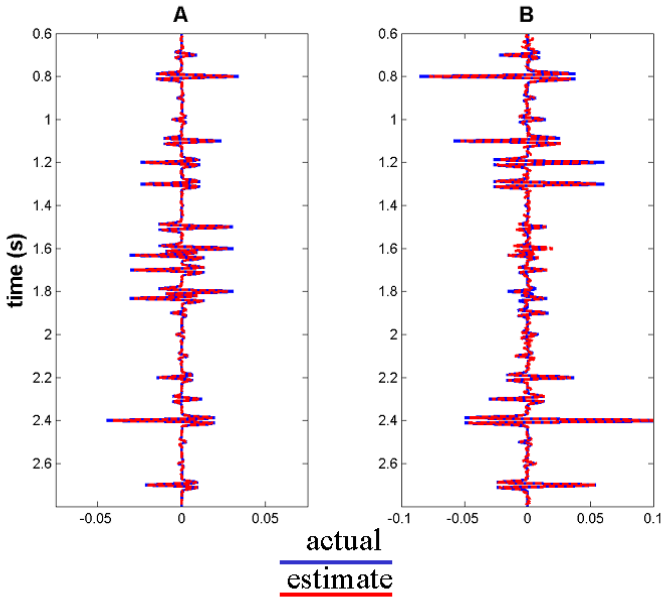

Figure 6.11: Comparison of estimated (red) versus ideal (blue) intercept and gradient. Estimate from AVO waveform inversion performed with wavelet 180 degrees out of phase.

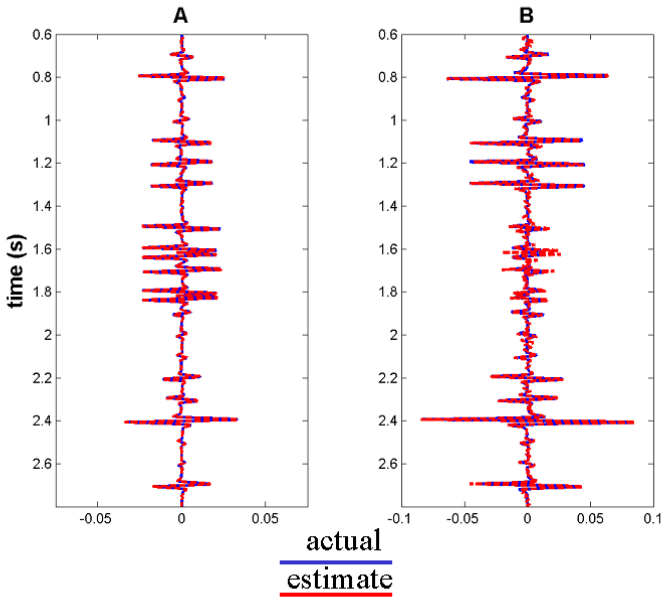

Figure 6.12: Comparison of estimated (red) versus ideal (blue) intercept and gradient. Estimate from AVO waveform inversion performed with wavelet 90 degrees out of phase.

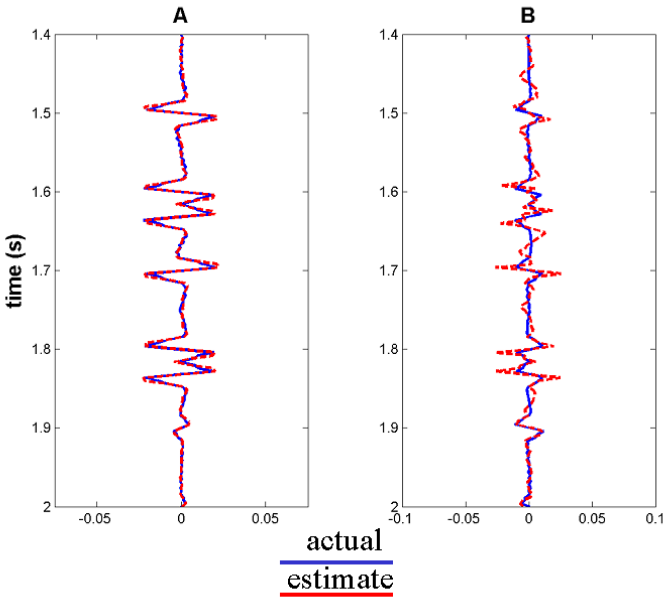

Figure 6.13: Zoom of comparison of estimated (red) versus ideal (blue) intercept and gradient. Estimate from AVO waveform inversion performed with wavelet 90 degrees out of phase.

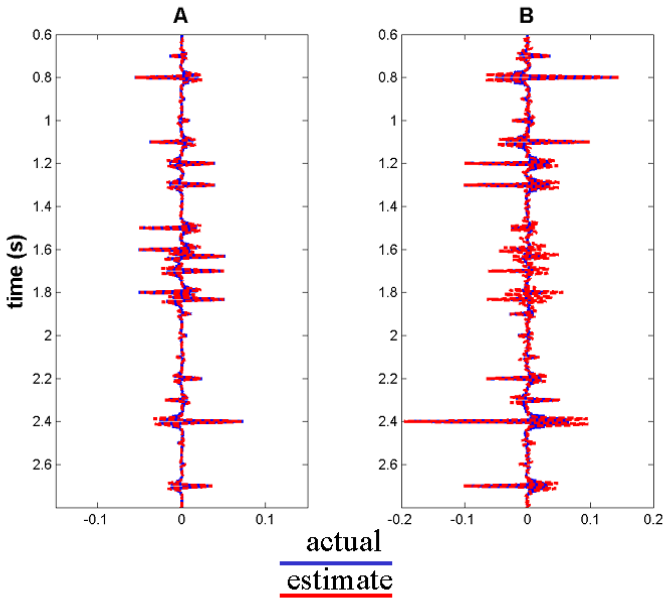

Figure 6.14: Comparison of estimated (red) versus ideal (blue) intercept and gradient. Estimate from AVO waveform inversion performed with 10/20-60/90 Hz wavelet instead of correct wavelet. Ideal reflectivity shown with 10/20-60/90 Hz filter for comparison purposes.

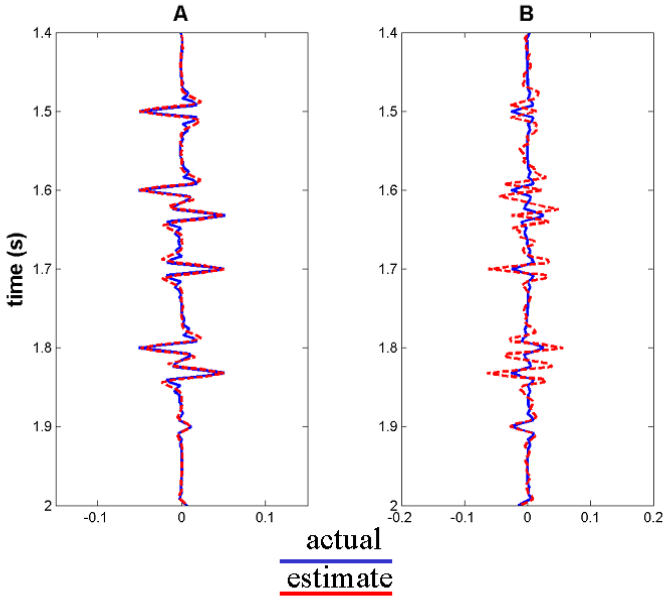

Figure 6.15: Zoom of comparison of estimated (red) versus ideal (blue) intercept and gradient. Estimate from AVO waveform inversion performed with 10/20-60/90 Hz wavelet instead of correct wavelet. Ideal reflectivity shown with 10/20-60/90 Hz filter for comparison purposes.

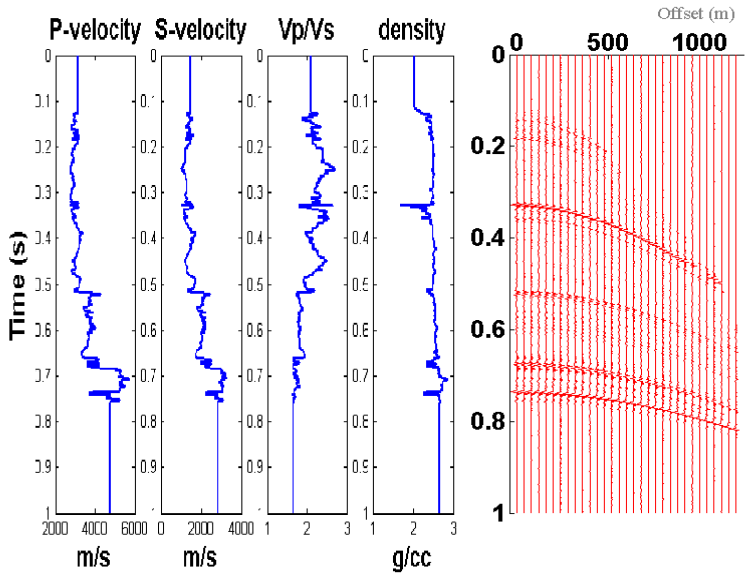

Figure 6.16: Input P- and S-wave velocity, α/β ratio, density logs, and resultant synthetic gather.

6.5 Examples

6.5.1 Halfway synthetic model

The preceding examples input synthetic model's reflectivity was sparse so the *a priori* assumptions were perfectly met. The next model tests what happens if the reflectivity does not match the *a priori* assumptions perfectly. For this second example, the reflectivity is based on a well log data from northeastern British Columbia. The input well logs are shown in Figure 6.16 together with the synthetic seismic gather that was generated. In this case, the reflectivity was generated with the Zoeppritz equations and convolved with a 10/14-90/110 Hz zero phase wavelet. Once again noise was added to give a signal-to-noise ratio of 4:1. The AVO inversion was done over angles from 0 ° to 35 °. The filtered estimate is an excellent match (Figure 6.17). The sparse spike reflectivity shows some errors particularly around 0.34 seconds on the S-wave impedance reflectivity.

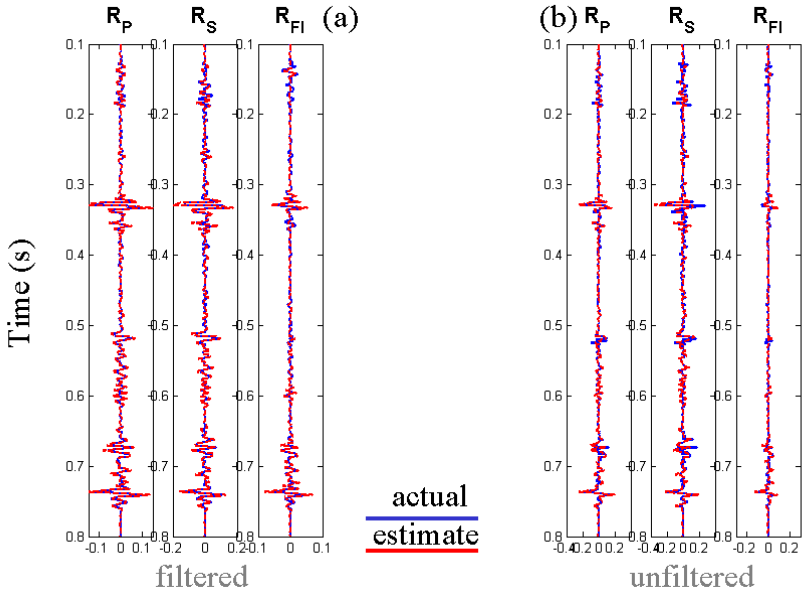

Figure 6.17: The estimated reflectivity (red) of fluid factor, P- and S-wave impedance reflectivity based on the AVO waveform inversion compared to the ideal (blue). Panel (b) shows the full spectrum estimate while (a) shows the result filtered by the estimated source wavelet.

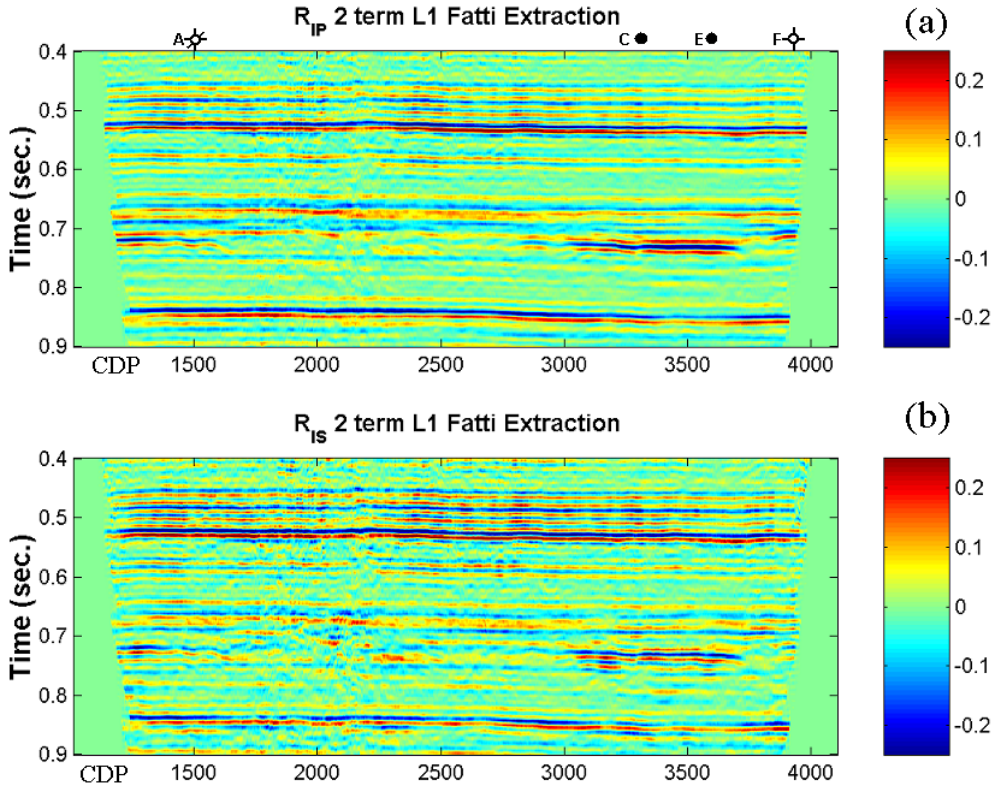

Figure 6.18: P-wave (a) and S-wave (b) impedance reflectivity estimate from traditional two-term AVO inversion.

6.5.2 Halfway seismic data example

The Halfway data example was inverted using the two-term AVO waveform inversion. The results are shown in Figure 6.19. These can be compared to the results following the traditional AVO methodology (Figure 6.18). Note that the AVO waveform S-wave impedance reflectivity has higher frequency on the Charlie Lake reflector at 0.67 seconds. The Peace River marker (0.52 seconds) should be an isolated reflector. The AVO waveform inversion S-wave impedance reflectivity shows an isolated reflector for the Peace river where as the traditional AVO results displays ringing over this time interval. Both these reflections are due to a rapid increase in the S-wave impedance and as such should be ideally reflectivity spikes. In addition, the AVO waveform inversion seems to have a better signal-to-noise ratio.

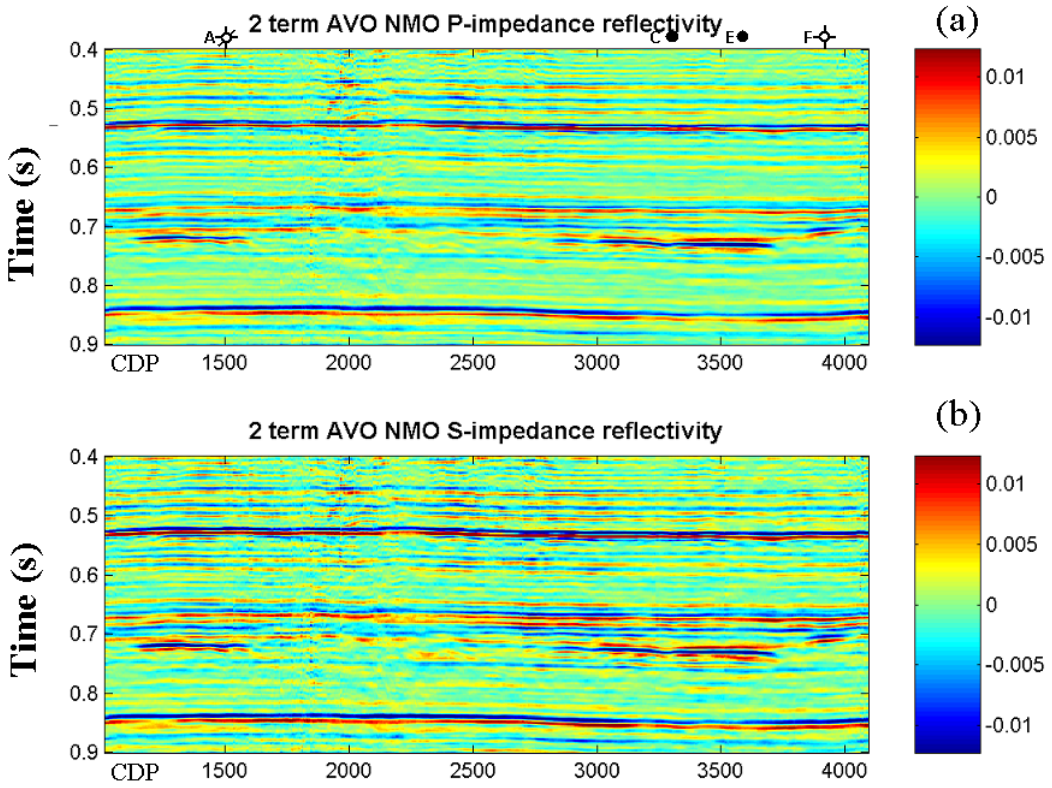

Figure 6.19: P-wave (a) and S-wave (b) impedance reflectivity estimate from two-term AVO waveform inversion.

6.5.3 Colony seismic data example

This seismic data example is from northeast Saskatchewan. The well at CDP 3030 encountered two gas sands, an upper blanket sand and a channel sand. Both of these gas sands should show fluid factor anomalies. The data have a good signal-to-noise ratio with signal extending from 10 to 130 Hz. The range of angles used in the inversion was 0° to 36°. The wavelet was calculated following Walden and White (1984).

Figure 6.20 shows the results of the AVO waveform inversion. The bottom panel shows the fluid factor. The anomaly at 0.3 seconds is the top blanket sand and the anomaly at 0.4 seconds is the channel sand. The anomaly at 0.65 seconds with the peak over the trough represents a carbonate reflector, the fluid factor is responding to a change in lithology, which is off the mudrock trend. For comparison, Figure 6.21 shows the results from performing a traditional AVO inversion on NMO corrected gathers. Notice the significant improvement in signal-to-noise and continuity in the AVO waveform inversion. This is probably due to two factors. The fact the waveform is included in the formulation of the problem limits how quickly the filtered reflectivity can change temporally. Secondly, and probably more importantly, the regularization of the problem improves the condition number of the problem, decreasing how the noise gets amplified as part of the inversion. Figure 6.22 shows the prestack seismic data and the AVO waveform inversion model.

6.6 Impedance Inversion

Rather than solving for reflectivity it is sometimes desirable to output impedances as this simplifies the interpretation (Section 1.4.2). However, the seismic data typically are missing low frequencies implying some sort of constraints must be used to obtain a unique solution for impedance. This low frequency trend may be supplied using regional well control and information from an interval velocity travel time inversion from the seismic data.

In the case of small reflection coefficients, the P-wave impedance I_p may be calculated by integrating and taking the exponential of the P-wave reflectivity r_p (Peterson, 1955)

$$I_{p_j} \approx I_{p_0} \exp\left\{ 2 \sum_{j=1}^{L} r_{p_j} \right\}. \tag{6.47}$$

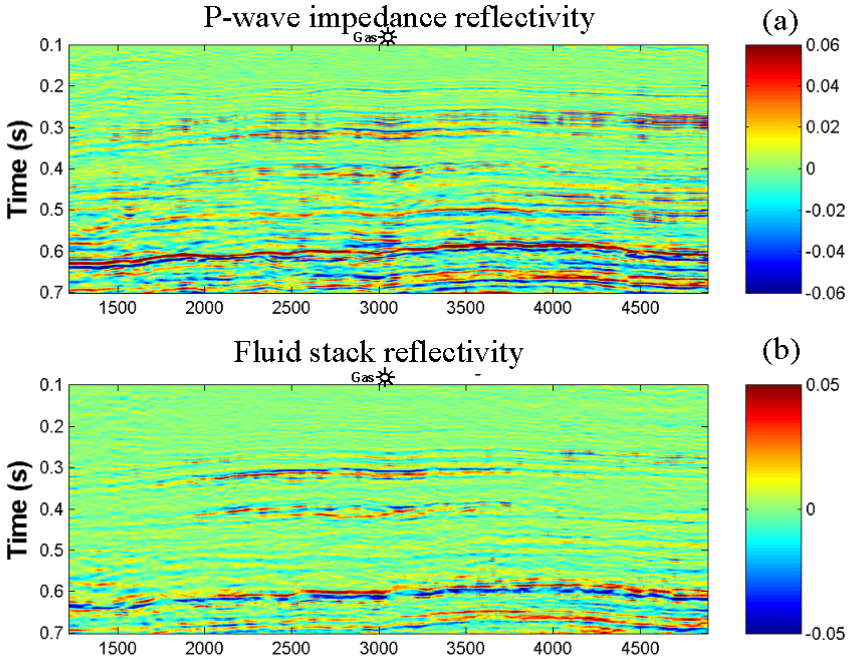

Figure 6.20: P-wave impedance (a) and fluid (b) reflectivity estimate from two-term AVO waveform inversion.

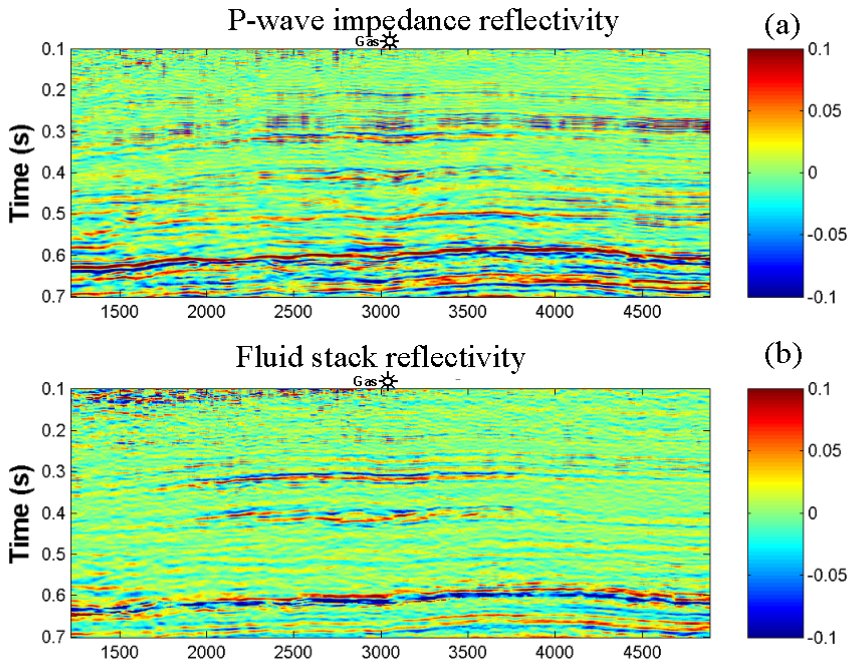

Figure 6.21: P-wave impedance (a) and fluid (b) reflectivity estimate from two-term traditional AVO inversion.

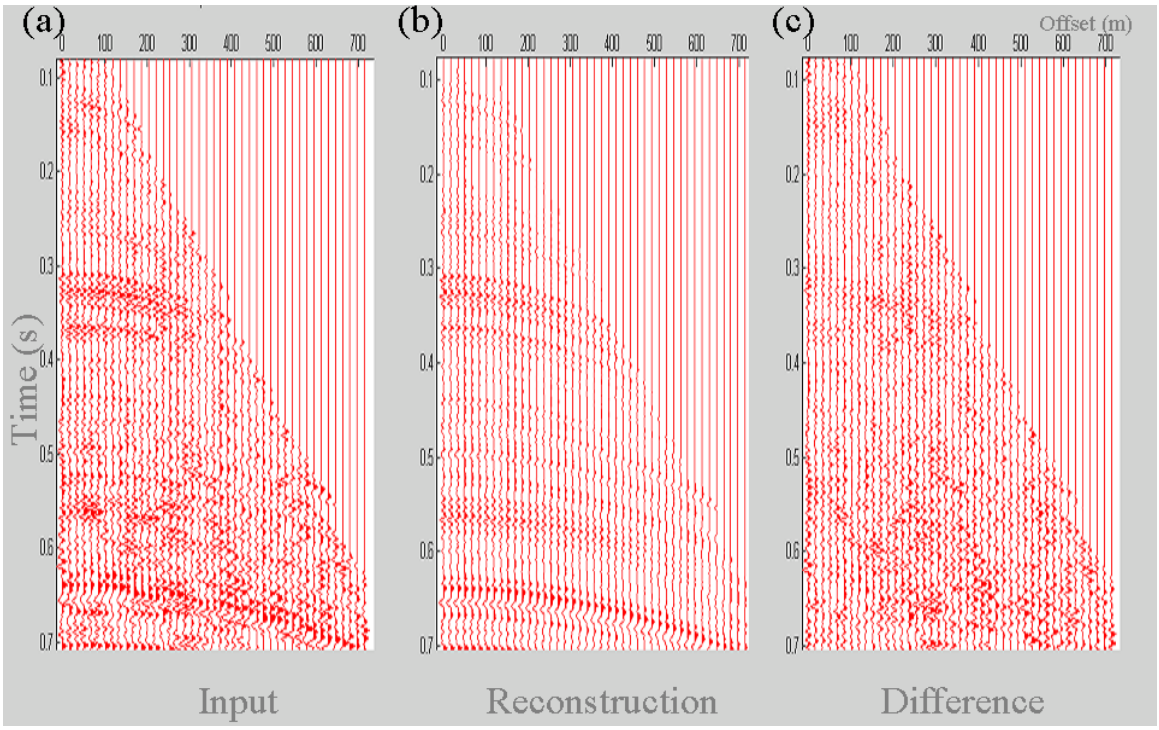

Figure 6.22: The input (a) is compared to estimated two-term AVO waveform model (b). Note the good match between the two with the difference (c) showing mostly random noise.

In order to make equation (6.47) more convenient to work with, the logarithm of the P-wave impedance ξ_k is calculated

$$\xi_j = \xi_0 + 2\sum_{j=1}^{L} r_{p_j}, \tag{6.48}$$

or alternatively its deviation $\Delta\xi_j = \xi_j - \xi_0$

$$\Delta\xi_j = 2\sum_{j=1}^{L} r_{p_j}. \tag{6.49}$$

Equation (6.49) provides a framework to constrain the low frequency impedance trend. Ideally the low frequency impedance trend is known from well control or travel time inversion or some combination of the two. Using equation (6.49) the low frequency impedance trend may be specified, by defining $\Delta\xi_j$ at regular intervals (for example, every 0.1 seconds). The number of constraints, J, imposed should be inversely proportional to the lowest frequency available in the seismic data. These constraints are inexact and have error. This size of this error may be estimated from the well control if the impedances are assumed to be locally stationary around the point constraint.

Written in matrix notation, the point constraints are

$$\mathbf{d}_{\xi_p} = \mathbf{P}\mathbf{r}_p, \tag{6.50}$$

where $\mathbf{d}_{\xi_p}$ is a vector specifying the P-wave point constraints, and $\mathbf{P}$ is the appropriate linear summation operator. A variation of this is that the average impedance over some interval must equal some value within a certain error. Note, that the S-wave impedance may be related to the S-wave impedance reflectivity in a similar fashion.

Augmenting the simultaneous equations (6.9) with the P-wave and S-wave impedance point constraints results in

$$\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \\ \mathbf{d}_{\xi_p} \\ \mathbf{d}_{\xi_s} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1^n\mathbf{W}\mathbf{N}_1\mathbf{M}_1^f\mathbf{E}_1 & \mathbf{M}_1^n\mathbf{W}\mathbf{N}_1\mathbf{M}_1^f\mathbf{F}_1 \\ \mathbf{M}_2^n\mathbf{W}\mathbf{N}_2\mathbf{M}_2^f\mathbf{E}_2 & \mathbf{M}_2^n\mathbf{W}\mathbf{N}_2\mathbf{M}_2^f\mathbf{F}_2 \\ \mathbf{P}_{\xi_p} & 0 \\ 0 & \mathbf{P}_{\xi_s}\boldsymbol{\gamma} \end{bmatrix} \begin{bmatrix} \mathbf{r}_p \\ \tilde{\mathbf{r}}_s \end{bmatrix}, \tag{6.51}$$

where $\boldsymbol{\gamma}$ is a diagonal matrix whose elements consist of the S-wave to P-wave velocity ratio. This may be written in the more compact notation

$$\mathbf{d} = \mathbf{G}\mathbf{x}, \tag{6.52}$$

where $\mathbf{G}$ is the linear operator

$$\mathbf{G} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 \\ \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K \\ \mathbf{P}_{\xi_p} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\xi_s} \boldsymbol{\gamma} \end{bmatrix}, \tag{6.53}$$

$\mathbf{d}$ is the data vector

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \\ \mathbf{d}_{\xi_p} \\ \mathbf{d}_{\xi_s} \end{bmatrix}, \tag{6.54}$$

and $\mathbf{x}$ is the unknown parameter vector. The data vector contains a total of $M = L \times (K + 2J)$ elements and the parameter vector $N = 2L$ elements.

To solve this, equation (6.52) must be first transformed so the variables are independent (equation 6.22). The result is similar to matrix (6.26) and may be solved in a similar fashion as in Section 6.2.3. However, the elements of the misfit vector $\boldsymbol{\varepsilon} = \mathbf{G}'\mathbf{x}' - \mathbf{d}$ must be suitably weighted. The misfit is no longer uniform. The misfit due to fitting the seismic data and that due to fitting the point constraints have different variances. The diagonal weighting matrix $\mathbf{W}_e$ is introduced to compensate for the difference in relative sizes between the two misfits. This matrix $\mathbf{W}_e$ has ones for the data misfit elements, and some scalar for the point constraint misfit elements. The choice of this scalar controls the how well the point constraints are honored relative to the data constraints.

Having introduced this weighting matrix, this leads to the solution

$$\left[\mathbf{G}'^T \mathbf{W}_e \mathbf{G}' + \Theta \mathbf{Q}\right] \mathbf{x}' = \mathbf{G}'^T \mathbf{W}_e \mathbf{d} \tag{6.55}$$

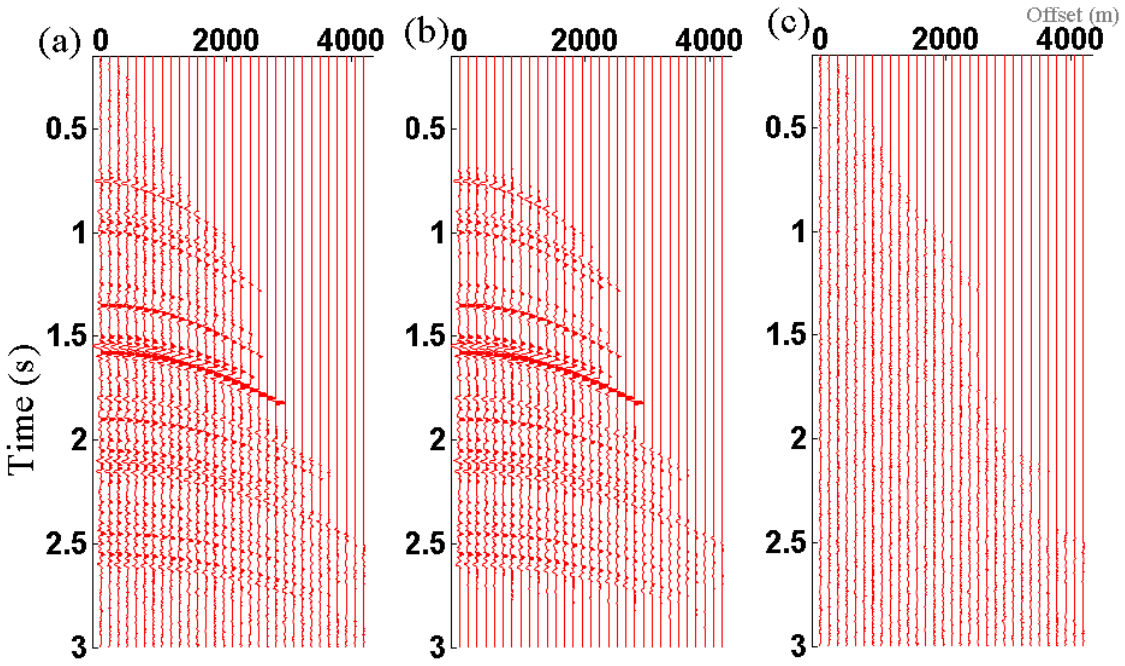

Figure 6.23: Model (b) estimated from the prestack impedance inversion compared to the input (a) and scaled (5X) difference (c).

where

$$\Theta = \lambda_1 \frac{\sigma_N^2}{\sigma_{r_p}^2}, \tag{6.56}$$

and

$$\sigma_N^2 = \frac{\boldsymbol{\varepsilon}^T \mathbf{W}_e \boldsymbol{\varepsilon}}{(M-1)}. \tag{6.57}$$

Once again $\boldsymbol{\varepsilon} = \mathbf{G}'\mathbf{x}' - \mathbf{d}$ and $\mathbf{Q}$ is defined by equation (6.36). Numerically, this is solved as in Section 6.2.3. Upon solving for $\mathbf{x}'$, the solution is transformed to $\mathbf{x}$ and then the P-wave and S-wave impedance reflectivity are integrated using equation (6.47) to estimate the impedances. This solution technique is similar to the poststack impedance inversion described by Oldenburg et al. (1983). However, they used linear programming to implement the ℓ^1 solution. This solution technique is too expensive for prestack data, necessitating the use of conjugate gradient and suitable weighting functions.

To test this, a simple blocky model was constructed where most reflections follow the mudrock trend. A Class III anomaly undergoing offset dependent tuning was introduced at 1.5 seconds. The reflectivity was generated using the Zoeppritz

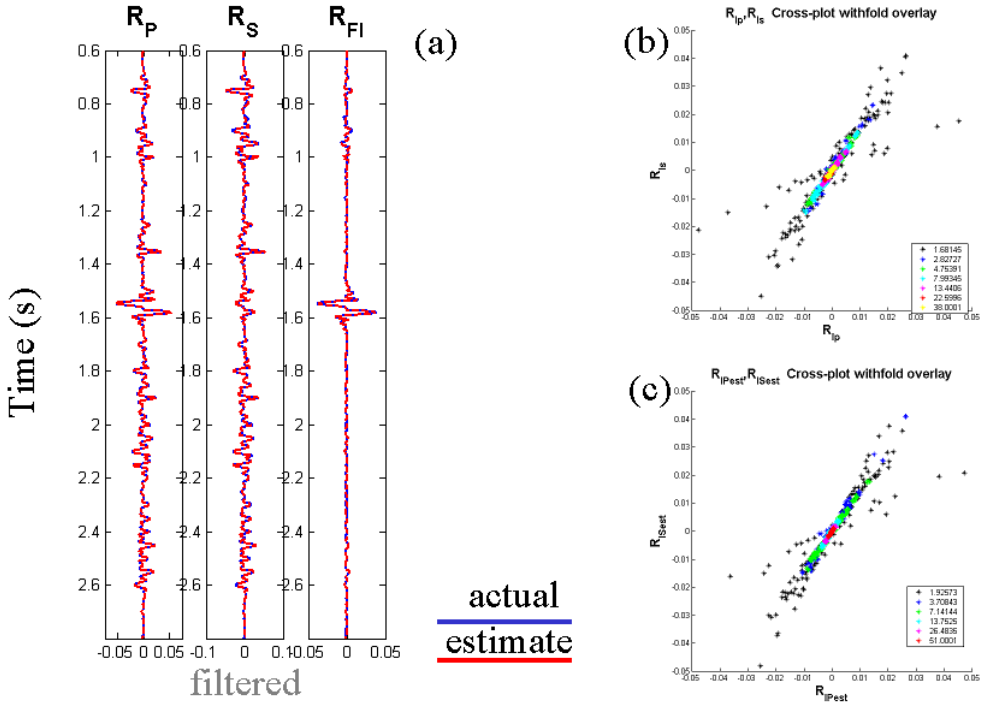

Figure 6.24: The AVO waveform estimates (red) of the P-wave Impedance, S-wave impedance and fluid stack reflectivity (a) vs the ideal (blue). Also, the ideal data (b) is compared to the estimated data (c) in the cross-plot domain.

equations and was convolved with 5/10-40/50 Hz wavelet. Noise was added to give a signal-to-noise ratio of 4:1. Figure 6.23 shows the input synthetic gather and the model gather generated by the inversion. Note the excellent match between the two with only random noise showing up in the difference display (Figure 6.23c).

The band-limited reflectivity is shown in Figure 6.24a. To quantify the match, estimated (Figure 6.24c) and ideal (Figure 6.24b) reflectivity are crossplotted. The estimated data shows the same trends. The tuned Class III anomaly is correctly identified. The estimated full spectrum reflectivity is also a good match with the ideal reflectivity (Figure 6.25a). The P-wave and S-wave impedances are estimated by integrating their respective reflectivities (equation 6.47) again matching the ideal accurately (6.25b). Once again the tuned Class III anomaly is correctly predicted.

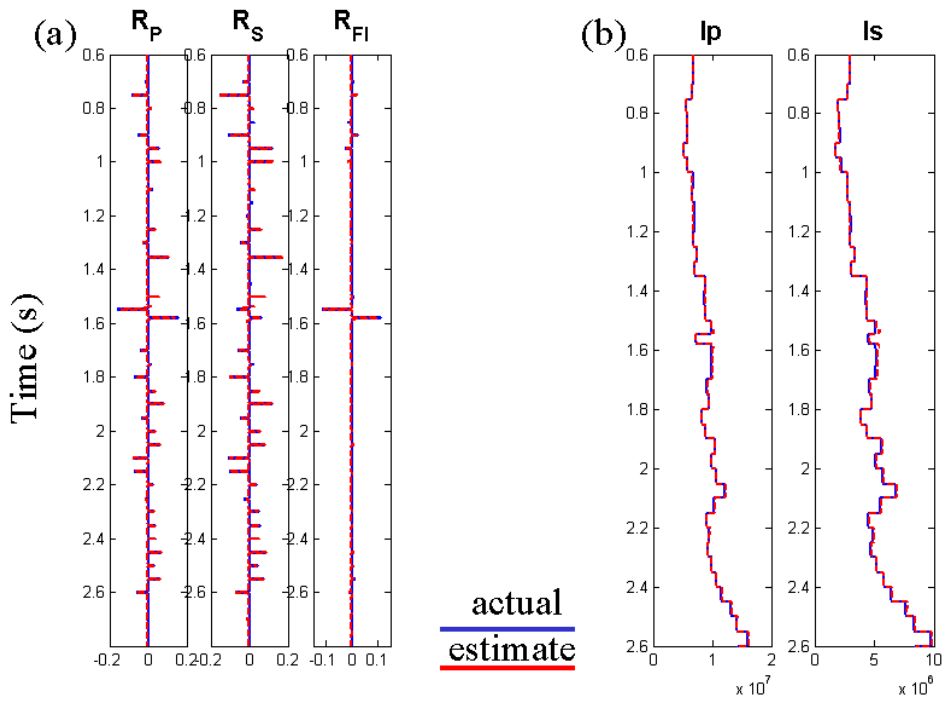

Figure 6.25: The estimate (red) of the full spectrum P-wave impedance, S-wave impedance and fluid factor reflectivity (a) compared to the ideal (blue). After integrating, the P-wave and S-wave impedances (b) estimate (red) are compared to the ideal impedance (blue).

6.7 Discussion

I was unable to obtain acceptable impedance inversion results on real seismic data. The spike series estimated by the inversion were laterally discontinuous and geologically implausible. In order to apply this algorithm to real data, I believe that lateral constraints need to be incorporated. I originally developed this algorithm using sparse matrices, and when implementing the lateral constraints ran into memory issues. In the future I plan to implement the lateral constraints following the second conjugate gradient methodology discussed in Section 6.2.3.

Three different *a priori* distributions were developed to constrain the problem. I preferred the Cauchy distribution since the results seemed to be less sensitive to its parameterization. The ℓ^1 and Huber norm results seemed to be quite sensitive to the choice of the threshold parameter ϵ.

Buland and Omre (2003) use an interesting constraint that I plan to implement in the future. They assume that the elastic parameters for different layers are correlated using a covariance function to describe their inter-relationship. This assumption, is probably more realistic than the assumption made in this chapter, that the variables from different interfaces are independent. This would help address the null space introduced by the band-limited nature of the source wavelet. However, this would also make the eigenvector analysis in Section 6.2.1 much more expensive.

In the preceding analysis it was assumed that statics and stacking velocities are known exactly *a priori*. Residual velocity errors cause distortions in the AVO parameter estimates particularly the gradient. Swan (2001) makes use of the fact that the ratio between the gradient and the intercept is distorted to identify and correct for residual moveout. In a similar fashion, this algorithm could be modified to identify velocity errors and correct for them. A less costly approach would be to run Swan's residual velocity analysis and correction prior to the AVO waveform inversion.

Residual statics also cause distortions in estimates of the AVO parameters. To help address this issue, treat the data generated by each outer loop of the inverse solution as a pilot model for trim statics. The data traces can then be trimmed into this model improving the overall fit. The advantage of using this model over some pilot generated from the stack is that offset dependent changes are included. In addition, this approach to could be applied to seismic data prior to NMO. These ideas are developed in Chapter 8.

An uncertainty analysis can be performed in a manner similar to Chapters 2 and

3. The cost of doing this is quite expensive in that it requires the explicit calculation of the inverse of the matrix $\left[\mathbf{G}'^{T}\mathbf{G}'+\Theta\mathbf{Q}\right]$. Further, this calculation will be overly pessimistic since the truncated conjugate gradient solution excludes some of the small eigenvalues.

6.8 Conclusions

The AVO waveform inversion developed in this chapter generates better AVO reflectivity attribute estimates on events undergoing NMO stretch and differential tuning than traditional AVO inversion performed on NMO corrected gathers. This is clearly demonstrated by the first synthetic example. The inversion is relatively insensitive to the waveform used, provided it is reasonably close. Further, the estimates are more reliable than the estimates provided by the traditional AVO analysis performed on a sample-by-sample basis. This increased reliability is due to the classic trade-off between resolution and stability. A few sparse reflectivity values are estimated with greater certainty than the dense reflectivity at every time sample as in the traditional AVO analysis.

The algorithm is reasonably fast. Even though the size of the problem is much larger than the traditional AVO problem, only a few iterations of conjugate gradient are required to come to a satisfactory solution. For the real data example, the algorithm took eight times as long as the traditional way of solving the problem.

In this chapter, the AVO model used was based on a two-term approximation. This is not particularly satisfactory in that the original motivation was to be able to invert large offsets in order to do three-term AVO analysis. Limiting the solution to two terms was done initially for stability reasons. In Chapter 8 the theory is extended to estimate three independent variables.

Chapter 7

Stretch-free NMO

7.1 Introduction

An alternative approach to deal with NMO stretch is to perform stretch-free NMO (Hicks, 2001, Trickett, 2003, Downton et al., 2003) on the data prior to AVO. There are a number of advantages to stretch-free NMO including the ability to include more offsets and obtain higher frequency content in the resulting stack (Hicks, 2001). For marine data, because of the size of the prestack data, it is a common practise to perform AVO analysis on limited offset stacks rather than prestack gathers. Connolly (1999) provides a framework to interpret these stacks. Smith and Gidlow (2003) point out that a far offset stack created with a suitable range of offsets is similar to a fluid stack. However, the far offset stack has significantly lower frequency content, compared to the near offset stack, making the analysis more complicated. It is thought that by doing stretch-free NMO that this issue can be addressed.

There have been a number of papers in the literature suggesting how to implement stretch-free NMO. Trickett (2003) proposes doing stretch-free NMO by combining the ideas of inverse NMO stack (Claerbout, 1992) with block NMO (Rupert and Chun, 1975). In actuality the output of this approach is a stretch-free stack. Prestack gathers can be mimicked by creating a series of limited offset stacks. Trickett (2003) never proves or shows that his approach is amplitude preserving However, Hunt et al. (2003) show an AVO case study in which this approach was used to do the AVO inversions. Differences are noted between the stretch-free result and the traditional NMO processed AVO inversion, but no conclusions are made as to which is superior.

Rupert and Chun (1975) suggested performing block NMO. Instead of applying NMO, the basic idea is to apply time shifts to isolated reflectors thus avoiding NMO

stretch. The issue with this approach is that the method cannot handle differential tuning as a function of offset or conflicting events. To get around this, Hicks (2001) suggested doing the block NMO in the Radon domain. He pointed out that conflicting, constant amplitude hyperbolic events are cleanly separated by the Hyperbolic Radon transform. However, due to the limited aperture and the band-limited nature of the seismic data, the constant amplitude hyperbolic events have an impulse response rather than a point response. The impulse responses from adjacent events interfere with each other complicating the analysis. Hicks (2001) solves this problem by identifying key events prior to the transform. Further, in order to allow for AVO and efficiently implement the algorithm, Hicks implemented the stretch-free NMO using the Parabolic Fourier Radon Transform (PFRT). Because isolated events have to be identified prior to the transform, the method is cumbersome. Moreover, if a reflector lies in between two of these identified events, it is distorted by the transform.

All these methods suffer from the fact that the data are band-limited. In the previous chapter, constraints and a solution method were developed to address this issue. In this chapter the AVO waveform inversion, derived in the previous chapter, is modified slightly to create a stretch-free NMO stack (Downton et al., 2003). The theory for this is first developed and then demonstrated on both synthetic and real data examples. The results of AVO inversions using traditional NMO, stretch-free NMO, and the AVO waveform inversion are then compared for synthetic data. Provided with suitable constraints, data preprocessed with stretch-free NMO can yield more accurate AVO inversion estimates than data preprocessed in the traditional manner. However, the results of this modeling study show that the estimates are not as good as that achieved by doing the AVO waveform inversion.

7.2 Theory

7.2.1 NMO inversion

Claerbout (1992) points out that NMO correction is a conjugate operation rather than an inverse operation. By using the conjugate instead of the inverse operator, amplitude and character distortions arise, namely NMO stretch. The conjugate operator is used as an easy way to deal with the ill-conditioned or underdetermined nature of the problem. Claerbout (1992) illustrates that for full bandwidth data, it is possible to perform inverse NMO using conjugate gradient (Figure 7.1). He uses the

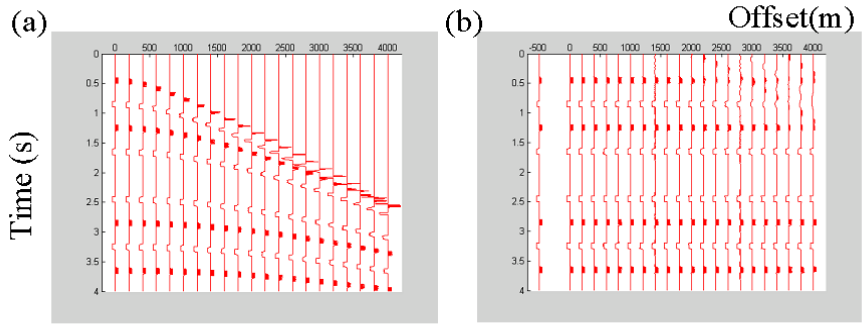

Figure 7.1: Comparison of inverse NMO (b) with input (a) for full spectrum data. The trace shown in (b) at offset= -500 is the stacked response. The stacked response is equivalent to the input.

number of iterations as the trade-off parameter to address the ill-conditioned nature of the problem. However, this solution technique does not work on band-limited data (Figure 7.2) since the problem is underdetermined, and as a consequence, additional constraints are required.

7.2.2 High-resolution NMO stack

Smith and Gidlow (1987) showed that AVO inversion is a weighted stacked. The result of this stacking process is two or three reflectivity estimates depending on the approximation used. The traditional stack can be considered a subset of this. In this case, the AVO model is a constant amplitude model which can be described by one reflectivity parameter, the stacked response $\mathbf{r}_0$. Under these assumptions the AVO waveform formulation (equation 6.8) may modified to form the NMO stack formulation

$$\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_m \end{bmatrix} \approx \begin{bmatrix} \mathbf{W}\mathbf{N}_1\mathbf{S}_1 \\ \vdots \\ \mathbf{W}\mathbf{N}_m\mathbf{S}_m \end{bmatrix} \mathbf{r}_0, \tag{7.1}$$

where $\mathbf{S}_k$ is the spray operator (Claerbout, 1992), $\mathbf{N}_k$ is the NMO operator (Claerbout, 1992), $\mathbf{W}$ is a convolution matrix representing the source wavelet and $\mathbf{d}_k$ is the data at offset k. The spray operator acts in a similar fashion to diagonal matrices $\mathbf{E}_k$ and $\mathbf{F}_k$ in Chapter 6, but are constant amplitude since AVO is not considered in this chapter. Because of the band-limited nature of the seismic data the inversion

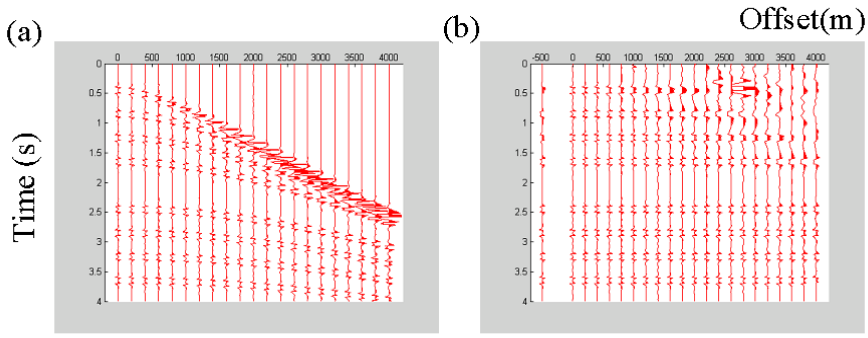

Figure 7.2: Comparison of inverse NMO (b) with input (a) for band-limited data. The trace shown in (b) at offset= -500 is the stacked response. Notice the waveform changes as a function of time.

of equation (7.1) is underdetermined. Constraints may be added in a similar fashion as previously done in the AVO waveform inversion, leading once again to an inverse problem of the form

$$\left[\mathbf{G}^T\mathbf{G}+\frac{\sigma_n^2}{\sigma_0^2}\mathbf{Q}\right]\mathbf{r}_o = \mathbf{G}^T\mathbf{d}, \tag{7.2}$$

where $\mathbf{G}$ in this case represents the linear operator specified by equation (7.1), σ_n^2 is the variance of the noise and $\mathbf{Q}$ is a diagonal matrix specifying the weights. Assuming the reflectivity follows a Cauchy distribution the diagonal elements of $\mathbf{Q}$ are

$$Q_{i,i} = \frac{2}{\left(\frac{r_i^2}{2\sigma_0^2}+1\right)}, \tag{7.3}$$

where σ_0^2 is the variance of the reflectivity. Similar to Chapter 6 other long-tailed distributions may be used to create the weighting function. The matrix $\mathbf{Q}$ is constructed in a bootstrap fashion. Essentially large reflectivities get weighted with small weights and small reflectivities with large weights. Once again the inverse problem is solved in two loops. In the inner loop equation (7.2) is solved using conjugate gradient getting a new estimate of $\mathbf{r}_o$. The number of iterations is used as a regularization parameter. The outer loop consists of updating the $\mathbf{Q}$ matrix from the new estimate of $\mathbf{r}_0$ and then rerunning the inner loop. This leads to a sparse spike reflectivity solution.

Stretch-free NMO

To perform AVO analysis, prestack data are required. Equation (7.1) may be performed on a subset of offsets thus creating partial offset stacks. In the extreme, one offset may be used. In this case, equation (7.1) becomes

$$\mathbf{d}_k = \mathbf{W}\mathbf{N}_k\mathbf{r}_k, \tag{7.4}$$

where $\mathbf{r}_k$ represents the NMO corrected data for offset k. Similar to the above, the inverse problem is

$$\left[\mathbf{N}_k^T\mathbf{W}^T\mathbf{W}\mathbf{N}_k + \frac{\sigma_n^2}{\sigma_0^2}\mathbf{Q}\right]\mathbf{r}_k = \mathbf{N}_k^T\mathbf{W}^T\mathbf{d}, \tag{7.5}$$

where $\mathbf{Q}$ is defined in a similar fashion as equation (7.3). The matrix $\mathbf{Q}$ may be calculated for each offset or for all offsets using the $\mathbf{Q}$ matrix formed for the high-resolution NMO stack. In practice, it was found that the latter gave better results.

In solving the inverse of equation (7.4), deconvolution is first performed, resulting in a sparse spike reflectivity series. Removing the influence of the wavelet addresses the fundamental underdetermined nature of the problem by making the data broad-band. The inverse NMO operator can then be successfully applied in a similar fashion as section 7.2.1**.** The reliability of the results are dependent on the applicability of the constraints or weights used to make the problem unique. For display purposes the NMO corrected data shown in the following examples is filtered with the source wavelet so it is band-limited.

7.3 Examples

7.3.1 Synthetic example

The synthetic example from Section 6.3 is used illustrate the stretch-free NMO algorithm. Recall that the AVO behavior of most of the reflectors in the synthetic gather followed the mudrock trend, but several Class III and IV anomalies were present. Both isolated and tuned anomalies were included. These anomalies should introduce distortions into the AVO inversion. Figure 7.3 shows the synthetic data after the application of stretch-free NMO compared to NMO corrected data and a synthetic created without NMO. The stretch-free NMO is higher frequency on the far offsets compared to the NMO corrected data, though it is also noisier as a result

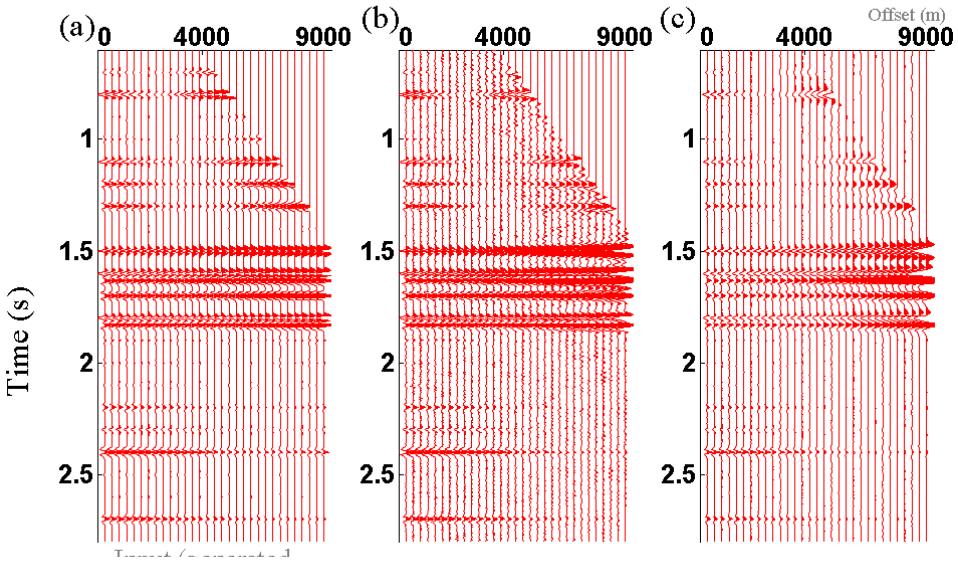

Figure 7.3: Comparison of NMO corrected data (c) with stretch-free NMO (b). For reference purposes gather (a) was generated without NMO.

of the deconvolution process. The two tuned reflectors at 1.6 and 1.8 seconds show greater detail and more information than the NMO corrected gathers but not as much information as the ideal synthetic gather. Figure 7.4 shows the difference between the ideal and stretch-free NMO gather. Note that all the differences occur for angles larger than 40°. The stretch-free NMO difference is better than the NMO corrected gather difference (Figure 6.3).

The stretch-free NMO corrected data was then inverted for intercept A and gradient B in a manner similar as in Section 6.3, the results of which are shown in Figure 7.5. The estimated intercept once again accurately matches the ideal intercept reflectivity. The estimated gradient is now a better match to the ideal than that done on the NMO corrected data (Figure 6.4). However, there is still significant scatter in the crossplot domain (Figure 7.5c) compared to the results from the AVO waveform inversion (Figure 6.6c). The reflectivity results are summarized in Figure 7.6. All techniques give good estimates of the intercepts. However, in the zone with Class III and IV reflectors, the NMO corrected gathers show significant error on the gradient estimate. The data which has been preprocessed with the stretch-free NMO shows less error, while the reflectivity estimate from the AVO waveform inversion shows practically no error.

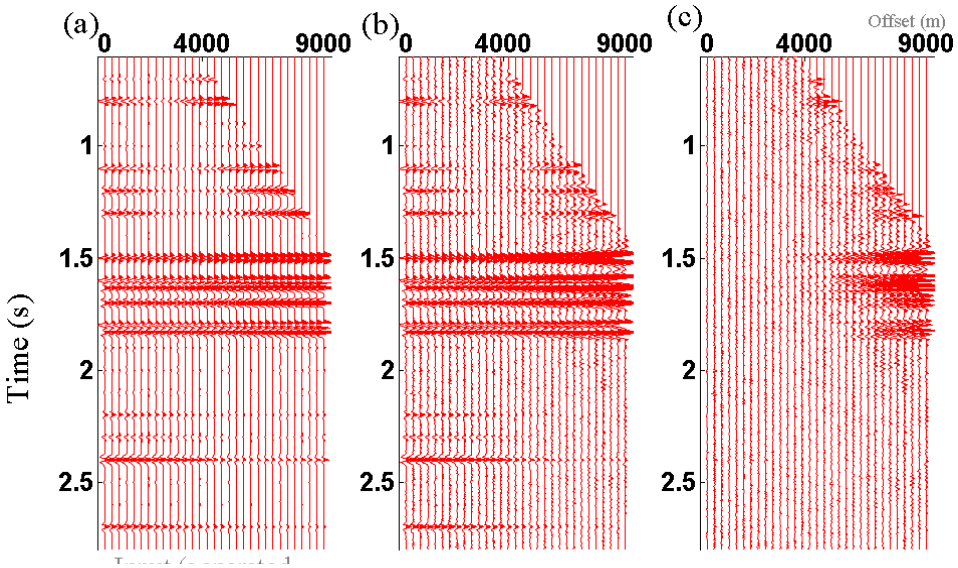

Figure 7.4: Comparison of stretch-free NMO corrected data (b) versus ideal (a), and the scaled (5X) difference (c).

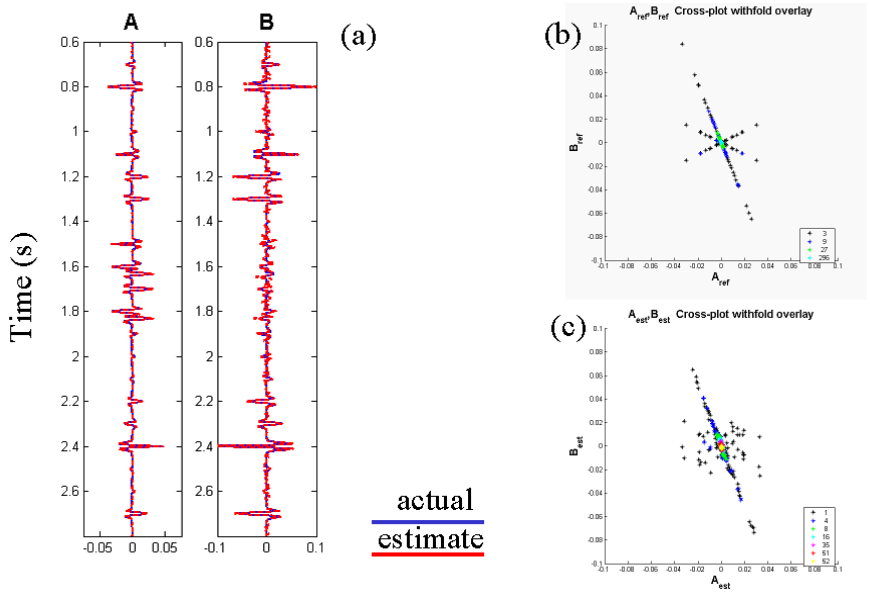

Figure 7.5: The AVO estimates (a) for intercept A and gradient B (red) are compared to the ideal results (blue) for gathers preprocessed with stretch-free NMO applied. The ideal data (b) is compared to the estimated data (c) in the cross-plot domain.

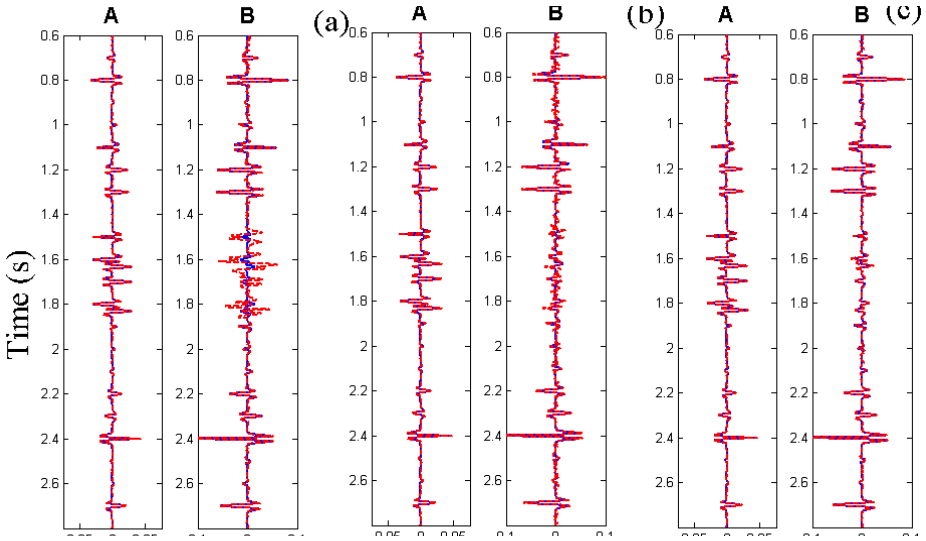

Figure 7.6: Comparison of AVO inversion results using traditional NMO as input (a), stretch-free NMO as input (b) and AVO waveform inversion (c).

7.3.2 Seismic data example

The stretch-free NMO algorithm was applied to marine data acquired off the East Coast of Canada. Figure 7.7 shows a representative prestack CDP gather. Note that the stretch-free NMO shows greater frequency content in the far offsets. For example, the event at 5.5 seconds shows higher frequency content and slightly different kinematics. If stretch-free NMO is to be applied, the velocity analysis needs to be modified as well. Also, note that along with the higher frequency content comes more random noise. The far angle stack (40 ° - 50 °) generated using the stretch-free NMO gathers is shown in Figure 7.9. For comparison, the angle stack was also generated using NMO corrected gathers (Figure 7.8). Note the improvement in frequency as a result of using the stretch-free NMO corrected gathers.

7.4 Conclusions

Stretch-free NMO may be derived from the more general AVO waveform formulation presented in Chapter 6. The output of this is either a stacked section or prestack gathers. The algorithm is much more computationally intensive than traditional NMO, with the cost being a order of magnitude less than that of running a hyperbolic Radon transform. High-resolution constraints are used to stabilize the

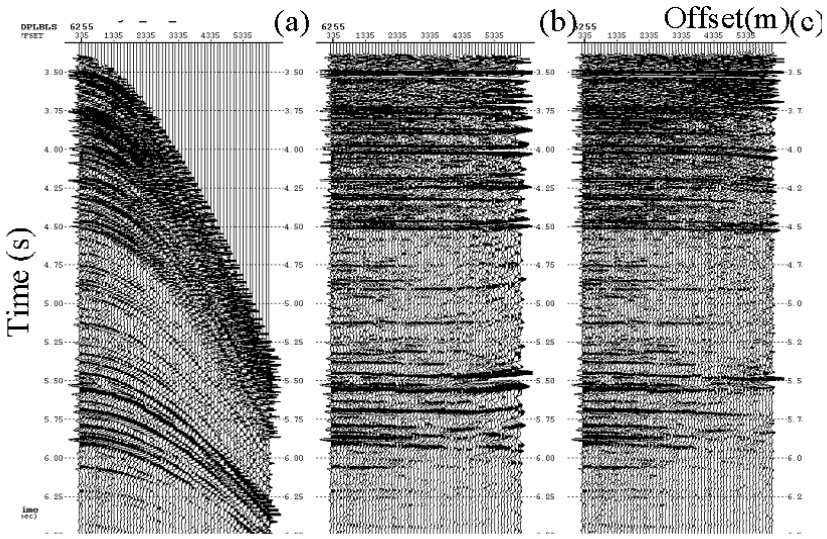

Figure 7.7: NMO corrected CDP gather (b) and stretch-free NMO (c) compared to input (a).

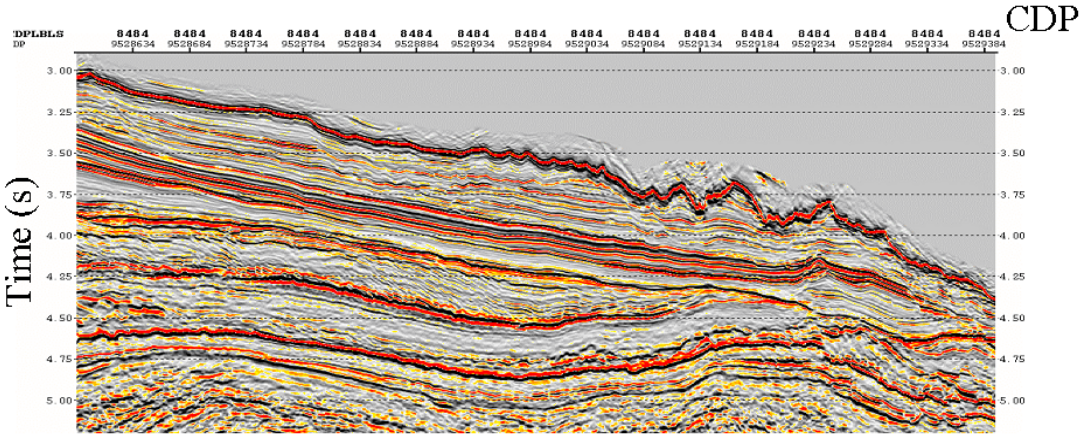

Figure 7.8: Angle stack (40 - 50 degrees) based on traditionally NMO corrected gathers

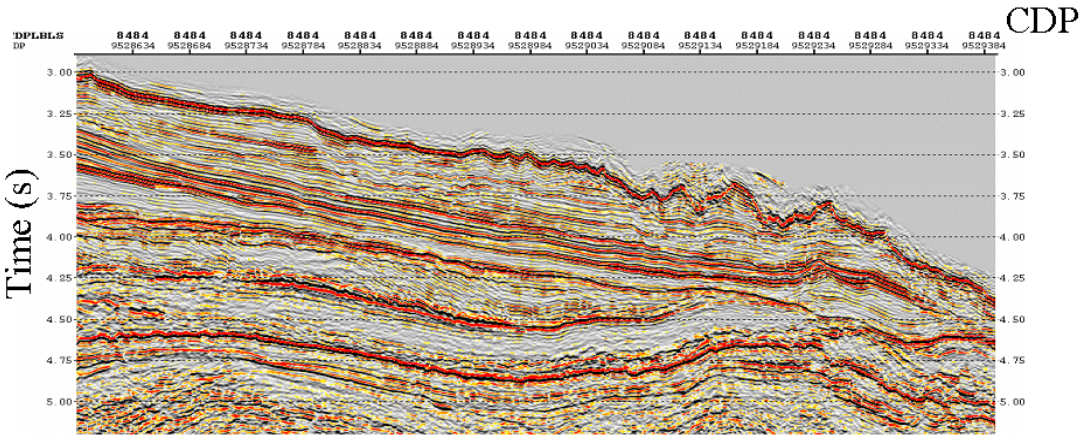

Figure 7.9: Angle stack (40 - 50 degrees) based on stretch-free NMO corrected gathers

inversion. If stretch-free NMO is to be applied, the stacking velocities must be picked with this approach for consistency.

On synthetic data, the AVO inversion results with stretch-free NMO input are superior to that processed with traditional NMO. However, the results were not as good as those obtained for the AVO waveform inversion. One might consider performing stretch-free NMO if limited offset stacks are desired as an output. This should lead to greater waveform consistency as a function of offset.

Chapter 8

Three-term AVO waveform inversion

8.1 Introduction

Density reflectivity is a useful AVO attribute to infer fluid saturation. In Chapter 3, a three-term nonlinear AVO inversion was developed to estimate the density reflectivity. It was shown that to obtain reliable results, large offsets and angles are needed. This being the case, NMO stretch and tuning become an issue. In Chapter 6, a two-term AVO waveform inversion was developed that addressed the issues of NMO stretch and tuning as part of the inversion. For stability reasons, the algorithm was developed using only two terms. This chapter extends the previously developed two-term AVO waveform inversion to three terms, incorporating the constraints developed in Chapter 3. By including the waveform and the NMO operator, NMO stretch and offset dependent tuning can be modeled as part of the inverse problem leading to more accurate estimates of the reflectivity.

In addition, several other theoretical innovations are introduced in this chapter. In Chapter 6, the input to the inversion needed to have NMO. In this chapter, modifications are introduced that allow the input seismic data to be either NMO corrected or uncorrected. Further, to allow for better estimates in the presence of non-Gaussian noise, long tailed likelihood probability distributions are investigated. Finally, a method for correcting residual statics as part of the AVO waveform inversion is developed.

In previous chapters, the noise was assumed to be uniform, independent and Gaussian. Real seismic data often has noise outliers that violate the Gaussian assumption leading to errors in the estimates. For example, Downton and Mackidd (1997) show that in the presence of spurious noise, such as multiples, the ℓ^1 mis-

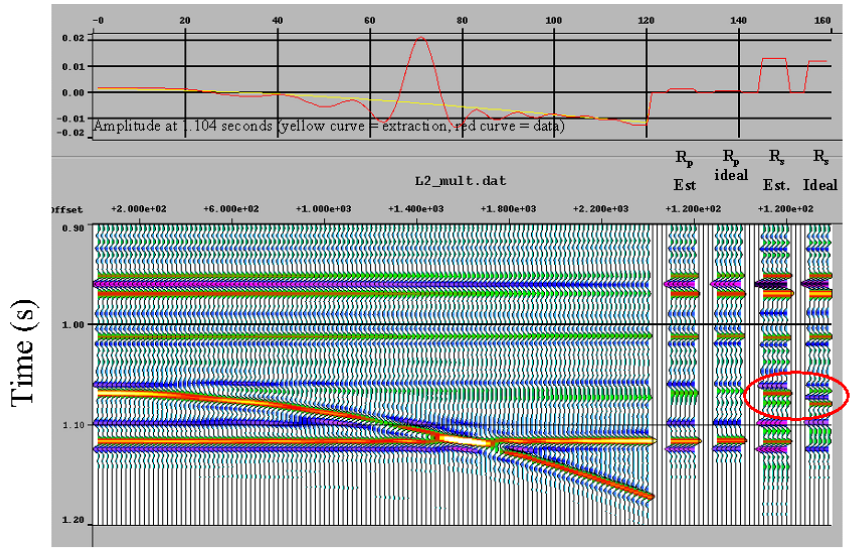

Figure 8.1: P- and S-wave impedance reflectivity estimate versus ideal based on L2 AVO inversion in the presence of non-Gaussian noise. Note that the S-wave impedance reflectivity estimate is biased (highlighted by red ellipse) by the multiple.

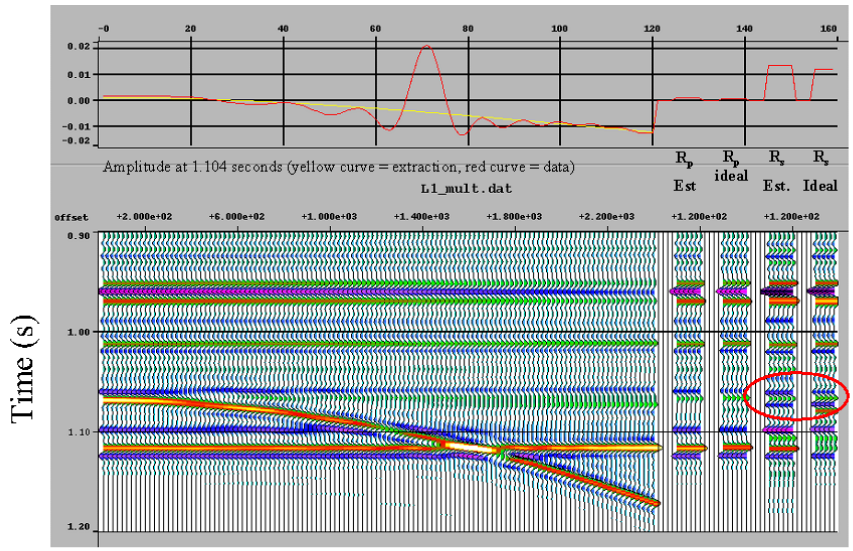

Figure 8.2: P- and S-wave impedance reflectivity estimate versus ideal based on L1 AVO inversion in the presence of non-Gaussian noise. Note the improvement in the S-wave impedance reflectivity estimate compared to Fig. 8.1. (highlighted by red ellipse).

fit function gives better estimates than the ℓ^2 misfit function (Figure 8.1 and 8.2). Downton and Mackidd (1997) use linear programming (Dantzig, 1963) to implement ℓ^1 misfit function. This is too cost prohibitive for the AVO waveform inversion problem. Gersztenkorn el al. (1986) showed that the ℓ^1 misfit function could be implemented using iteratively reweighted least-squares. Scales and Gersztenkorn (1988) showed that combining this misfit function along with conjugate gradient, leads to a computationally efficient implementation. This is the approach taken in this chapter. Various other misfit functions (Crase, 1990: Bube and Langan, 1997) are also investigated.

In the first part of this chapter, the relevant theory is developed using a Bayesian formalism. The likelihood function is developed assuming long-tailed statistics and extends the AVO waveform inversion of Downton and Lines (2003) to three terms. A modification is introduced allowing NMO corrected seismic data to be used as the input to the inversion. Because of the band-limited nature of the seismic data, the inverse problem is underdetermined necessitating the use of constraints. Well log information is introduced to establish the relationship between the different parameters. Based on this information, a change of variables is performed so that the transformed parameters are statistically independent. After the change of variables, the problem is still underdetermined requiring regularization. Similar to Section 6.2.2, this is addressed by creating a weighting function that treats certain reflection coefficients as being more probable than others. Theoretically, this is accomplished by choosing a long tailed *a priori* distribution. This leads to a nonlinear inversion which I solve using conjugate gradient. The problem is also regularized by only allowing the conjugate gradient solution to perform a limited number of iterations. This excludes the smallest eigenvalues from the solution.

The algorithm is demonstrated on both synthetic and real seismic data. The synthetic example is constructed so as to include NMO stretch and offset dependent tuning, effects that are known to bias the AVO inversion estimates. The three-term AVO waveform inversion produces good estimates of all the reflectivity attributes, including the density. The algorithm accurately estimates the reflectivity even on events exhibiting NMO stretch and differential tuning. The algorithm is then demonstrated on a more realistic synthetic based on a well log from western Canada. Non-Gaussian noise is added to the synthetic, to demonstrate the enhanced ability of the alternative misfit functions to estimate the reflectivity. Various misfit functions are

demonstrated on this example. The AVO waveform inversion is also demonstrated on seismic data from this area. Various misfit weights are examined, comparing their relative performance. Good results were obtained by all, but the Cauchy misfit function seems superior, since it gave good results while being relatively insensitive to its parameterization. The seismic data example demonstrates how the algorithm successfully differentiates a known density anomaly between two well locations and shows superior results to those of Chapter 3.

8.2 Theory

8.2.1 Convolutional model

The convolutional model is used as the basis for the likelihood model. This model assumes the earth is composed of a series of flat, homogeneous, isotropic layers. Ray tracing is performed to map the offset to angle of incidence. Transmission losses, converted waves, and multiples are not incorporated in this model and so must be addressed through prior processing. In theory, gain corrections such as divergence, absorption, directivity, and array corrections can be incorporated into this model, but are not considered for brevity and simplicity, so must be previously applied in the processing. Any linear approximation of the Zoeppritz equations may be used as the starting point for this derivation.

Following the same line of reasoning as in Chapter 6, a three-term AVO waveform model is

$$\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \end{bmatrix} = \begin{bmatrix} \mathbf{W}\mathbf{N}_1\mathbf{E}_1 & \mathbf{W}\mathbf{N}_1\mathbf{F}_1 & \mathbf{W}\mathbf{N}_1\mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{W}\mathbf{N}_K\mathbf{E}_K & \mathbf{W}\mathbf{N}_K\mathbf{F}_K & \mathbf{W}\mathbf{N}_K\mathbf{H}_K \end{bmatrix} \begin{bmatrix} \mathbf{r}_\alpha \\ \mathbf{r}'_\beta \\ \mathbf{r}_d \end{bmatrix}, \tag{8.1}$$

where $\mathbf{r}_\alpha$, $\mathbf{r}'_\beta$, $\mathbf{r}_d$ are the P- wave velocity reflectivity, S-wave velocity reflectivity scaled by γ, and density reflectivity respectively. These are all vectors whose elements correspond to different time samples. Likewise the elements of the data vector $\mathbf{d}_k$ represent the processed seismic data for the k^{th} offset for the corresponding time samples. There are a total of K offsets and L time samples. The block matrices describe the physics of the problem. The matrices $\mathbf{E}_k$, $\mathbf{F}_k$, and $\mathbf{H}_k$ are diagonal and contain weights that describe how the amplitude changes as a function of offset. These weights follow from equation (1.6). Following Claerbout (1992), the block matrix $\mathbf{N}_k$ performs NMO. This operator can be constructed using whatever offset traveltime

relationship one desires. In this chapter, the three-term correction following Castle (1994) is employed. Implicit in this model is that the kinematics are previously defined. Lastly, $\mathbf{W}$ is a convolution matrix that contains the source wavelet. Applying these three operators in series, the block matrices $\mathbf{E}_k$, $\mathbf{F}_k$, and $\mathbf{H}_k$ model the offset dependent reflectivity from the zero offset reflectivity, $\mathbf{N}_k$ applies NMO and $\mathbf{W}$ convolves the offset dependent reflectivity with the source wavelet modeling the band-limited seismic data with NMO. The inversion of equation (8.1) can be thought of as three separate inversion problems in series, deconvolution, inverse NMO and AVO inversion.

NMO corrected data

Data that have been previously NMO corrected may be input to this algorithm with a slight modification. As discussed in Section 5.2, the NMO correction is typically applied with the conjugate operator. Thus, NMO processing may be simulated by applying the conjugate NMO operator to both the left and right hand sides of equation (8.1) resulting in

$$\begin{bmatrix} \mathbf{d}'_1 \\ \vdots \\ \mathbf{d}'_K \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1^T\mathbf{W}\mathbf{N}_1\mathbf{E}_1 & \mathbf{N}_1^T\mathbf{W}\mathbf{N}_1\mathbf{F}_1 & \mathbf{N}_1^T\mathbf{W}\mathbf{N}_1\mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{N}_K^T\mathbf{W}\mathbf{N}_K\mathbf{E}_K & \mathbf{N}_K^T\mathbf{W}\mathbf{N}_K\mathbf{F}_K & \mathbf{N}_K^T\mathbf{W}\mathbf{N}_K\mathbf{H}_K \end{bmatrix} \begin{bmatrix} \mathbf{r}_\alpha \\ \mathbf{r}'_\beta \\ \mathbf{r}_d \end{bmatrix}, \tag{8.2}$$

where $\mathbf{d}'_k$ is the NMO corrected data for the k^{th} offset.

$$\begin{bmatrix} \mathbf{d}'_1 \\ \vdots \\ \mathbf{d}'_K \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1^T\mathbf{d}_1 \\ \vdots \\ \mathbf{N}_K^T\mathbf{d}_K \end{bmatrix}. \tag{8.3}$$

Note that $\mathbf{N}_k^T\mathbf{W}\mathbf{N}_k$ is not an identity matrix. This operator is responsible for NMO stretch and offset dependent tuning. By applying the inverse of this operator these artifacts can be removed.

Mask operator

Equation (8.1) is the three-term equivalent to the two-term equation (6.8). In Section 6.2.1 various masking or muting operators are applied both to the model and

to the seismic data, so as to limit the range of angles used in the AVO inversion, and to exclude supercritical angles and first breaks from the calculation. This resulted in equation (6.9). In a similar fashion, equation (8.1) is modified

$$\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_K \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 & \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K & \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{H}_K \end{bmatrix} \begin{bmatrix} \mathbf{r}_\alpha \\ \mathbf{r}'_\beta \\ \mathbf{r}_d \end{bmatrix}, \tag{8.4}$$

where $\mathbf{M}_k^f$ masks supercritical angles. This mask is calculated on data without moveout. The other mask $\mathbf{M}_k^n$ is calculated on data with NMO. The matrix $\mathbf{M}_k^n$ masks hard zeros in the data so as to exclude data that has previously been muted in the processing. In addition, $\mathbf{M}_k^n$ masks data that falls outside some user defined range of angles. This allows the user to exclude certain data, maybe on the basis of noise, from the solution.

In a similar fashion the NMO corrected equation (8.2) is modified to include masks resulting in

$$\begin{bmatrix} \mathbf{d}'_1 \\ \vdots \\ \mathbf{d}'_K \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 & \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K & \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{H}_K \end{bmatrix} \begin{bmatrix} \mathbf{r}_\alpha \\ \mathbf{r}'_\beta \\ \mathbf{r}_d \end{bmatrix}. \tag{8.5}$$

In this case, the masking operator $\mathbf{M}_k^n$ is designed on NMO corrected data.

Depending on the processing performed to the seismic data, either equation (8.4) or (8.5) may be inverted with similar results. For future reference and simplicity, the linear model (equation 8.4 or 8.5 as appropriate) is written as

$$\mathbf{d} = \mathbf{G}\mathbf{x}, \tag{8.6}$$

where $\mathbf{G}$ is the linear operator

$$\mathbf{G} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 & \mathbf{M}_1^n \mathbf{W}\mathbf{N}_1 \mathbf{M}_1^f \mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K & \mathbf{M}_K^n \mathbf{W}\mathbf{N}_K \mathbf{M}_K^f \mathbf{H}_K \end{bmatrix}, \tag{8.7}$$

in the case of equation (8.4) or

$$\mathbf{G} = \begin{bmatrix} \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{E}_1 & \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{F}_1 & \mathbf{M}_1^n \mathbf{N}_1^T \mathbf{W} \mathbf{N}_1 \mathbf{M}_1^f \mathbf{H}_1 \\ \vdots & \vdots & \vdots \\ \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{E}_K & \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{F}_K & \mathbf{M}_K^n \mathbf{N}_K^T \mathbf{W} \mathbf{N}_K \mathbf{M}_K^f \mathbf{H}_K \end{bmatrix}, \tag{8.8}$$

in the case of equation (8.5). The parameter $\mathbf{x}$ is the unknown reflectivity vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{r}_\alpha \\ \mathbf{r}'_\beta \\ \mathbf{r}_d \end{bmatrix}, \tag{8.9}$$

and $\mathbf{d}$ is the muted seismic data before or after NMO correction as appropriate. The data vector $\mathbf{d}$ contains $M = L \times K$ samples while the parameter vector $\mathbf{x}$ contains $N = 3L$ elements.

8.2.2 Parameterization

The regularization that is employed requires that the elements of $\mathbf{x}$ are statistically independent. However, this is clearly not the case since the mudrock relationship (Castagna et al. 1985) can be used to predict P-wave velocity reflectivity from the scaled S-wave reflectivity (equation 3.20). Further, the Gardner relationship (equation 3.32) may be used to predict the density reflectivity from the P-wave velocity reflectivity. In Chapter 3, it was shown, assuming Gaussian statistics, that these correlations can be described by the 3×3 covariance matrix (equation 3.39). Writing this covariance matrix in terms of the scaled S-wave reflectivity $\mathbf{r}'_\beta$ rather than the S-wave reflectivity results in

$$\mathbf{C}_\mathbf{x} = \begin{bmatrix} \sigma^2_{R_\alpha} & \sigma_{R_\alpha R'_\beta} & \sigma_{R_\alpha R_d} \\ \sigma_{R_\alpha R'_\beta} & \sigma^2_{R'_\beta} & \sigma_{R'_\beta R_d} \\ \sigma_{R_\alpha R_d} & \sigma_{R'_\beta R_d} & \sigma^2_{R_d} \end{bmatrix} = \sigma^2_{R_\alpha} \begin{bmatrix} 1 & \frac{r^2_{R_\alpha R'_\beta}}{m} & g \\ \frac{r^2_{R_\alpha R'_\beta}}{m} & \left(\frac{r_{R_\alpha R'_\beta}}{m}\right)^2 & f\left(\frac{r_{R_\alpha R'_\beta}}{m}\right)^2 \\ g & f\left(\frac{r_{R_\alpha R'_\beta}}{m}\right)^2 & \left(\frac{g}{r_{R_\alpha R_d}}\right)^2 \end{bmatrix}, \tag{8.10}$$

where m is the mudrock slope with correlation coefficient $r_{R_\alpha R_\beta}$, g is the Gardner coefficient with correlation coefficient $r_{R_\alpha R_\rho}$ and f is the Gardner coefficient between the S-wave velocity and density. If it is assumed, as is typically done in deconvo-

lution, that the reflectivity attributes are independent and that P-wave impedance reflectivity is stationary then it is simple to extend this covariance matrix to L time samples resulting in a $3L \times 3L$ sparse covariance matrix $\mathbf{C_x}$. This matrix describes the correlations between the different variables.

Similar to Chapter 6, the eigenvectors $\mathbf{V}$ and eigenvalues $\mathbf{\Sigma}$ can be calculated for the covariance matrix

$$\mathbf{C}_x = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^T. \tag{8.11}$$

The eigenvectors are used to transform the parameters

$$\mathbf{x}' = \mathbf{V}^{-1}\mathbf{x}, \tag{8.12}$$

so that the resulting parameter variables $\mathbf{x}'$ are independent. Under the change of variables, the covariance matrix for the transformed variables is

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}. \tag{8.13}$$

For each time sample there are three distinct eigenvalues $\sigma_{1_n}^2$, $\sigma_{2_n}^2$, and $\sigma_{3_n}^2$ corresponding to the variance of the three transformed variables. After the transform the new variables can be approximately thought of as (1) the average of the P-wave and S-wave impedance reflectivity (the variable associated with the largest eigenvalue), (2) similar to the fluid stack (variable associated with the second largest eigenvalue) and (3) the deviation between the density predicted by the Gardner equation and the actual density reflectivity (variable associated with the smallest eigenvalue). Once again, it is convenient for future use to rewrite equation (8.13) removing the P-wave impedance reflectivity variance $\sigma_{r_p}^2$

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} = \sigma_{r_p}^2 \begin{bmatrix} \tilde{\sigma}_1^2 & 0 & 0 \\ 0 & \tilde{\sigma}_2^2 & 0 \\ 0 & 0 & \tilde{\sigma}_3^2 \end{bmatrix}. \tag{8.14}$$

Under the change of variables equation (8.6) becomes

$$\mathbf{d} = \mathbf{G}'\mathbf{x}', \tag{8.15}$$

where

$$\mathbf{G}' = \mathbf{GV}. \tag{8.16}$$

8.2.3 Robust likelihood function

Following Gersztenkorn et al. (1986) the probability function

$$P\left(d_m|\mathbf{x}', \xi, p, I\right) = \frac{p}{2\xi\Gamma\left(\frac{1}{p}\right)} \exp\left(-\frac{\left|F_m\left(\mathbf{x}'\right) - d_m\right|}{\xi}\right)^p, \tag{8.17}$$

is used to describe the likelihood function for the ℓ^p norm. In equation (8.17) $F_m\left(\mathbf{x}\right)$ describes the functional relationship between the m^{th} experimental data measure d_m and parameters $\mathbf{x}$, $\Gamma\left(\right)$ is the gamma function, $\xi\left(>0\right)$ is a scaling parameter and $p\left(>0\right)$ is a shaping parameter. This function leads to the ℓ^p misfit function that includes both the exponential and Gaussian functions. Choosing $1 \leq p < 2$ leads to a robust estimator. The parameters $p = 2$ and $\xi^2 = 2\sigma^2$ results in the Gaussian probability function

$$P_{\ell^2}\left(d_m|\mathbf{x}, \sigma, I\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{\left|F_m\left(\mathbf{x}'\right) - d_m\right|^2}{2\sigma^2}\right)}, \tag{8.18}$$

that leads to the ℓ^2 misfit function, while the parameters $p = 1$ and $\xi = \sigma$ results in exponential distribution

$$P_{\ell^1}\left(d_m|\mathbf{x}, \sigma, I\right) = \frac{1}{2\sigma} e^{-\left(\frac{\left|F_m\left(\mathbf{x}'\right) - d_m\right|}{\sigma}\right)}, \tag{8.19}$$

that leads to the ℓ^1 misfit function. Multiplying the probability distributions of M independent measurements, each with scaling parameter ξ, results in the Likelihood function

$$P\left(\mathbf{d}|\mathbf{x}, \xi, p, I\right) = \frac{p^M}{\left(2\Gamma\left(\frac{1}{p}\right)\right)^M \xi^M} \exp\left(-\sum_{m=1}^{M} \frac{\left|F_m\left(\mathbf{x}'\right) - d_m\right|^p}{\xi^p}\right). \tag{8.20}$$

8.2.4 Prior model

In the linear model (equation 8.7 or 8.8) both, the matrices $\mathbf{W}$ and $\mathbf{N}$ are typically underdetermined or ill-conditioned. This is due to the fact the data are band-limited

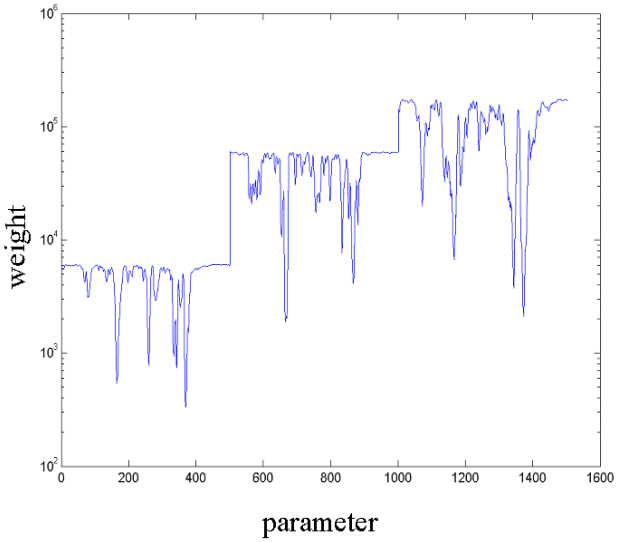

Figure 8.3: Diagonal of Q matrix calculated for final iteration of Halfway synthetic example.

and the differential tuning introduces null spaces into the NMO operator $\mathbf{N}$. Because of this, the problem needs to be regularized. Similar to Chapter 6, this is done by choosing a weighting function that treats certain reflection coefficients as being more probable than others. Choosing a long tailed *a priori* distribution leads to such a weighting function.

A long tailed distribution or sparse reflectivity may be argued for based on physical arguments. The P-wave impedance reflectivity may be modeled as a long tailed distribution, such as the ℓ^1distribution (Levy and Fullagar, 1981; Shapiro and Hubral, 1999). Under the change of variables the second variable is similar to the fluid factor. The fluid factor reflectivity is sparse by its nature since it only responds to anomalous fluids or large changes in lithology. The third variable is similar to a difference between the scaled density and the velocity reacting to places where the density is uncorrelated to the velocity. The reflectivity of this third variable is sparse as well. After the change of variables in the proceeding section, the variables are independent so the resulting parameter covariance matrix is diagonal. Similar to Chapter 6, the three reflectivity series can be modeled by a variety of distributions including the Huber, Cauchy or ℓ^p norm. For illustration purposes, the Cauchy distribution (equation G.11)

$$P\left(\mathbf{x}'\right) = \prod_{i}^{N} \frac{\exp\left(-\lambda_0\right)}{\left(\left|x_i'\right|^2 + b\right)^{\lambda_1}}, \tag{8.21}$$

is used to derive the three-term nonlinear AVO waveform inversion. In deriving the maximum *a posterior* (MAP) solution the derivative of the negative natural logarithm of the prior probability distribution (8.21) J_{Cauchy} is required. For the three-parameter problem this is

$$\frac{\partial J_{Cauchy}}{\partial x_m} = \frac{\lambda_1}{\sigma_{1_m}^2}\mathbf{Q}_c\mathbf{x}', \tag{8.22}$$

where $\mathbf{Q}$ is a diagonal weighting matrix whose elements are defined by

$$Q_{nn} = \begin{cases} \frac{1}{\left(\frac{x_n'^2}{2\sigma_1^2}+1\right)} & n \leq L \\ \frac{\sigma_1^2}{\sigma_2^2}\frac{1}{\left(\frac{x_n'^2}{2\sigma_2^2}+1\right)} & L < n \leq 2L \\ \frac{\sigma_1^2}{\sigma_3^2}\frac{1}{\left(\frac{x_n'^2}{2\sigma_3^2}+1\right)} & 2L < n \leq 3L \end{cases} . \tag{8.23}$$

These weights rely on parameter estimates themselves and so must be calculated

in a bootstrap fashion. The methodology is similar to that of Sacchi and Ulrych (1995) and that used in Chapter 6. Figure 8.3 shows weights that were calculated for synthetic example in Section 8.3.2. In general large reflectivity results in small weights and small reflectivity in large weights. It is also possible to see the influence of the different eigenvalues associated with the different variables in Figure 8.3. The weights associated with the third eigenvalue ($1001 < n \leq 1500$) are much larger than that for the first eigenvalue ($1 < n \leq 500$) as equation (8.23) implies. This means that information will preferentially be put into the first eigenvector rather than the last eigenvector.

8.2.5 Nonlinear inversion

The likelihood function (equation 8.20) may be combined with the *a priori* probability function (equation 8.21) using Bayes' Theorem

$$P\left(\mathbf{x}'|\mathbf{d},I\right) \propto P\left(\mathbf{d}|\mathbf{x}',I\right) P\left(\mathbf{x}'|I\right), \tag{8.24}$$

resulting in

$$P\left(\mathbf{x}|\mathbf{d},p,\xi,I\right) \propto \exp\left(-\sum_{m=1}^{M} \frac{\left|F_m\left(\mathbf{x}'\right)-d_m\right|^p}{\xi^p}\right) \prod_{n}^{N} \frac{\exp\left(-\lambda_0\right)}{\left(\left|x_n'\right|^2+b\right)^{\lambda_1}}. \tag{8.25}$$

The MAP solution occurs where the function is stationary and convex. It is easier to take the partial derivatives of the logarithm of the function

$$L\left(\mathbf{x}|\mathbf{d},p,\xi,I\right) \propto -\sum_{m=1}^{M} \frac{\left|F_m\left(\mathbf{x}'\right)-d_m\right|^p}{\xi^p} + \sum_{n}^{N} \frac{\exp\left(-\lambda_0\right)}{\left(\left|x_n'\right|^2+b\right)^{\lambda_1}}. \tag{8.26}$$

Thus, the partial derivative of the objective function

$$\frac{\partial J}{\partial x_j} = 0 = -\frac{\partial J_p}{\partial x_n} - \frac{\partial J_{Cauchy}}{\partial x_n}, \tag{8.27}$$

is the sum of two parts, the ℓ^p misfit function

$$J_p = \sum_{m=1}^{M} \frac{\left|F_m\left(\mathbf{x}'\right)-d_m\right|^p}{\xi^p}, \tag{8.28}$$

and the *a priori* function, in this case J_{Cauchy}. Substituting the partial derivative of J_{Cauchy} (equation 8.22) and J_p (Appendix G.4)

$$\boldsymbol{\nabla} J_p = \frac{\mathbf{G}'^T \mathbf{W}_{\ell_p} \mathbf{G}' \mathbf{x}' - \mathbf{G}'^T \mathbf{W}_{\ell_p} \mathbf{d}}{\xi^2}, \tag{8.29}$$

results in

$$\left[\mathbf{G}'^T \mathbf{W}_e \mathbf{G}' \mathbf{x}' + \Theta \mathbf{Q}_c\right] \mathbf{x}' = \mathbf{G}'^T \mathbf{W}_e \mathbf{d}, \tag{8.30}$$

where

$$\Theta = \frac{\lambda_1 \xi^2}{\sigma_1^2}, \tag{8.31}$$

$$\mathbf{W}_e = diag \left\{ \begin{array}{cc} \epsilon^{p-2} & \left|\frac{\varepsilon_m}{\xi}\right| \leq \epsilon \\ \left|\frac{\varepsilon_m}{\xi}\right|^{p-2} & \left|\frac{\varepsilon_m}{\xi}\right| > \epsilon \end{array} \right\}_{m=1}^{M}, \tag{8.32}$$

and ϵ is a small number. The inequality $\left|\frac{\varepsilon_m}{\xi}\right| \leq \epsilon$ is introduced to protect against the singularity that arises for $1 \leq p < 2$.

There are three sources of nonlinearity in equation (8.30), the estimate of the penalty term Θ, and the calculation of the diagonal weighting matrices $\mathbf{Q}$ and $\mathbf{W}_e$ (equation 8.23 and 8.32). Depending on the weights used, there are unknown hyper-parameters in the calculation of Θ. The optimal choice of Θ may be found by doing a parameter search as outlined in Chapter 6. I also found that one could compare, at the well control, the amount of dispersion in the estimated R_p, R_d crossplot versus the ideal. If the constraints are applied too harshly, then the crossplotted reflectivity appears as a line. If the constraints are applied too loosely, there is too much scatter in the crossplot space. Thus, the amount of dispersion may be use to adjust the value of Θ so it appears similar to that of the well control.

The other source of nonlinearity is the construction of the weights $\mathbf{Q}$ and $\mathbf{W}_e$. These require previous estimates of $\mathbf{x}'$ and $\boldsymbol{\varepsilon}$ so must be calculated in a bootstrap fashion. Similar to Chapter 6, the inverse problem is solved by performing two nested loops. The outer loop consists of calculating the penalty function Θ, and the two diagonal weighting matrices $\mathbf{Q}$ and $\mathbf{W}_e$ based on the previous estimates of $\mathbf{x}'$ and $\boldsymbol{\varepsilon}$ The inner loop calculates $\mathbf{x}'$ and $\boldsymbol{\varepsilon}$ by solving equation (8.30) with conjugate gradient (Shewchuck, 1994), using the maximum number of iterations as a regularization parameter (Hansen, 1998). For the initial loop the weighting matrix $\mathbf{W}_e$ is set to the identity matrix while the penalty term Θ is set to zero. Generally,

a satisfactory sparse solution was obtained after two to three outer loops. The three-term inversion is more ill-conditioned than the two-term inversion. I was never able to achieve a sparse spike solution. The estimate of $\mathbf{x}'$ always appeared band-limited. To avoid the solution becoming unstable, I had to more severely restrict the maximum number of conjugate gradient solution iterations run, compared to the two term solution. This excluded small eigenvalues that made the solution more unstable, but were also important in creating the spiky appearance of the two term solution. After the final iteration, the parameters can be transformed to a more natural parameterization using equation (8.12). In fact, the parameters may be transformed to any convenient parameterization using the three-term transform (equation B.54).

Misfit weights

A number of different misfit weighting functions may be used. They can be derived from long-tailed distributions in a similar fashion as done in Chapter 6 for the *a priori* distributions. The ℓ^1 and ℓ^2 weights may be derived from the general ℓ^p norm (equation 8.32). The ℓ^1 norm may be simulated by this, for in the limit as $p \rightarrow 1$ the weights become

$$\mathbf{W}_{\ell^1} = diag \left\{ \begin{array}{cc} \frac{1}{\epsilon} & \left|\frac{\varepsilon_m}{\sigma}\right| \leq \epsilon \\ \frac{1}{\left|\frac{\varepsilon_m}{\sigma}\right|} & \left|\frac{\varepsilon_m}{\sigma}\right| > \epsilon \end{array} \right\} \quad \text{where } \epsilon << 1, \tag{8.33}$$

The Gaussian weights also follow from equation (8.32), for $p = 2$ the weights are the identity matrix

$$\mathbf{W} = \mathbf{I}. \tag{8.34}$$

which is consistent with the results derived for Gaussian misfit functions in Chapter 6. Bube and Langan (1997) introduced a misfit function that behaves like a ℓ^2 misfit function for small misfit and a ℓ^1 misfit function for large misfit. This ℓ^1/ℓ^2 hybrid weighting function is

$$\mathbf{W}_{BL} = diag \left\{ \frac{1}{\sqrt{\frac{\varepsilon_m^2}{\sigma^2} + 1}} \right\}, \tag{8.35}$$

where $\sigma^2 \approx \frac{0.63}{M} \sum_{m=1}^{M} \varepsilon_m^2$. They claim this misfit function has better convergence properties than equation (8.32) since the behavior for small misfit is continuous rather

than discontinuous as for ℓ^1weights. The Bube and Langan (1997) weights are very similar in form to the Cauchy misfit weights

$$\mathbf{W}_{Cauchy} = diag\left\{ \frac{1}{\left(\frac{\varepsilon_m^2}{2\sigma^2}+1\right)} \right\}, \tag{8.36}$$

where σ is a hyper-parameters. The derivation of the Cauchy misfit weights is similar to that done for priors in Chapter 6.

Lastly the Huber norm produces the weighting function

$$\mathbf{W}_{Huber} = diag\left\{ \begin{array}{cc} 1 & |\varepsilon_m| \leq \epsilon\sigma \\ \frac{1}{\frac{|\varepsilon_m|}{\epsilon\sigma}} & |\varepsilon_m| > \epsilon\sigma \end{array} \right\}, \tag{8.37}$$

where ϵ is some scalar greater than zero. The derivation again is similar to that in Chapter 6. Like the Bube and Langan (1997) weights, the Huber weights behave like a ℓ^2 misfit function for small misfit and a ℓ^1 misfit function for large misfit, but is discontinuous. Essentially the Huber misfit weighting function underweights data with misfit greater than ϵ standard deviations.

8.3 AVO waveform examples

8.3.1 Synthetic example

The synthetic example from Section 6.3 was modified to illustrate the three-term AVO waveform inversion. An additional reflector due to a density perturbation, uncorrelated with the P-wave or S-wave impedance reflectivity, was introduced at 1.9 seconds. Figure 8.4a shows the NMO corrected synthetic gather. Recall from section 6.3, that the AVO behavior of most of the reflectors in the synthetic gather followed the mudrock trend, but several Class III and IV anomalies were also present. The events at 1.5 and 1.7 seconds are Class III and IV anomalies undergoing NMO stretch while the events at 1.6 and 1.8 seconds are Class III and IV AVO anomalies undergoing offset dependent tuning and NMO stretch. These anomalies should introduce distortions into the AVO inversion.

The synthetic data was generated using a convolutional model with a Ricker wavelet with a dominant frequency of 32.5 Hz. Noise was added to give a signal-to-

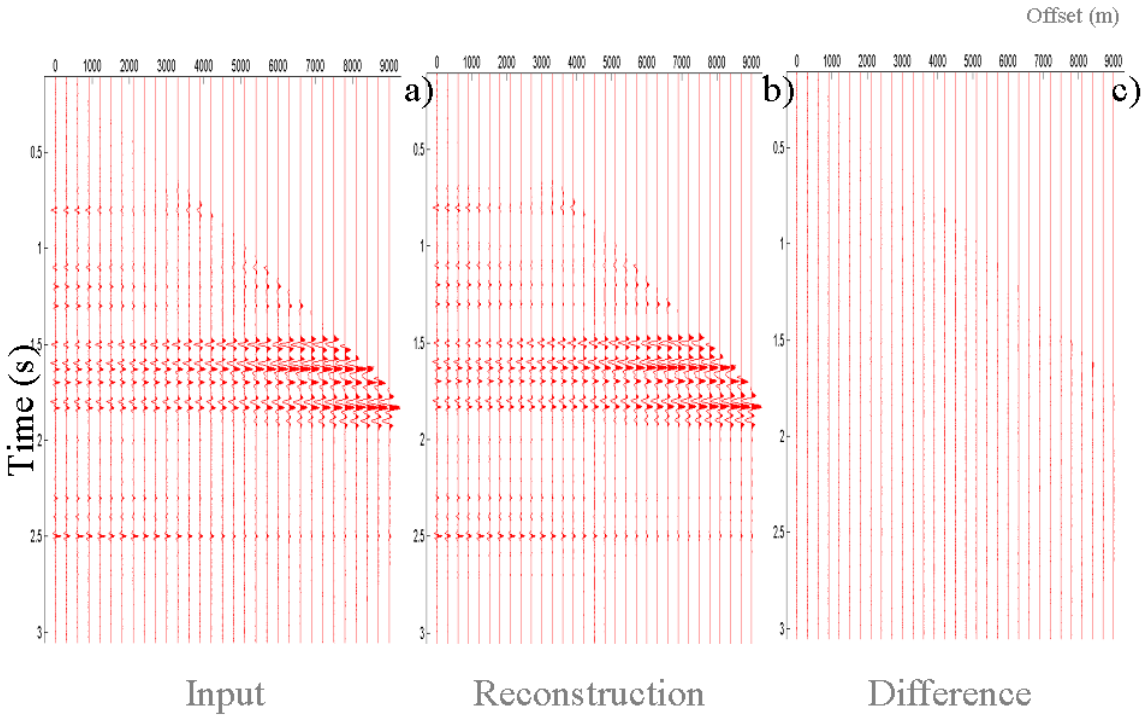

Figure 8.4: Synthetic (a) input, (b) three-term AVO waveform model reconstruction, and (c) (5X) difference between input and model reconstruction. Note the anomaly due to the density perturbation at 1.9 seconds.

noise ratio of 10:1. A constant background velocity is used so that there is a simple angle to offset relationship. The maximum offset was chosen to be four times the target depth so that angles out to 65° are available for the inversion. These large angles are included to highlight any distortions. The NMO corrected gathers were input into the three-term AVO waveform inversion, inverting angles up to 55°. The estimated model (Figure 8.4b) accurately models all the events, including those undergoing NMO stretch and offset dependent tuning, as evidenced by the small amount of noise in the difference display (Figure 8.4c).

Figure 8.5 shows the estimate of the band-limited reflectivity versus the ideal zero offset reflectivity. To the eye, the estimate appears to be almost a direct match to the ideal for all the reflectivity attributes, including the density. The density reflectivity anomaly at 1.9 seconds is accurately estimated even though it is completely uncorrelated with the P-wave and S-wave impedance reflectivity. However, crossplots (Figure 8.7) of the P-wave and S-wave impedance reflectivity, and the P-wave impedance and density reflectivity reveal that the estimates are not as accurate as the crossplots for the two term inversions (Figure 6.6). Further, the estimate of the full spectrum reflectivity is not as sparse and high frequency as that of the ideal (Figure 8.6), due to the reasons outlined in Section 8.2.5.

8.3.2 Halfway synthetic example

A more realistic example based on an actual well log is examined next. This is the Halfway synthetic model previously discussed in Chapters 3 and 5, based on a well log from northeastern British Columbia, Canada. Figure 8.8 shows the well logs and the synthetic gather. The synthetic gather was generated by a convolutional modeling scheme using reflectivity generated by the Zoeppritz equations and ray tracing. The reflectivity is convolved with a 10/14 - 90/100 Hz zero phase wavelet and muted so only angles less than 45° are shown. Gaussian noise was added to give a 4:1 signal-to-noise ratio. Figure 8.9 shows the same model after NMO correction. Note the apparent tuning and NMO stretch artifacts, for example, the event at 0.32 seconds. Figure 8.9 also shows the input P-wave velocity and density reflectivity. Note that the density reflectivity behaves differently than the P-wave velocity reflectivity, particularly at 0.32 and 0.67 seconds.

In Chapter 3, this same model was generated, but without NMO. The constrained three-term AVO inversion was able to accurately estimate all the reflectivity attributes

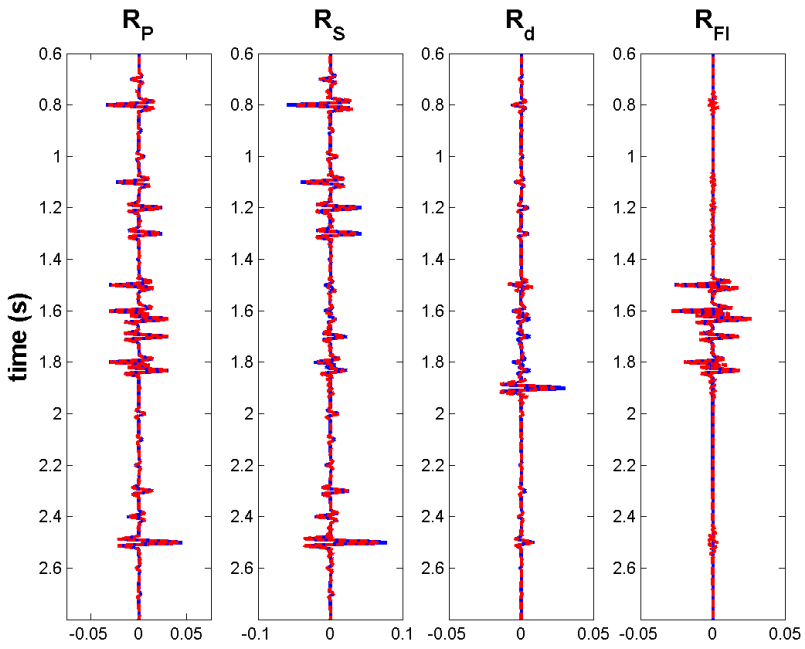

Figure 8.5: Comparison of filtered three-term AVO waveform inversion for the P- and S-wave impedance, density and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue) for synthetic example 1b. Note the excellent match.

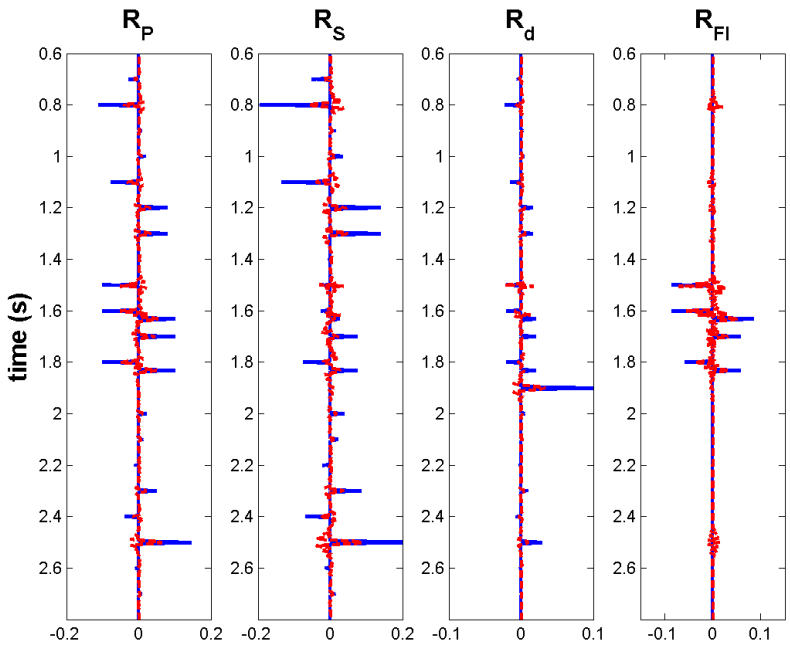

Figure 8.6: Comparison of unfiltered three-term AVO waveform inversion for the P- and S-wave impedance, density and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue) for synthetic example 1b. Note the poor match.

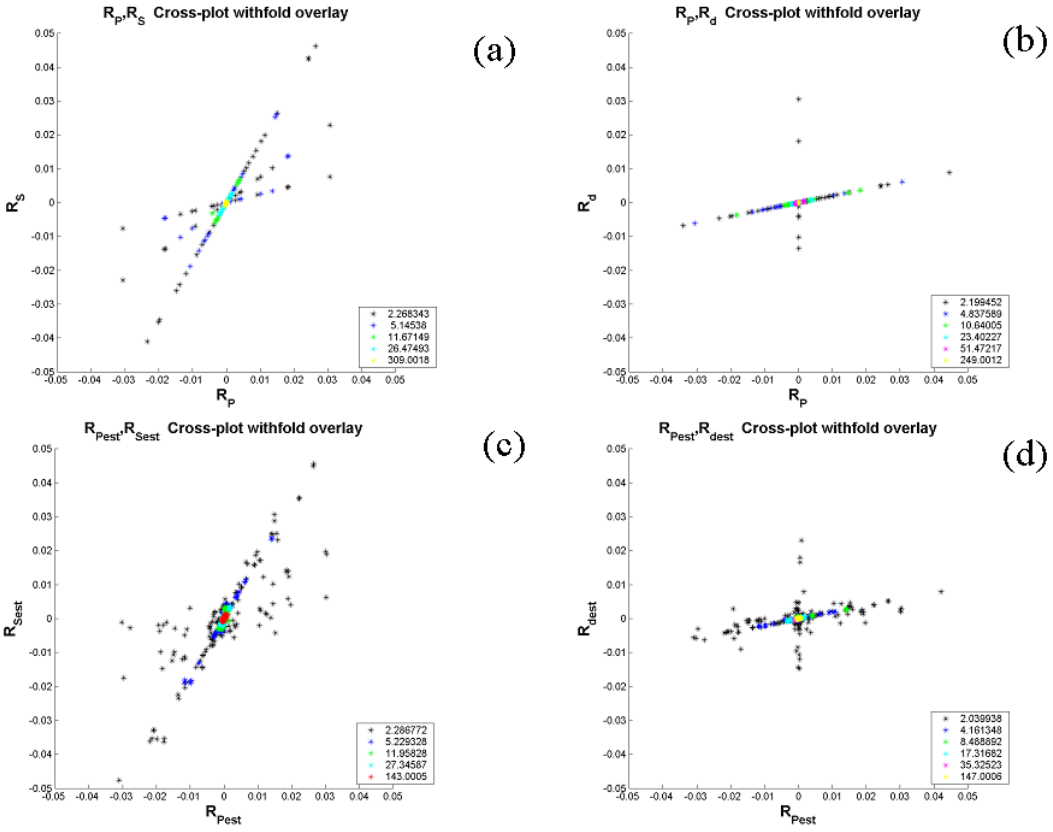

Figure 8.7: Cross-plot of filtered (a) ideal P- and S- wave impedance reflectivity, (b) estimated P-and S-wave impedance reflectivity, (c) ideal P-wave impedance and density reflectivity, and (d) estimated P-wave impedance and density reflectivity. Note scatter in the estimate due to error.

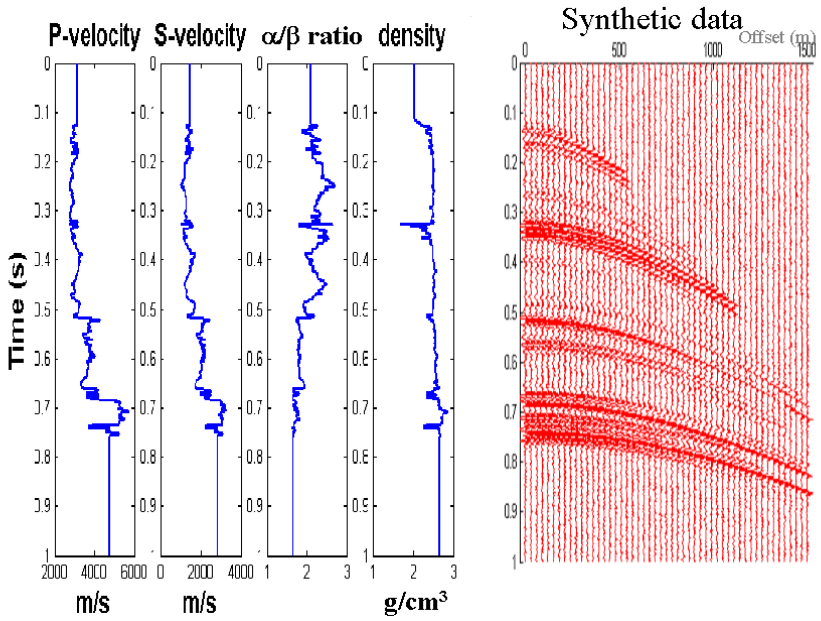

Figure 8.8: Wireline logs displayed in time for the Halfway well log. The synthetic gather was generated using reflectivity generated from the Zoeppritz equation. The reflectivity is shown without moveout, filtered and with noise added to give a S/N=4.

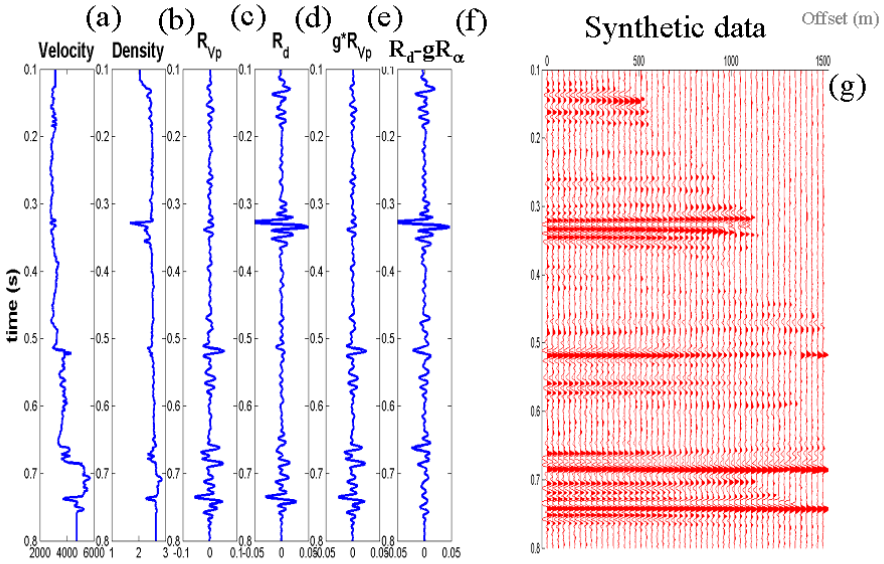

Figure 8.9: P-wave velocity and density of Halfway well log. Note that density and velocity are uncorrelated at 0.32 and 0.68 seconds. Note the corresponding reflectivity is as well. Density estimated using the Gardner equation Rd=g*Rvp correlates poorly with the actual density. Note that the NMO corrected synthetic gather (g) shows offset dependent tuning on the event at 0.32 seconds as well.

including density. However, if the NMO corrected gathers are input into the inversion instead, the results are not as good (Figure 8.10). This is due to the error introduced by the NMO stretch and tuning.

The three-term AVO waveform inversion incorporates NMO stretch and offset dependent tuning into its forward model, so should be unaffected by these issues. The same synthetic model, this time with NMO, was inverted using the three-term AVO waveform inversion, using uniform misfit weights and the same constraint parameters as used in Chapter 3. Figure 8.11 shows the input, the estimated model and the difference between the two. The inversion is able to accurately model all the features in the data set. Figure 8.12a shows the estimated versus ideal zero offset reflectivity. Once again the match is excellent with only small errors. All the major features are accurately predicted. The density is estimated accurately even where it is uncorrelated with either the P- or S-wave impedance reflectivity. Figure 8.13 shows the crossplots for the ideal and estimated reflectivity. The crossplots show essentially the same trends and features.

The three-term AVO waveform inversion was also done using NMO corrected gathers as the input. The ideal and estimated zero offset reflectivity (Figure 8.12b) are essentially identical with those based on the gathers with NMO with some subtle differences (Figure 8.12a). For example, the NMO corrected input inversion results shows slightly greater error on the Halfway reflector (trough at 0.74 seconds). Figure 8.14 shows the NMO corrected input, the estimated model and the difference. The estimated model once again accurately models the events even with NMO stretch and tuning.

8.3.3 Comparison of different misfit weighting functions on Halfway synthetic

The previous results were obtained using uniform misfit weights ($\mathbf{W}_e = \mathbf{I}$). If there is non-Gaussian noise in the data, it might be more appropriate to use a different weighting scheme. To test this, non-uniform and non-Gaussian noise was added to the synthetic model (Figure 8.15a). Specifically, non-uniform Gaussian noise $10\times$ times greater than that of the other traces was introduced to the first trace. In addition, an event with hyperbolic moveout was introduced to simulate interference from a multiple reflection. Based on the Downton and Mackidd (1997) results, the multiple should bias the estimate around 0.51 seconds where the near offsets are interfered

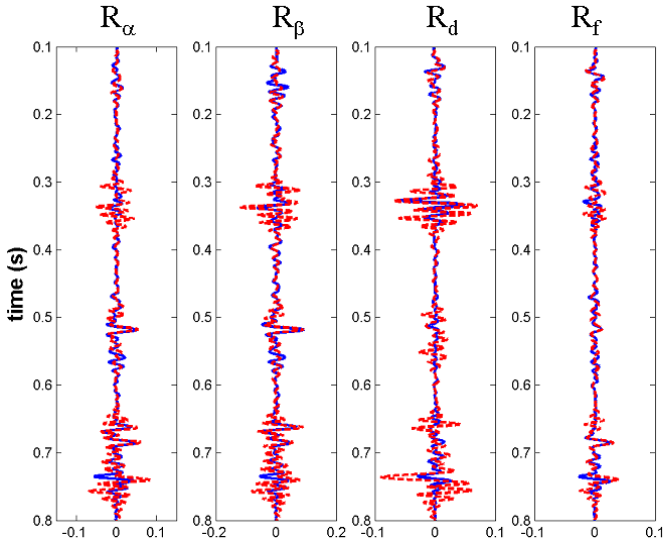

Figure 8.10: Comparison of the estimated P- and S-wave velocity, density, and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue) based on three-term AVO inversion of NMO corrected Halfway synthetic. Note that the estimate is a poor due to the theoretical error introduced by NMO correction.

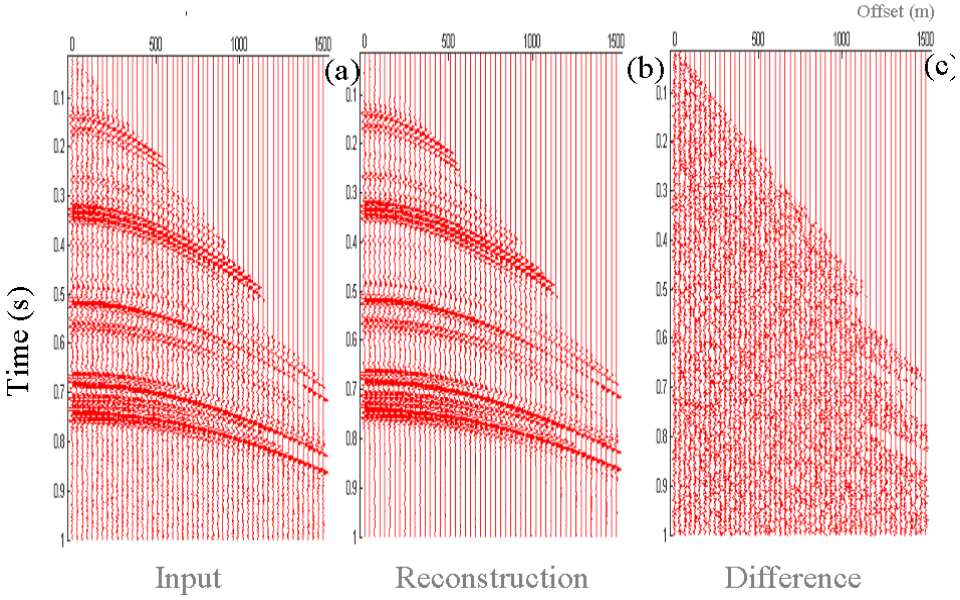

Figure 8.11: Halfway synthetic input (a), model reconstruction (b), and (5X) difference between input and model reconstruction (c) for the three-term AVO waveform inversion. Note the good match and small difference.

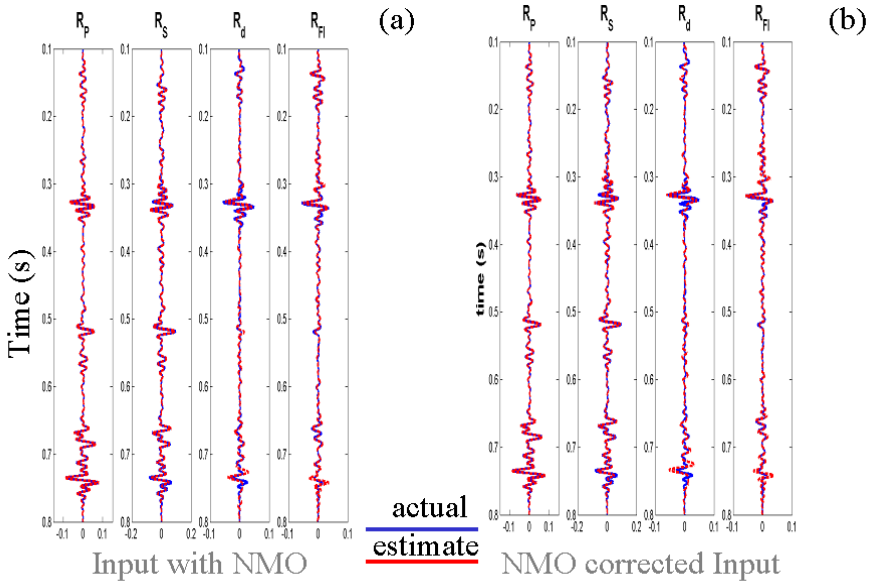

Figure 8.12: Comparison of the estimated P- and S-wave impedance, density, and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue) based on three-term AVO inversion of NMO corrected Halfway synthetic. Note there is little difference between doing the AVO waveform inversion with NMO (a) or after NMO correction (b).

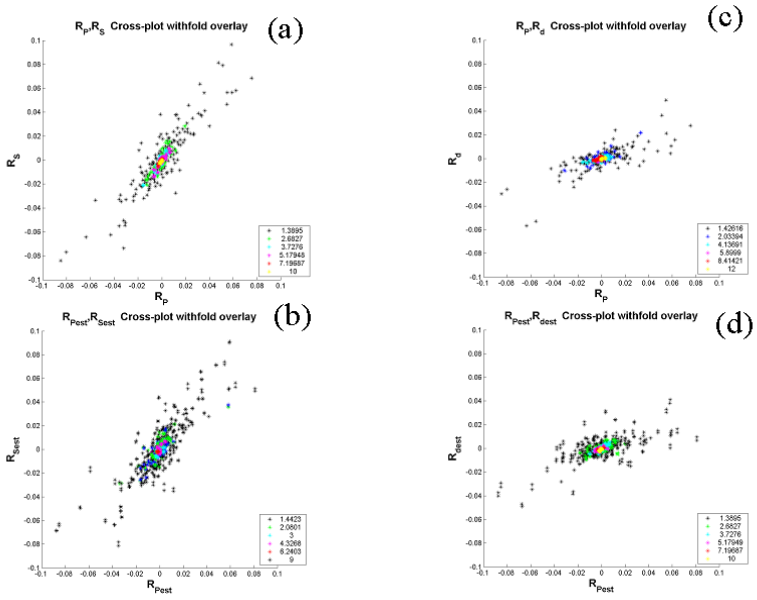

Figure 8.13: Cross-plot of filtered (a) ideal P- and S- wave impedance reflectivity, (b) estimated P-and S-wave impedance reflectivity, (c) ideal P-wave impedance and density reflectivity, and (d) estimated P-wave impedance and density reflectivity. Note good correspondence between estimated and ideal data.

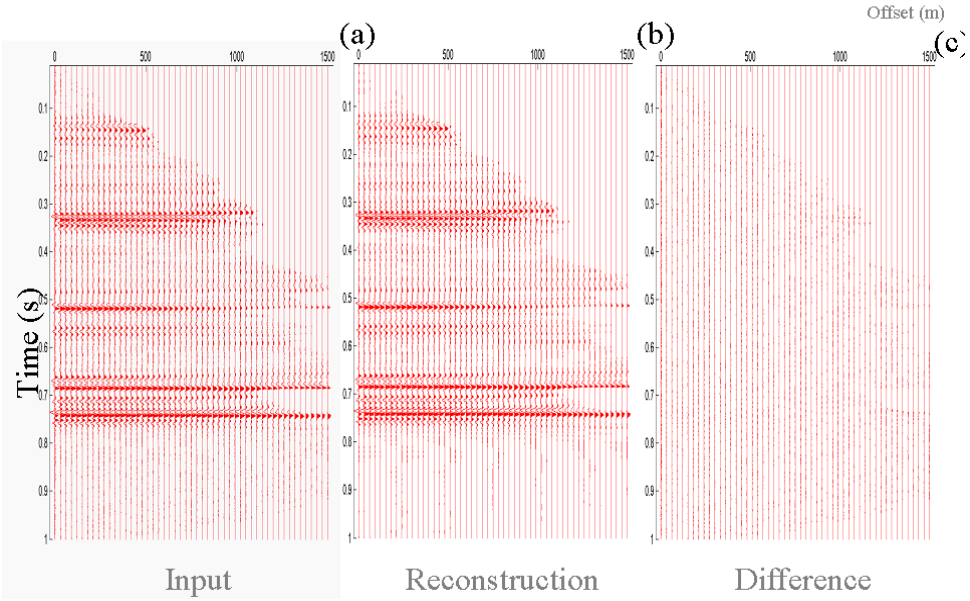

Figure 8.14: Halfway synthetic (a) input, (b) model reconstruction, and (c) difference between input and model reconstruction for three-term AVO waveform inversion. Note the input is NMO corrected and that the reconstruction models the NMO stretch and offset dependent tuning.

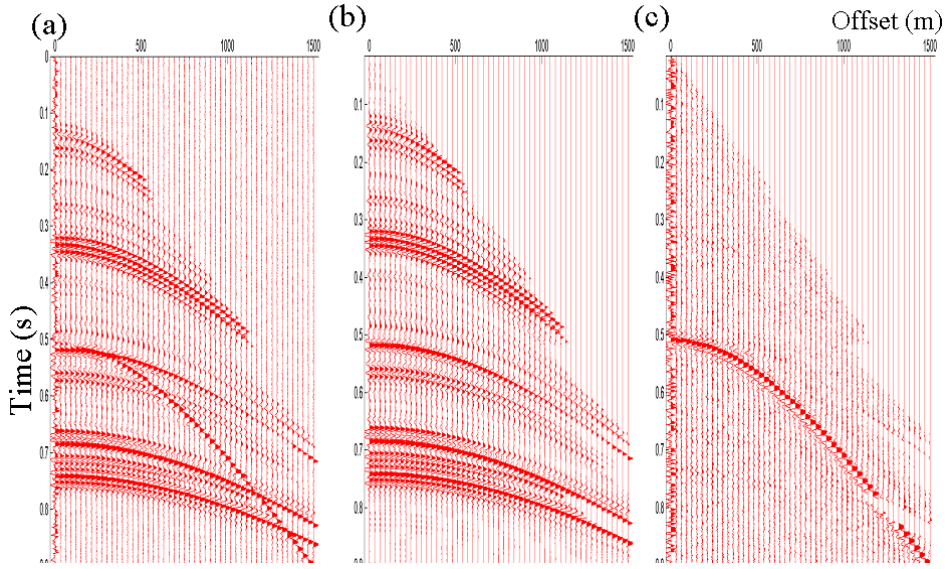

Figure 8.15: Halfway synthetic (a) input, (b) model reconstruction, and (c) difference between input and model reconstruction for interatively reweighted least-squares three-term AVO waveform inversion (Cauchy weights).

with. The three-term inversion is quite sensitive to noise so the expectation is that there should also be error in the estimate due to interference on the far offsets. The AVO waveform inversion was performed on this modified synthetic model testing all the weights outlined in Section 8.2.5.

Figure 8.16 shows the three-term AVO waveform inversion performed with uniform weights. The density reflectivity displays significant error over the interval coinciding with the multiple. It does not appear that the noisy trace influences the solution. Figure 8.17 displays the result using the Huber weights. The threshold ϵ was set to be two standard deviations, so that data with misfit greater than two standard deviations were weighted down. Figure 8.18 shows the misfit weights that were applied to the solution. The solution is now significantly better but there is still error. The density estimate is poor over the interval from 0.67 to 0.75 seconds.

Figure 8.19 shows the reflectivity estimates based on the Cauchy misfit weights (Figure 8.20) while Figure 8.21 shows the results using the Bube and Langan weights (Figure 8.22). The Bube and Langan weights are the square root of the Cauchy weights. The Bube and Langan weights lead to a solution that seems to be slightly better than the Cauchy weight solution. Various values were tried for the hyper-parameter σ in equations (8.35) and (8.36) trying values from 1/2 to 3 times $\hat{\sigma}_N$. The overall solution varied somewhat, but was quite insensitive to the choice of this parameter. In the end one standard deviation was chosen. Figure 8.15 shows the model reconstruction and difference based on the Cauchy weights. The spurious noise on trace 1 and multiple are well estimated and removed.

The best solution is that provided by the ℓ^1 weights (Figure 8.23 and 8.24). This solution approaches that of the AVO waveform inversion done on the synthetic with only Gaussian noise. The S-wave impedance reflectivity with ℓ^1 misfit weights shows greater error at 0.51 seconds compared to the Figure 8.12a). The threshold was set to $\epsilon = 0.1$ for this solution. Originally smaller thresholds (e.g. $\epsilon = 0.01$ and 0.001) were tried but with poor results.

8.3.4 Real seismic data example

The three-term AVO waveform inversion was also run on the Halfway seismic line shown in Chapter 3 and Chapter 5. This seismic line was acquired to explore for Halfway sand potential (Downton and Tonn, 1997). The data have a relatively good signal-to-noise ratio with the exception of some areas acquired through muskeg. In

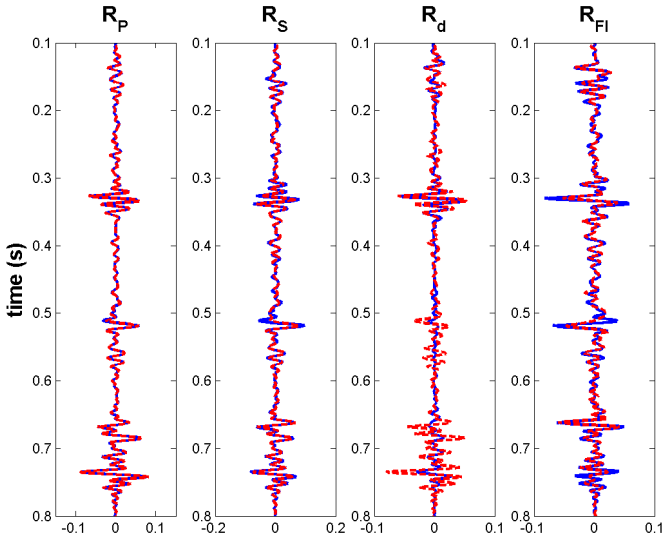

Figure 8.16: Comparison of the AVO waveform estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity for the Halfway model with non-Gaussian noise. Note the poor estimate of the density reflectivity over the interval from 0.5 to 0.76 seconds.

these areas there is quite a bit of shot generated noise. Offsets are long enough to perform the three-term AVO inversion on angles out to 45°. Forward modeling indicates that within this range supercritical reflections are not encountered. Figure 8.25 shows the three-term P-wave impedance reflectivity, S-wave impedance reflectivity, and density reflectivity estimates generated in Chapter 3. There are two bright spots on the line, evident on the P-wave impedance reflectivity section, at around 0.72 seconds The producing field (wells C and E) should have both a low velocity and density response, hence large reflectivity responses. The uneconomic gas sand, at well A, should have a large velocity but poor density response. This is evident on the reflectivity sections. The P-wave impedance reflectivity shows bright spots at both the wells. The density reflectivity section shows a poorer anomaly at the well with uneconomic gas saturations (well A) than the producing field.

Figure 8.26 shows the three-term AVO waveform inversion based on gathers with NMO. The results are a significant improvement over those from Chapter 3. Notice that the AVO waveform inversion estimate of the S-wave impedance reflectivity is much better correlated with the P-wave impedance than result from Chapter 3. This seems to be more geologically plausible. The density reflectivity shows the same sort

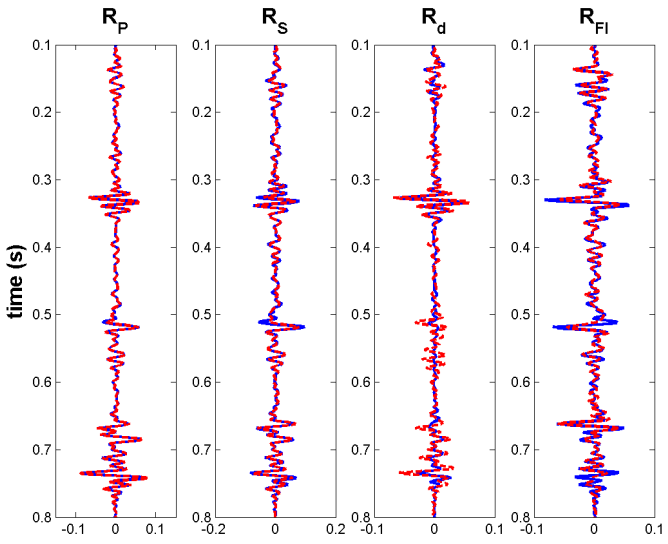

Figure 8.17: Comparison of the reweighted AVO waveform estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity for the Halfway model with non-Gaussian noise. Note the poor estimate of the density reflectivity over the interval from 0.5 to 0.76 seconds. Note the improvement in the estimate of the density reflectivity compared to Fig. 8.16. Huber weights with $\epsilon = 2$ are used.

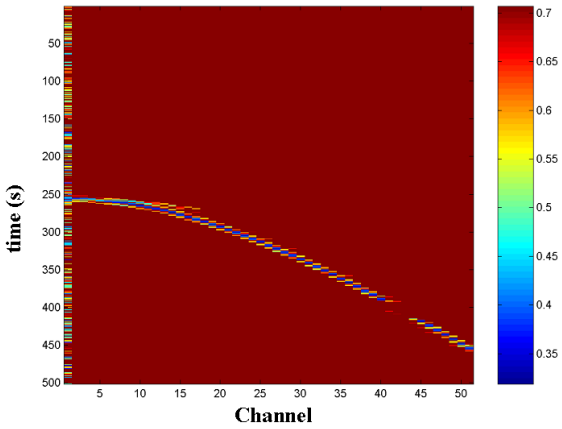

Figure 8.18: Huber weights with $\epsilon = 2$ are used to generate Fig. 8.17.

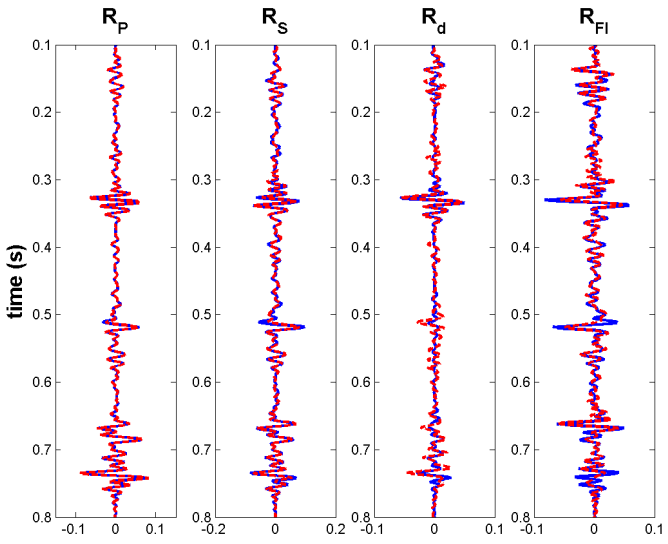

Figure 8.19: Comparison of the reweighted AVO waveform estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity for the Halfway model with non-Gaussian noise. Note the poor estimate of the density reflectivity over the interval from 0.5 to 0.76 seconds. Note the improvement in the estimate of the density reflectivity compared to Fig. 8.17. Cauchy weights are used.

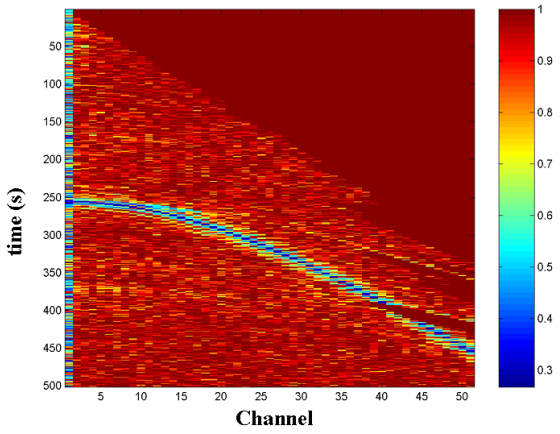

Figure 8.20: Cauchy weights that are used to generate Fig. 8.19.

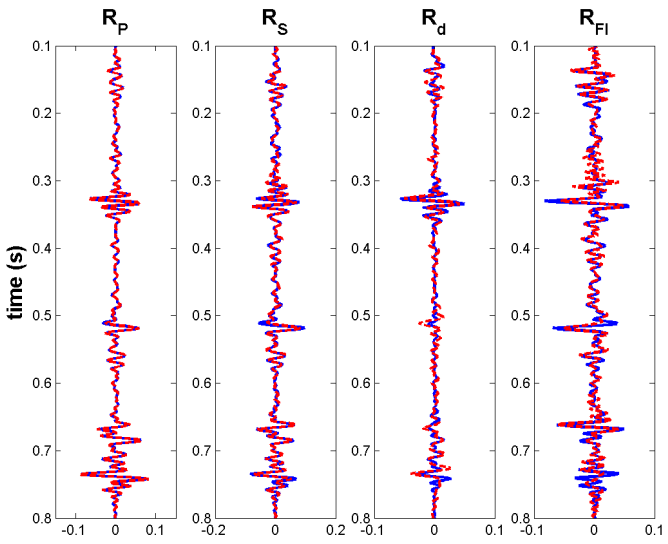

Figure 8.21: Comparison of the reweighted AVO waveform estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity for the Halfway model with non-Gaussian noise. Note the poor estimate of the density reflectivity over the interval from 0.5 to 0.76 seconds. Note the improvement in the estimate of the density reflectivity compared to Fig. 8.18. Bube and Langan (1997) weights are used.

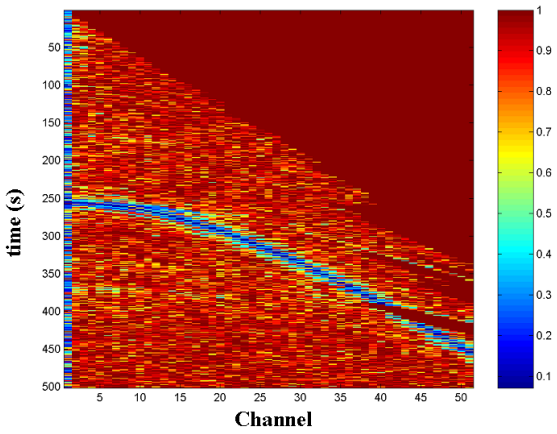

Figure 8.22: Bube and Langan (1997) weights are used to generate Fig. 8.21.

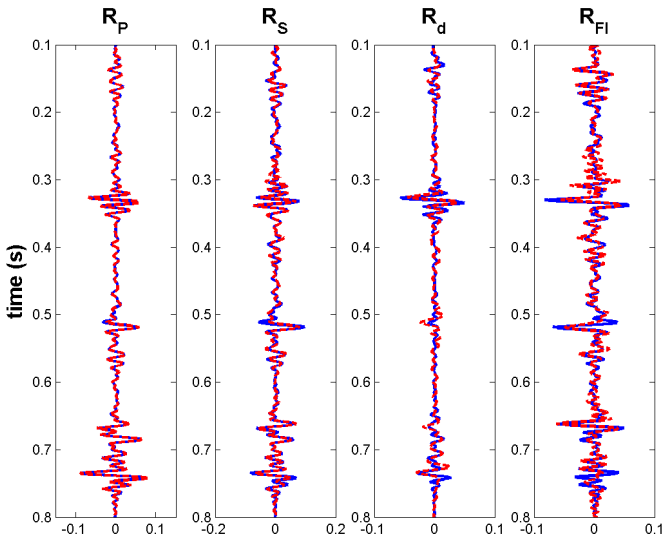

Figure 8.23: Comparison of the reweighted AVO waveform estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity for the Halfway model with non-Gaussian noise. Note the poor estimate of the density reflectivity over the interval from 0.5 to 0.76 seconds. Note the improvement in the estimate of the density reflectivity compared to 8.16. L1 weights with epsilon=0.1 are used.

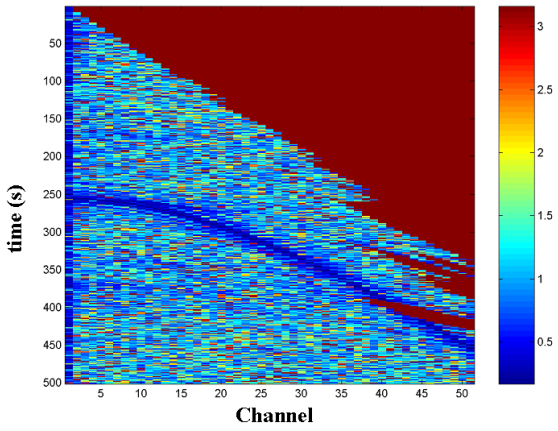

Figure 8.24: L1 weights with epsilon=0.1 that were used to generate Fig. 8.23.

of response at the zone of interest for both inversions. The AVO waveform inversion seems to suffer more from the noise associated with shots acquired in the muskeg around CDP 2000.

The AVO waveform inversion was also run with NMO corrected gathers (Figure 8.27). The results of this are comparable to the previous results. Around CDP 2100 the result seems to be much better, however there are noise spikes on the density reflectivity between CDP 3000 to 3500. Overall, the results are quite similar in quality.

To address these noise issues, the AVO waveform inversion was rerun but using iteratively reweighted least squares. The input to all of these inversions is NMO corrected. Figure 8.28 shows the results with the Huber misfit weighting function (equation 8.37) with $\epsilon = 3$. This results in estimates that are more continuous in the noisy area, with less spiky noise on them. Figure 8.29 shows the results of the ℓ^1 norm inversion, which are significantly worse than than the others. Potentially better results might be obtainable with a different choice of ϵ. I found the ℓ^1 norm solution to be quite sensitive to parameter selection. Figure 8.30 shows the result using Cauchy weights. The reflectivity estimate is now quite continuous in the noisy areas. Figure 8.31 shows the results using the Bube and Langan weights. Both the Cauchy and the Bube and Langan weights give good results with only subtle differences between them. I slightly prefer the Cauchy weights since they result in higher frequency estimates, with better lateral continuity of reflections.

Figure 8.32a shows the input seismic data at CDP 1500, while Figure 8.32b shows the estimated data from the AVO waveform inversion. The difference between the two only shows minor differences (Figure 8.32c). Note the AVO waveform inversion was able to identify and ignore the coherent noise starting at 0.5 seconds on the near offset.

8.4 Statics

Residual statics present in the AVO inversion input degrade the resulting parameter estimates. To understand the impact of this, residual statics were introduced to the Halfway synthetic gather (Figure 8.33). The AVO waveform inversion, with this input, results in poorer estimates of the density and S-wave impedance reflectivity (Figure 8.34a) than previously (Figure 8.12a). Typically, trim statics are run to remove small residual statics that surface consistent statics have left uncorrected.

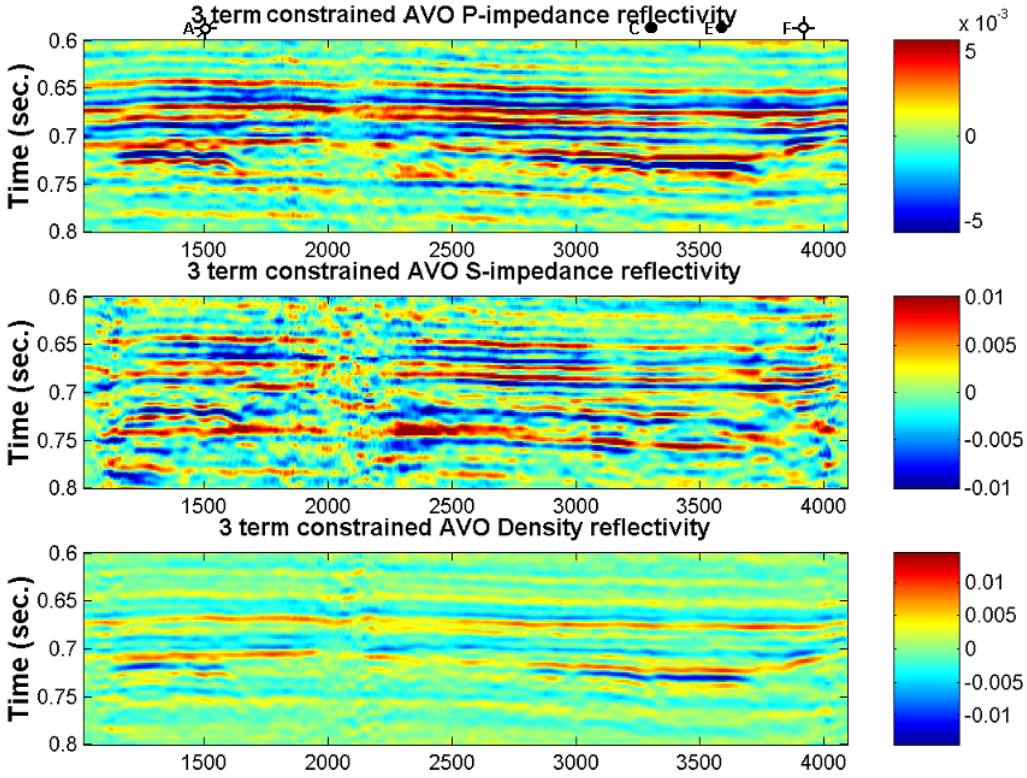

Figure 8.25: P- and S-wave impedance, and density reflectivity AVO inversions over producing and non-economic gas fields. Note that it is possible to differentiate on the density section the low gas saturation gas well (light blue Well A at 0.72 s) from the economic gas wells (dark blue at Well C and E at 0.72 s).

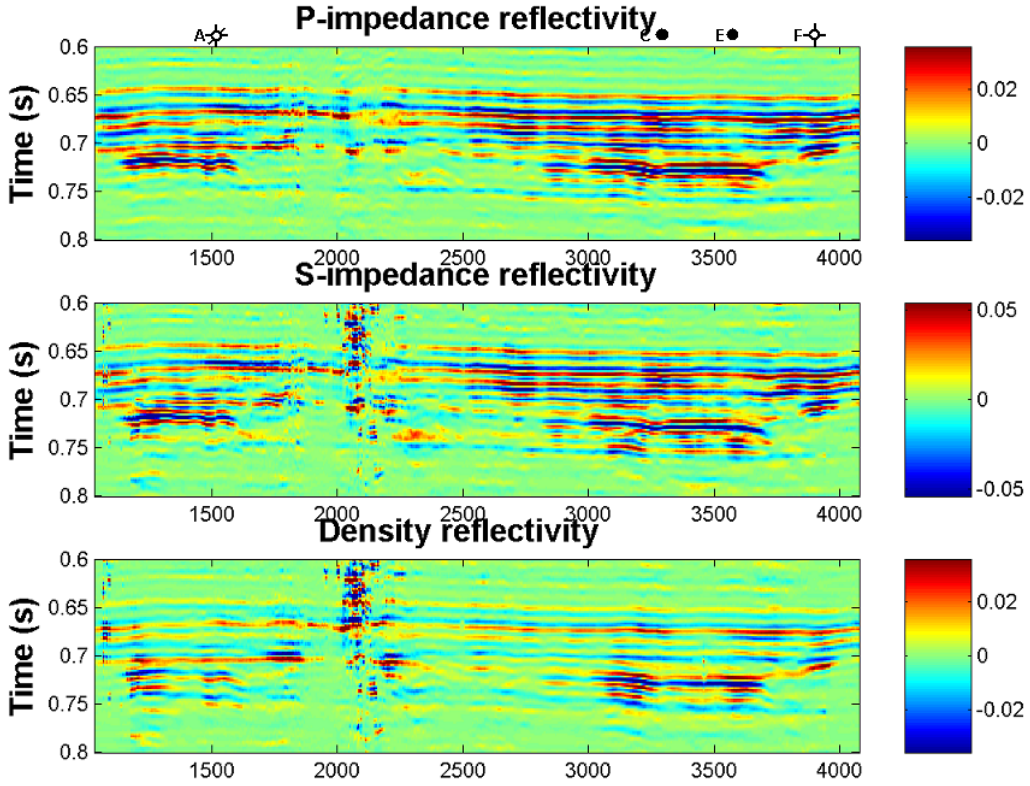

Figure 8.26: P- and S-wave impedance, and density reflectivity AVO waveform inversions. Note the improvement in the S-wave impedance reflectivity section compared to Fig. 8.25 though there are more noise artifacts.

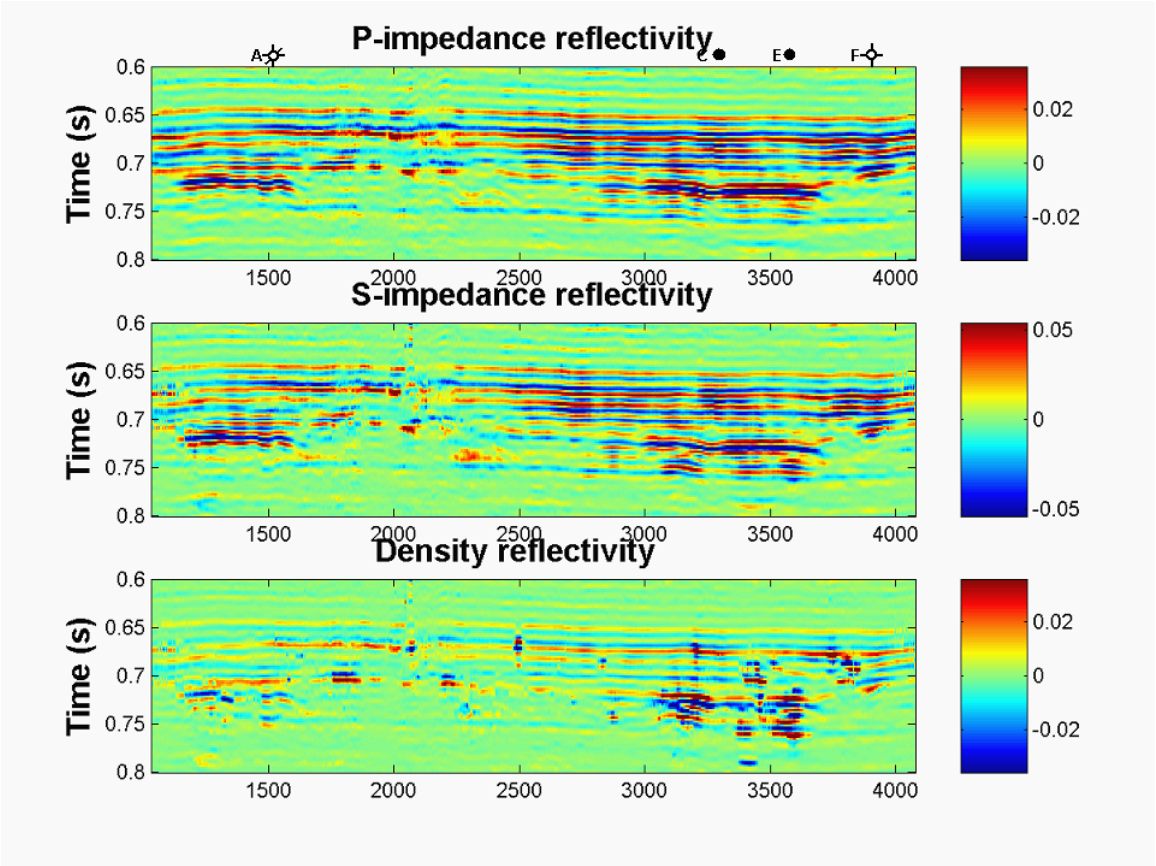

Figure 8.27: P- and S-wave impedance, and density reflectivity AVO waveform inversions based on NMO corrected gathers. Note result is very similar to that of Fig. 8.26.

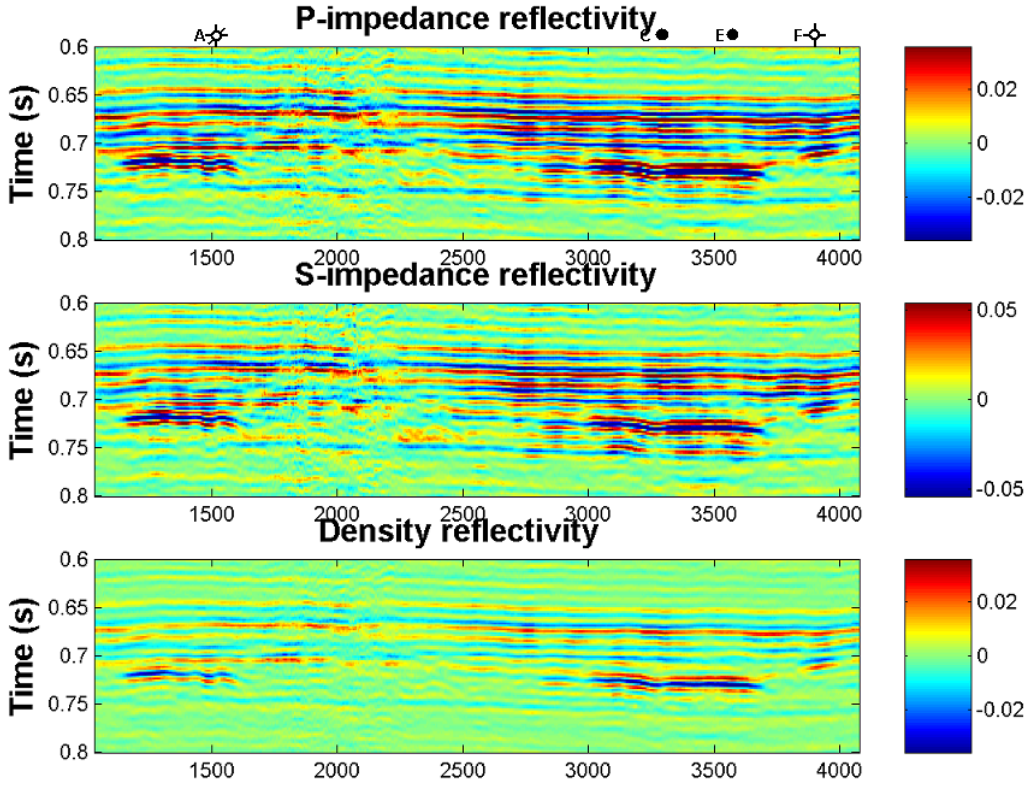

Figure 8.28: P- and S-wave impedance, and density reflectivity AVO waveform inversions using Huber misfit weights ($\epsilon = 2$). Note the improvement in the signal-to-noise ratio compared to Fig. 8.26.

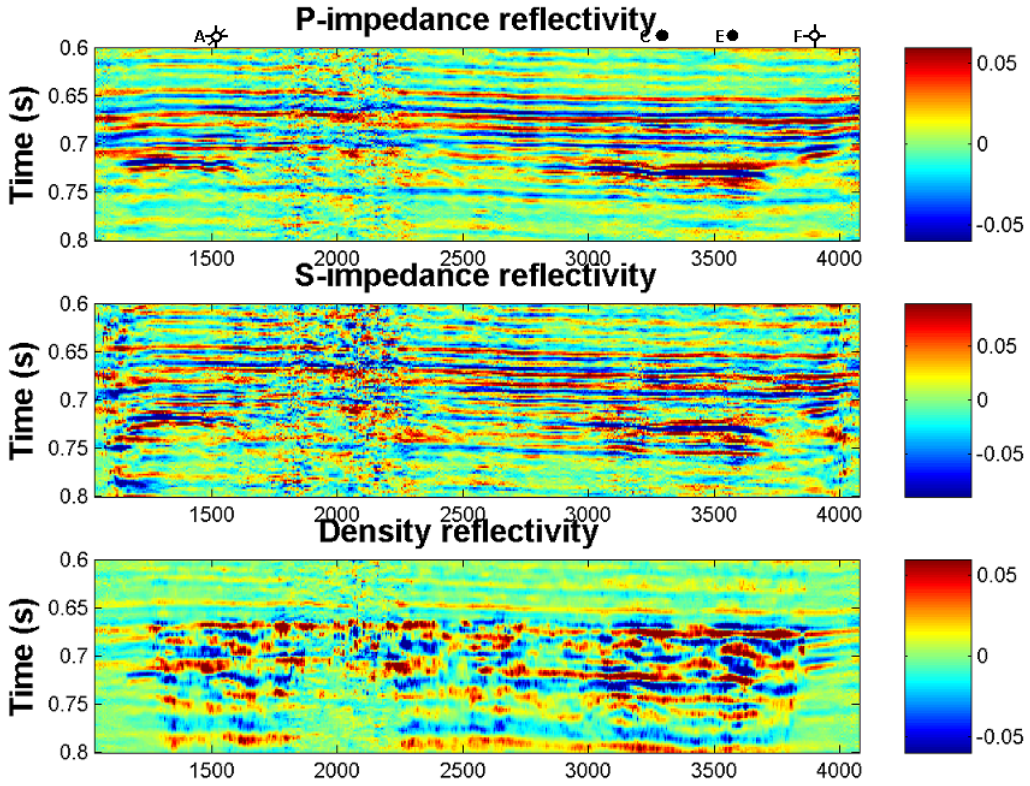

Figure 8.29: P- and S-wave impedance, and density reflectivity AVO waveform inversions using L1 misfit weights. The results are much poorer than Fig. 8.26.

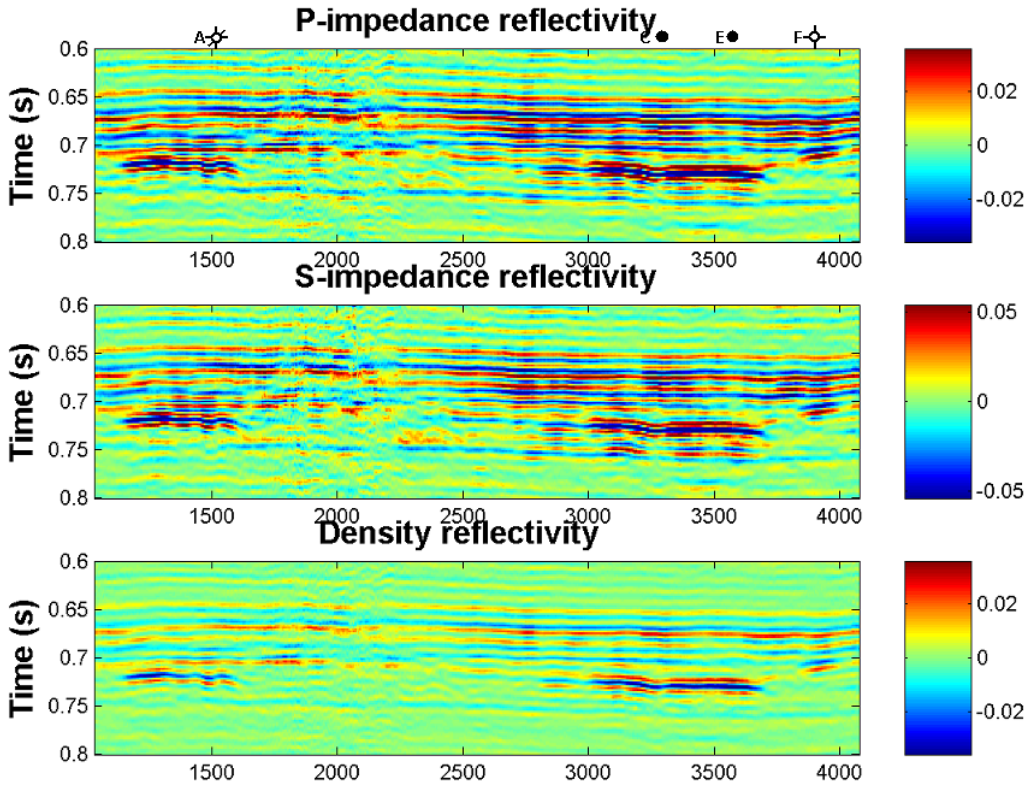

Figure 8.30: P- and S-wave impedance, and density reflectivity AVO waveform inversions using Cauchy misfit weights. Note the improvement in the signal-to-noise ratio compared to Fig. 8.26 and Fig 8.28.

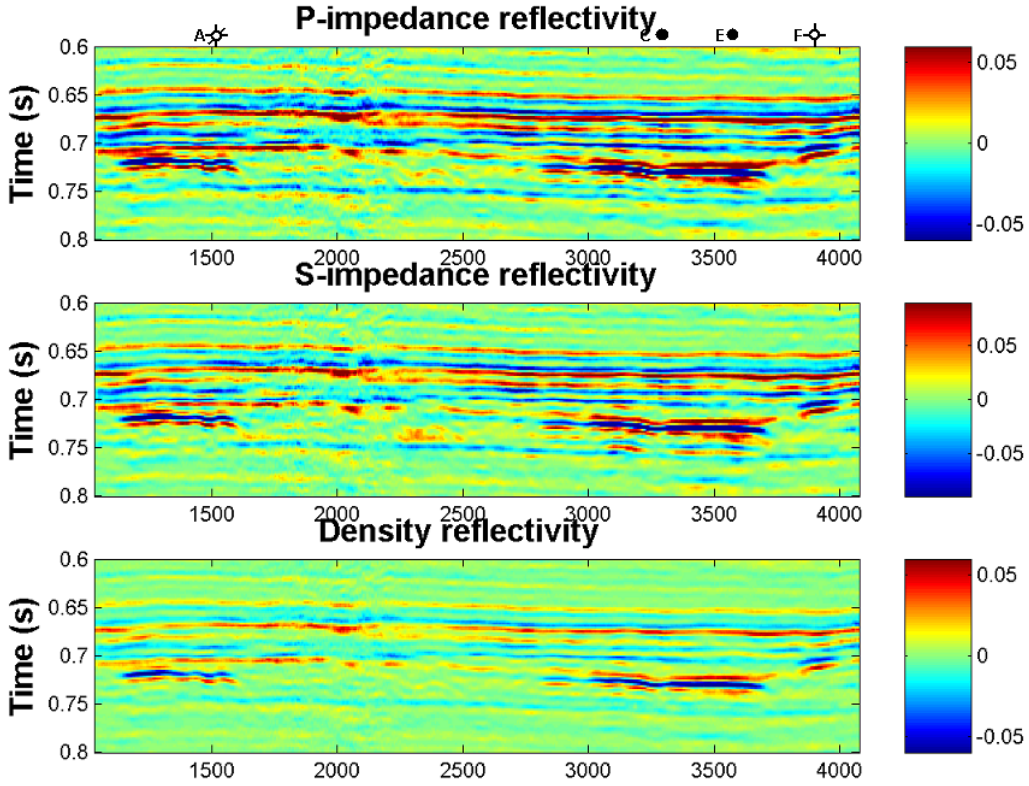

Figure 8.31: P- and S-wave impedance, and density reflectivity AVO waveform inversions using Bube and Langan (1997) misfit weights. Note the improvement in the signal-to-noise ratio compared to Fig. 8.26 and Fig. 8.28.

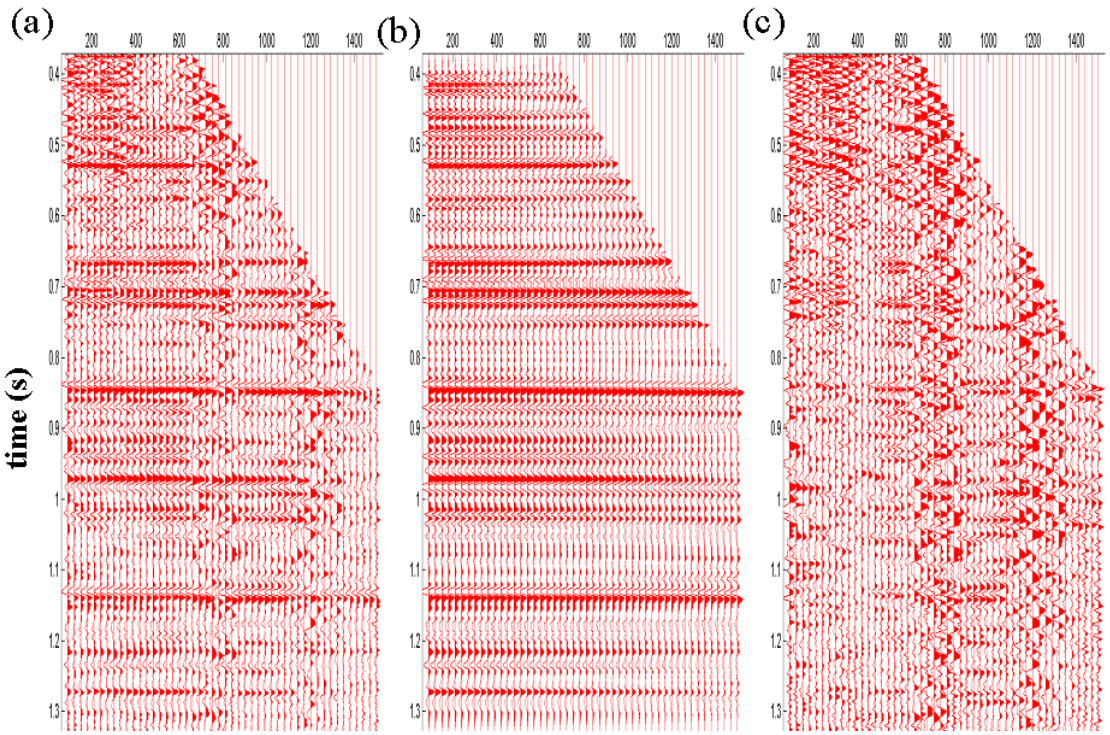

Figure 8.32: CDP gather at 1500 showing the input (a), the three-term AVO waveform estimated model using Cauchy misfit weights (b), and the (5X) difference between input and estimated reconstruction (c).

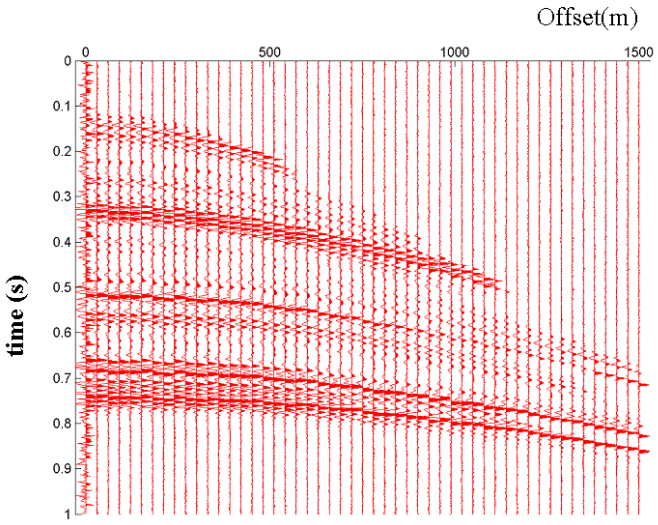

Figure 8.33: Halfway synthetic modified to include statics

Essentially this consists of creating a pilot trace (which is a usually stacked trace) and cross-correlating each prestack trace with this pilot. From these cross-correlations, time shifts may be calculated to apply to each trace within the CDP gather. The problem with this methodology is that the pilot does not take into account changes in character and amplitude as a function of offset. Instead of using a single pilot trace, the model estimated from the AVO inversion may be used as a pilot to trim into. The estimated data's amplitude and phase vary smoothly as a function of offset providing a stable pilot to trim into.

To implement this, the AVO waveform inversion was modified to have an extra outer loop. After the first outer loop, the data are estimated (modeled) based on a previous parameter estimate, and then this model is used as a pilot to trim into. This can be done before or after the application of NMO correction unlike traditional trim statics. The data are static corrected and then the AVO waveform inversion continues as before. This can be done one or more times. Figure 8.34b shows the reflectivity estimate after performing this with one additional loop. The result is significantly better than Figure 8.34a, and is close to the result achieved when there were no statics to contend with (Figure 8.12a). Figure 8.34c shows the result after three iterations

of trim statics. The results are marginally better than one iteration (Figure 8.34b). The Halfway (trough at 0.74 seconds) amplitude is better estimated on the version with three iterations of trim statics. These results suggest fast convergence.

The results shown in Figure 8.34b and 8.34c are done on gathers without NMO correction. Figure 8.35 shows the results when the input is NMO corrected with one pass of trim statics. The results are similar to those done on the data before NMO correction with one pass of trim statics (Figure 8.34b).

This methodology was also applied to the Halfway seismic data (Figure 8.36) using a Cauchy misfit weighting function. With the exception of subtle differences, the results are quite similar to the previous results (Figure 8.27) in many respects. This is probably due to the initial statics solution being sufficient.

8.5 Discussion

In Chapter 3 it was shown that in order to obtain reliable estimates of density reflectivity, large incidence angles and offsets were required. Unfortunately these long offsets introduce additional issues in terms of the processing and gain correction. In this chapter, some of the issues of using long offset data, notably NMO stretch and offset dependent tuning, have been addressed However, it is important to remember that there are other factors that can bias or distort the three-term AVO inversion such as offset dependent transmission losses and residual NMO. There are solutions for these factors (Section 1.3), but it is important to remember that the results are sensitive to processing errors. In addition, it is important to remember that errors may be introduced by using an isotropic elastic model, if the media is actually anelastic or anisotropic.

The wavelet is assumed to be known as part of the AVO waveform inversion, but as shown in Chapter 6, some error is acceptable in its estimation. By including the wavelet, the band-limited reflectivity estimates are always smooth. This is not the case for traditional AVO inversion done on a sample-by-sample basis. Due to the ill-conditioned nature of the problem, the solution may oscillate quickly in time, introducing high frequency noise into the solution.

The approach taken in Chapter 3 was able to predict the uncertainty of the various reflectivity estimates. Predicting the uncertainty for the three-term AVO waveform inversion is more complex and expensive. Instead of inverting a 3×3 matrix a $3L \times 3L$ matrix must be inverted which is prohibitively expensive. In addition,

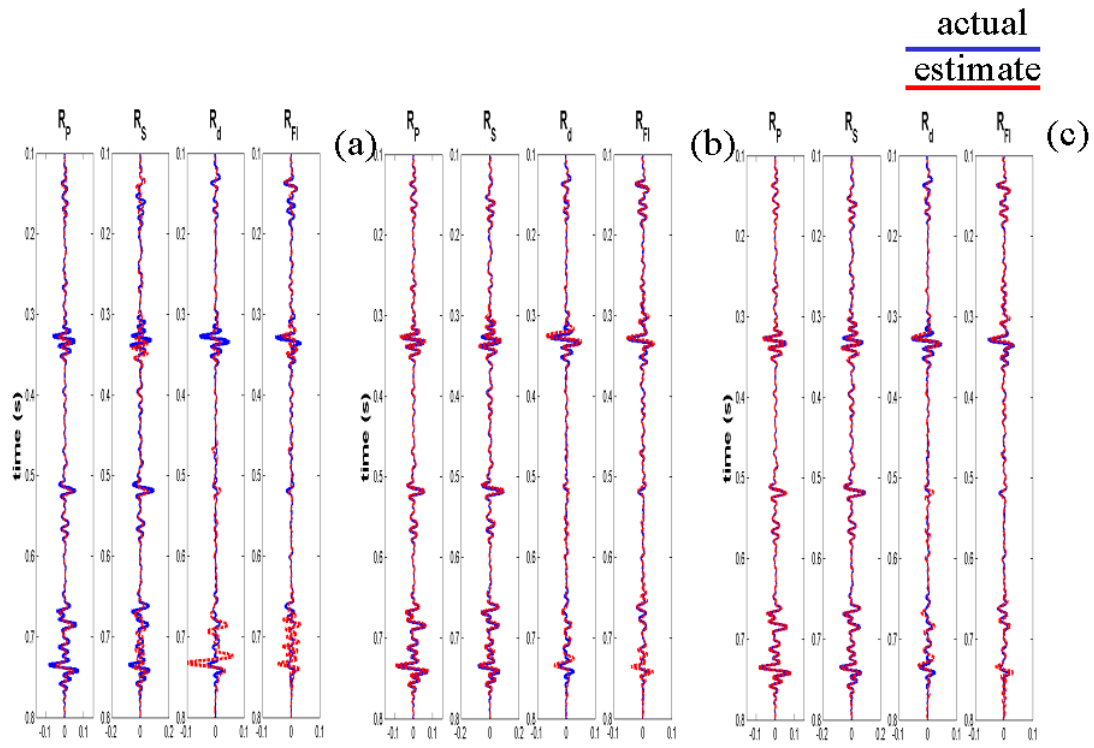

Figure 8.34: Comparison of the estimated P- and S-wave impedance, density, and fluid stack reflectivity (shown in red) with the ideal zero-offset reflectivity (shown in blue) based on the three-term AVO waveform inversion of Halfway model with statics. Statics left uncorrected degrades the inversion results (a). Fig (b) shows one pass of trim statics while Fig. (c) shows 3 passes. Note the improvement in the density reflectivity compared to (a) and how it compares favorably to Fig. 8.12a.

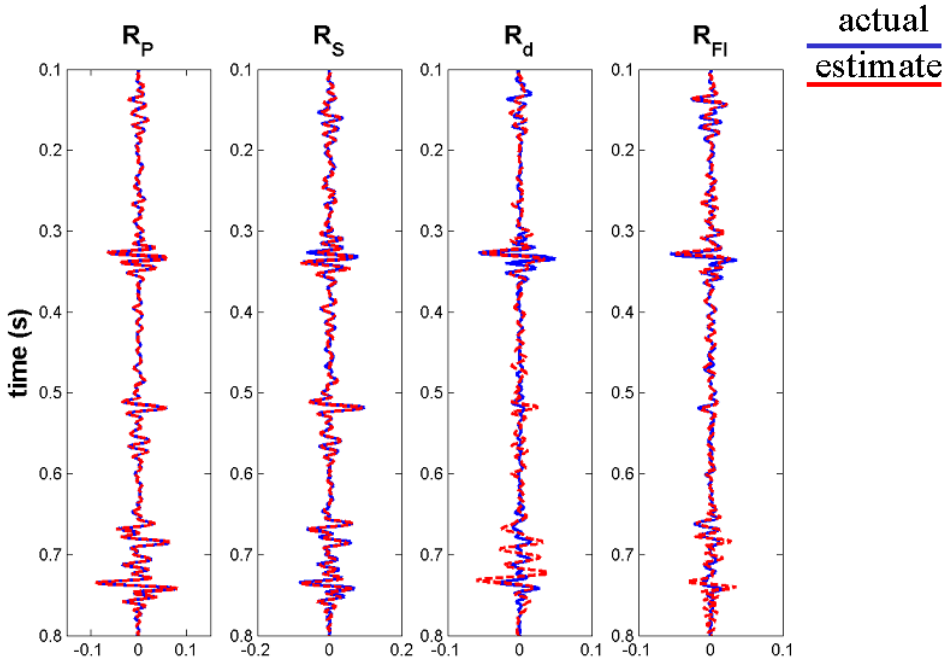

Figure 8.35: Comparison of the estimated (red) P- and S-wave impedance, density, and fluid stack reflectivity with the ideal (blue) zero-offset reflectivity based on the NMO corrected three-term AVO waveform inversion of Halfway model with one pass of trim statics. The result compares favorably with Fig. 8.34b.

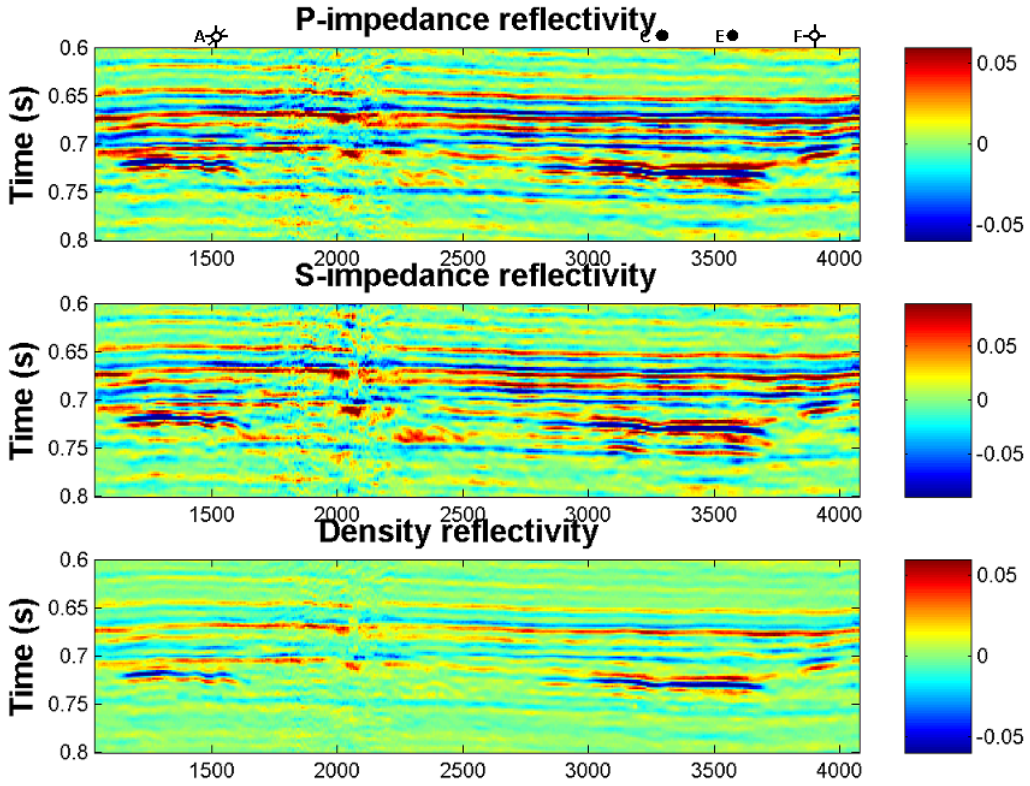

Figure 8.36: P- and S-wave impedance, and density reflectivity AVO waveform inversions with trim statics and Cauchy misfit weights.

the problem is being regularized by controlling the number of iterations used in the conjugate gradient solution. This has the effect of excluding the smallest eigenvalues, thus changing the condition number and uncertainty of the problem. Currently, my thought is to run the three-term inversion of Chapter 3 producing an uncertainty analysis that could be used to identify problem areas on the line. This uncertainty would only provide a rough guide to where potential problems might be, but it would identify problem areas particular to the AVO inversion itself.

Various misfits weights were tested on both the synthetic are real data. The ℓ^1 misfit function gave the best results on the model data set, but the results were sensitive to its parameterization. On the real data, the results from ℓ^1 misfit function were quite poor, perhaps suggesting a parameterization problem. However, both the the Bube and Langan weights, and Cauchy weights gave good results for both the synthetic and real data. Each of these weights seemed to be relatively insensitive to their parameterization. The Huber weights gave good results, were relatively insensitive to parameterization, but did not provide as good as results as the Bube and Langan weights, and Cauchy weights.

8.6 Conclusions

This chapter developed and demonstrated a three-term AVO waveform inversion that works on data with NMO stretch and offset dependent tuning. By including these factors, longer offsets may be used to help make the AVO problem better conditioned. Both misfit weights and constraints are used to regularize the problem helping to address the underdetermined nature of the problem and to stabilize the inversion in the presence of non-Gaussian noise. Incorporating the misfit weights improve the reflectivity estimates in the presence of non-Gaussian noise for both the synthetic and real seismic data. A methodology has been shown to allow for the identification and application of residual statics as part of the AVO waveform inversion. The input gathers can either be NMO corrected or left uncorrected. This first option allows prestack migrated gathers to be input into the algorithm. The algorithm is successfully demonstrated on synthetic data exhibiting NMO stretch and offset dependent tuning that would bias a traditional three term AVO inversion. Density reflectivity is successfully estimated on both synthetic and real seismic data even when density is uncorrelated with the velocity reflectivity for signal-to-noise ratios typical of real seismic data.

Chapter 9

Conclusions

This thesis demonstrates a number of methods to improve the estimates of AVO inversion. This is accomplished by implementing a series of theoretical improvements to the formulation of the inverse problem and including *a priori* probabilistic constraints based on local geologic knowledge. As importantly, a number of quality controls were also developed to help the explorationist understand the reliability and information content of the predictions.

In the first part of this thesis, a constrained three-term AVO inversion was developed following a Bayesian methodology. The three-term linear approximation used in this inversion introduces less bias than the two-term approximations typically practised by the exploration industry today. Two-term AVO inversions implicitly impose hard constraints either through the truncation of the third term, or the use of some external relationship such as the Gardner relationship. These hard constraints lead to biases. Instead of incorporating hard constraints, the three-term AVO inversion is made better conditioned by using probabilistic constraints, based on local geologic knowledge. By comparing the probabilistic *a priori* information (local geologic control) and the probabilistic data misfit function, Bayes' theorem estimates the optimal solution, with the least amount of bias and error

The degree to which these constraints influence the solution is dependent on the signal-to-noise ratio. The lower the ratio, the greater the influence of the constraints. Secondly, the acquisition geometry influence the relative contribution of the data and the *a priori* information. If the data are poorly conditioned, for example, if the all the far angle information is missing, then the *a priori* information contributes more to the solution.

The three-term AVO inversion bears some similarities to Lortzer and Berkhout

(1993) though there are significant differences in the derivation. Lortzer and Berkhout treat the geologic control as data constraints while I treat the information as *a priori* information. Further, Lortzer and Berkhout never discuss how the parameter covariance matrix is constructed. I show in detail how this is done and discuss some of the implications of the underlying relationships. For example, if the background α/β ratio changes as a function of depth, this implies the S-wave impedance reflectivity is non-stationary.

The Bayesian framework, used to develop the AVO inversion, is probabilistic in nature providing a convenient framework for the uncertainty analysis. The most likely estimate corresponds to the largest probability while the uncertainty is related to the probability distribution's dispersion. The standard deviation measures the dispersion for Gaussian misfit functions and can be calculated analytically. Estimates of uncertainty based on these analytic relationships were tested and verified by a modeling study, performed in Chapter 2. Further, it is possible to see whether the reflectivity estimates provide information independent from the constraints by observing the ratio of the unconstrained to constrained uncertainty. If the ratio is large, then most of the information for that reflectivity attribute is coming from the constraints.

This is particularly important for the density reflectivity, perhaps the most difficult reflectivity attribute to estimate reliably. Provided sufficient angles and signal-to-noise ratio, it is possible to estimate density reflectivity accurately. This was demonstrated with both synthetic and real data examples in Chapter 3. Accurate estimates of density reflectivity are quite prized by the explorationist for their ability help discriminate noncommercial gas from commercial gas. Estimating density reflectivity, along with the corresponding reliability displays is one of the major accomplishments of this thesis. The approach differs from Kelly and Skidmore (2001) in three respects. First, I use a linear rather than a nonlinear approximation. I do this based on philosophical grounds. Since higher order reflectivity terms, such as interbed multiples, are ignored in AVO inversion, it does not seem appropriate to include higher order reflectivity terms in the Zoeppritz equations approximation. Secondly, Kelly and Skidmore (2001) do not discuss how they constrain the problem, if they do so at all. Finally, they provide no mechanisms to quality control the estimates.

The three-term AVO inversion can be output in terms of a number of different

reflectivity attributes. The three-term AVO inversion is parameterized in terms of P-wave velocity reflectivity, S-wave velocity reflectivity, and density reflectivity, but through the use of transform matrix (equation B.54) the reflectivity attributes may be output in terms of any of the attributes common in the literature. In addition, the uncertainty associated with the reflectivity may also be transformed. In this way, the reflectivity attributes which best highlight the desired anomaly may be chosen, while still maintaining acceptable error. For example, impedance reflectivity might be chosen over velocity reflectivity, since it shows similar geologic information, but with less error.

The uncertainty estimates performed in this thesis are based on the assumption that the noise is Gaussian. If there are systematic data errors or theoretical errors the uncertainty estimates will be too optimistic. This necessitates that (1) the seismic data are properly conditioned prior to the AVO inversion, (2) the assumptions made are honored, and (3) theoretical error associated with the linear operator is random. In Chapter 1, it was assumed that the earth was elastic and isotropic. If either of these assumptions are incorrect, then the Zoeppritz equations need to be modified to reflect that reality. For VTI media, the linearized approximation of the Zoeppritz equations lead to an underdetermined set of equations.

Theoretical error associated with the linear operator was investigated in Chapter 4. Analytic relationships and modeling studies showed that random errors in the α/β ratio and offset-to-angle mapping lead to second order random error, which is typically negligible in comparison to the error arising from random noise. On the other hand, systematic error within the data leads to systematic error in the estimated reflectivity attributes. The size of this error is dependent upon the parameterization. The error usually shows up in the variable associated with the second or third term, for example the gradient or S-wave impedance reflectivity. Generally there is little or no error in the variable associated with the first term (P-wave impedance reflectivity). Systematic errors in the macro P-wave velocity field lead to systematic errors in the gradient and associated variables. Systematic errors in the α/β ratio lead to errors in the S-wave impedance reflectivity and associated variables.

Errors arising from the preconditioning of the seismic data and their solutions was investigated in the last part of this thesis, specifically, the effect of NMO stretch and offset dependent tuning on the AVO inversion. These two effects lead to distortions, although AVO inversion on NMO-corrected (and NMO-stretched) gathers

is actually surprisingly robust. The distortion primarily shows up on the secondary AVO reflectivity attribute, such as the gradient or S-wave impedance reflectivity. The attribute associated with the first term, the intercept or P-wave impedance reflectivity is unaffected. Further, there is a surprisingly large group of geologic interfaces for which NMO stretch and offset dependent tuning do not distort the reflectivity estimates. Regional reflectors from interfaces between clastics following the mudrock trend are not distorted. Class I and II gas sands are not distorted. Only Class III and IV gas sand anomalies are distorted. Further, these distortions are only significant when large angles are used. When done with angles less than 30^0, the AVO inversion estimates performed on the synthetics shown here had insignificant error. Only when larger angles were used (for example, 45^0) were the errors significant.

These distortions may be avoided by performing stretch-free NMO prior to the AVO inversion or by incorporating the NMO inversion into the AVO inversion. The best results were obtained by AVO waveform inversion but at a significant extra cost. Stretch-free NMO shows some promise for helping precondition the data prior to AVO. The advantages of using these algorithms have to be weighed against the cost of performing them.

A new stretch-free NMO algorithm was developed and tested in Chapter 7. Modeling studies performed on synthetic seismic data showed that the new approach reduced the amount of error in the AVO inversion estimates. The approach is an extension of the approach outlined by Claerbout (1992) with the addition of high-resolution constraints (Sacchi and Ulrych, 1995). The high-resolution constraints improve the results for band-limited data. The approach is also a variation of the theory developed for AVO waveform inversion. The cost of performing this is an order of magnitude faster than performing a hyperbolic Radon transform.

The best results were obtained with AVO waveform inversion. By incorporating the waveform, the NMO operator, the AVO operator and suitable constraints, the non-uniqueness due to the band-limited nature of the seismic data and NMO operator may be addressed. The offset dependent tuning and NMO stretch are built into the forward operators and so are addressed as part of the inversion.

The problem was regularized in two ways, explicitly with the *a priori* constraints and secondly by the numerical solution excluding small eigenvalues. In order to implement the high-resolution constraints, first the problem had to be transformed to an independent set of variables. The parameter covariance matrix developed for the

three-term AVO inversion in Chapter 3 was used for this purpose. The covariance matrix describes the inter-relationships between the variables. The eigenvectors of this covariance matrix were then used to transform the parameters to an independent set of variables. This transformation is similar to the transformation Simmons and Backus (1996) performed. They solved for P-wave impedance reflectivity, fluid stack reflectivity and a third reflectivity variable. This third variable is the difference between the density predicted by the Gardner relationship and the actual density. This last variable they found to be unstable and did not trust the estimates. Simmons and Backus make the problem better posed by reducing the number of layers estimated. For example instead of estimating elastic parameters for each time sample, they might do this for every two or three time samples. This leads to layers with fixed isochrons.

In contrast, I use a bootstrap approach to identify the major interfaces. The number of interfaces and thickness of layers is variable and depend on the bootstrap procedure and its parameterization. The use of long-tailed distributions lead to weighting functions which create small weights for large reflectivity and large weights for small reflectivity. The solution is biased towards reflectivity with small weights. The eigenvalues of the variables also influence the size of weights where the eigenvectors are similar to the Simmons and Backus parameterization. Variables associated with the third eigenvalue (density reflectivity) have in general much larger weights than variables associated with the first eigenvalue (P-wave impedance reflectivity), thus there is a bias against large density reflectivity. The bootstrapping procedure produces a sparse spike reflectivity series. This is similar to the poststack impedance inversion algorithms of Sacchi (1997) and Trad (2002) extending the poststack impedance inversion to prestack data. However, I was only able to perform the impedance parameterization on synthetic data. This is an area for future work.

The approach of Buland and Omre (2003) also perform AVO waveform inversion using a linearized approximation, and constraints similar to what I used in Chapter 3. Buland and Omre (2003) assume Gaussian statistics, however, they assume that the reflectivity from different time samples are correlated. I assumed the opposite. This leads to an easier calculation of the eigenvectors, but their assumption has some merit and I would like to explore it further in the future. They treat every time sample as a layer whose elastic properties are estimated as part of the solution. The correlation between different layers is used as additional information to make the problem better

posed. The Buland and Omre (2003) approach, however, does not consider NMO stretch and offset dependent tuning.

Like Sacchi (1997) and Trad (2002), I solved the inverse problem using conjugate gradient, calculating the linear operator on the fly. The algorithm consisted of two nested loops. The outer loop estimates the constraint weights and the penalty term. The penalty term controls the relative weighting of the constraints versus the seismic data. The inner loop solves the inverse problem using conjugate gradient, for which the maximum number of iterations can be used as regularization parameter excluding small eigenvalues from the problem. I created both two- and three-term implementations of the AVO waveform inversion. The two-term inversion is a better conditioned problem and the results of the algorithm are more stable. The three-term parameterization requires that the maximum number of conjugate gradient iterations be quite restricted in order to get a stable result. Restricting the number of conjugate gradient iterations leads to solutions which are substantially faster than solving the full inverse problem.

Both the two- and three-term versions were tested on synthetic and real data examples. On synthetic data, the AVO waveform inversion made the best estimates of events undergoing NMO stretch and tuning of all the methods discussed here. On real seismic data, the three-term AVO waveform inversion was able to distinguish commercial from noncommercial gas. The AVO waveform inversion's S-wave impedance reflectivity was significantly better than the three-term AVO inversion's. The continuity of events for the three-term AVO waveform inversion was significantly better than two-term AVO inversion. I believe this improvement was due to the classic trade-off between resolution and variance. With the AVO waveform inversion a few isolated reflectors are imaged with less variance than estimating reflectors on every time sample as is the case with conventional AVO inversion.

The AVO waveform inversion was also implemented with long-tailed misfit distributions. On synthetic data, with a simulated multiple, this formulation was found to give superior results compared to a Gaussian misfit function. Of the long-tailed distributions tested, on real and synthetic data, the Cauchy distribution seemed to give the best results while being relatively insensitive to its parameterization.

Lastly, the effect of uncorrected statics was investigated. These degraded the AVO estimates. A method was shown were the estimated model from the AVO inversion could be used as a model for trim statics. This provided for a model that kinematics,

character, and amplitude could change as a function of offset. The approach worked well on synthetic data, but showed little improvement on real data.

9.1 Future work

Throughout this thesis, it was assumed that the kinematics were known prior to the AVO inversion. However, if residual NMO is present, this will distort the AVO estimates (Spratt, 1987). Further, Spagnolini (1994) pointed out that, correcting the residual NMO on the basis of the AVO estimates is difficult because of interrelatedness of the parameters. He showed that Radon parameterization (τ, p) is mathematically equivalent to the two-term Shuey parameterization (A, B). Swan (2001) seems to have gotten around the underdetermined nature of the problem by introducing extra information. Namely, he made the AVO reflectivity estimates follow the mudrock trend over some time window. I would like to pursue this general idea for the AVO waveform inversion.

Secondly, throughout this thesis, it was assumed that the geologic structure was relatively simple, so that effects normally dealt with by the migration could be ignored. To address this shortcoming the Kirchhoff operator rather than NMO operator could be used to formulate the problem. This makes the problem much larger and more expensive to solve, but more theoretically rigorous. Feng and Sacchi (2004) have done some preliminary work trying to develop this approach. I would like to pursue this.

The AVO inversion could also be modified to perform joint PP and PS AVO inversion. One of the big issues of joint AVO inversion is that the PP and PS data have different wavelets. This needs to be dealt with by the AVO inversion. AVO waveform inversion provides a convenient framework to do this.

Lastly I would like to continue to work on the problem of estimating impedances from the reflectivity on real seismic data. I was able to do this on synthetic data, but had problems with real data due to the lack of lateral continuity of the sparse spike estimates. To address this, I would like to implement spatial constraints. Again, this would increase the size and cost of the problem.

9.2 Software developed

In performing the research for this thesis, I wrote a series of algorithms in MATLAB, including algorithms for:

- Three-term AVO inversion
- Stretch-free NMO
- Two and three-term AVO waveform inversion
- Various utilities (for example, to help construct 3 term covariance matrix)
- Modeling software

  Some of the algorithms that I used frequently, I subsequently coded in FORTRAN 95.

References

Aki, K., and Richards P.G., 1980, Quantitative seismology: theory and methods: W.H. Freeman and Co.

Aki, K., and Richards P.G., 2002, Quantitative seismology: second edition: University Science Books.

Alkhalifah, T., and Tsvankin, I., 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550-1566.

Allen, J.L., and Peddy, C.P., 1993, Amplitude variation with offset / Gulf Coast Case histories: Soc. Expl. Geophys.

Avseth, P., 2000, Combining rock physics and sedimentology for seismic reservoir characterization of North Sea Turbidite systems: Ph.D. Thesis, Standford University.

Backus, G.E., 1962, Long-wave elastic anisotropy produced by horizontal layering, Journal of Geophysical Research, **67**, No. 11, 4427-4440

de Bazelaire, E., 1988, Normal moveout revisited - Inhomogeneous media and curved interfaces: Geophysics, Soc. of Expl. Geophys., 53, 143-157.

Berkhout, A. J., 1997, Pushing the limits of seismic imaging, part II: Integration of prestack migration, velocity, estimation, and AVO analysis: Geophysics, **62**, 954-969.

Bleistein, N., Cohen, J.K., and Stockwell, J.W., 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer-Verlag.

Borse G.J., 1997, Numerical methods with MATLAB: PWS Publishing company.

Bortfeld, R., 1962, Exact solution of the reflection and refraction of arbitrary spherical compressional waves at liquid-liquid interfaces and at solid-solid interfaces with equal shear velocities and equal densities: Geophys. Prosp., **10**, 35-67.

Brandsberg-Dahl, S., de Hoop, M. V., and Ursin, B., 2003, Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers: Geophysics, **68**, 232-254.

Bube, K. P., and Langan, R. T., 1997, Hybrid ℓ^1/ℓ^2 minimization with applications to tomography: Geophysics, **62**, 1183-1195.

Buland, A., and Omre, H., 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185-198.

Cambois, G., and Stoffa, P. L., 1992, Surface-consistent deconvolution in the log/Fourier domain: Geophysics, **57**, 823-840.

Castagna, J.P., Batzle, M.L., Eastwood, R.L., 1985, Relationships between compressional-wave and shear wave velocities in clastic silicate rocks: Geophysics, **50**, 571-581.

Castagna, J.P., Batzle, M.L.,and Kan, T.K., 1993, Rock Physics – The link between rock properties and AVO response, *in* Castagna, J.P., and Backus, M. M., Eds., Offset-dependent reflectivity–theory and practice of AVO analysis: Soc. Expl. Geophys., 135-171.

Castagna, J. P., and Swan, H. W., 1997, Principles of AVO crossplotting: The Leading Edge, **16**, no. 04, 337-342.

Castle, R. J., 1994, Theory of normal moveout: Geophysics, **59**, 983-999.

Carazzone, J.J., and Srnka, L.J., 1993, Elastic inversion of Gulf of Mexico data *in* Castagna, J.P., and Backus, M. M., Eds., Offset-dependent reflectivity–theory and practice of AVO analysis: Soc. Expl. Geophys., 303-313.

Carcione, J. M., and Gangi, A. F., 2000, Gas generation and overpressure: Effects on seismic attributes, Geophysics, **65**, 1769-1779.

Cary, P. W. and Lorentz, G. A., 1993, Four-component surface-consistent deconvolution: Geophysics, **58**, 383-392.

Červeny, V., 1987, Ray methods for three-dimensional seismic modelling: The Norwegian Inst. of Tech.

Červeny, V., 2001, Seismic ray theory: Cambridge University Press.

Claerbout, J. F., 1973, Fundamentals of geophysical data processing: Blackwell Scientific Publications.

Claerbout, J. F., 1986, Imaging the earths's interior: Blackwell Scientific Publications.

Claerbout, J. F., 1992, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.

Connolly, P., 1999, Elastic impedance: The Leading Edge, **18**, no. 4, 438-452.

Crase, E., Pica, A., Noble, M., McDonald, J. and Tarantola, A., 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527-538.

Daley, P. F., and Hron, F., 1977, Reflection and transmission coefficients for transversely isotropic media: Bull. Seis. Soc. Am., 661-675.

Dantzig, G.B., 1963, Linear programming and extensions: Princeton University Press.

Debeye, H. W. J., and van Riel, P., 1990, LP-norm deconvolution: Geophys. Prosp., **38**, 381-404.

Debski, W. and Tarantola, A., 1995, Information on elastic parameters obtained from the amplitudes of reflected waves: Geophysics, **60**, 1426-1436.

Domenico, S. N., 1976, Effect of brine-gas mixture on velocity in an unconsolidated sand reservoir: Geophysics, **41**, 882-894.

Dong, W., 1996, Fluid line distortion due to migration stretch, 66th Ann. Internat. Mtg.: Soc. of Expl. Geophys., 1345-1348.

Dong, W., 1998, AVO detectability against tuning and stretching artifacts, 68th Ann. Internat. Mtg.: Soc. of Expl. Geophys., 236-239.

Dong, W., 1999, AVO detectability against tuning and stretching artifacts: Geophysics, **64**, 494-503.

Downton, J., and Chaveste, A, 2004, Calibrated three-term AVO to estimate density and water saturation: 2004 CSEG Convention Expanded Abstracts.

Downton, J., Guan, H., and Somerville, R., 2003, NMO, AVO and Stack: 2003 CSPG / CSEG Convention Expanded Abstracts.

Downton, J., and Lines, L., 2001a, Constrained three parameter AVO inversion and uncertainty analysis: 71st Ann. Internat. Mtg.: Soc. of Expl. Geophys., 251-254.

Downton, J., and Lines, L., 2001b, AVO feasibility and reliability analysis: CSEG Recorder, Vol. **26**, No 6, p 66-73.

Downton, J., and Lines, L., 2001c, AVO analysis at Pikes Peak: 2001 CSEG Convention Expanded Abstracts.

Downton, J., and Lines, L., 2002, AVO before NMO: 72nd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 233-236.

Downton, J., and Lines, L., 2003, High-resolution AVO analysis before NMO: 73rd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 219-222.

Downton, J., and Mackidd, D., 1997, Methods to improve A.V.O. extractions: 2004 CSEG Convention Expanded Abstracts, 144-145.

Downton, J., Russell, B.H., and Lines, L, 2000, AVO for Explorationists: Pitfalls and Solutions, CREWES 2000 Research Report, 577-598.

Downton, J., and Tonn, R., 1997, Reservoir quality mapping with geostatistical and neural network techniques: 1997 S.E.G. Production and Development Forum Abstracts.

Dunkin, J. W., and Levin, F. K., 1973, Effect of normal moveout on a seismic pulse: Geophysics, **38**, 635-642.

Eberhart-Phillips, D., Han, D-H., and Zoback, M. D., 1989, Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone: Geophysics, **54**, 82-89.

Fatti, J.L., Smith, G.C., Vail, P.J., Strauss, P.J., and Levitt, P.R., 1994, Detection of gas in sandstone reservoirs using AVO analysis: A 3-D seismic case history using the Geostack technique: Geophysics, **59**, 1362-1376.

Feng, J., and Sacchi, M.D., 2004, High-resolution regularized least squares AVA Kirchhoff Migration, 74th Ann. Internat. Mtg.: Soc. of Expl. Geophys.

Gardner, G. H. F., Gardner, L. W. and Gregory, A. R., 1974, Formation velocity and density - The diagnostic basics for stratigraphic traps: Geophysics, **39**, 770-780.

Gassmann, F., 1951. Uber die elastizitat poroser medien. vier der natur. gesellshaft in Zurich, **96**, 1-23.

Gersztenkorn, A., Bednar, J. B., and Lines, L. R., 1986, Robust iterative inversion for the one-dimensional acoustic wave equation: Geophysics, **51**, 357-368.

Gidlow, P.M., Smith, G. C., and Vail, P. J., 1992, Hydrocarbon detection using fluid factor traces, a case study: How useful is AVO analysis?: Joint SEG/EAEG summer research workshop, Technical Program and Abstracts, 78-89.

Goodway, B., Chen, T., and Downton, J., 1997, Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters: 67th Annual International SEG meeting, Expanded Abstracts, 183-186.

Goupillaud, P. L., 1961, An approach to inverse filtering of near-surface layer effects from seismic records: Geophysics, **26**, 754-760.

Gouveia, W., and Scales, J.A., 1998, Bayesian seismic waveform inversion, Jour. Geophy. Res., **103**, 2759-2779.

Gray, D., Goodway, B. and Chen, T., 1999, Bridging the gap - using AVO to detect changes in fundamental elastic constants: 61st Mtg.: Eur. Assn. Geosci.

Gray, D., 2003, P-S converted-wave AVO: 73rd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 165-168.

Greenberg, M. L. and Castagna, J. P., 1992, Shear-wave velocity estimation in porous rocks: Theoretical formulation, preliminary verification and applications : Geophys. Prosp., **40**, 195-210.

Grion, S., Mazzotti, A. and Spagnolini, U., 1998, Joint estimation of AVO and kinematic parameters: Geophys. Prosp., **46**, 405-422.

Han, D. H., 1986, Effects of porosity and clay content on acoustic properties of sandstones and unconsolidated sediments: Ph.D. Thesis, Standford University.

Han, D. H., Nur, A., and Morgan, D., 1986, Effects of porosity and clay content on wave velocities in sandstones : Geophysics, **51**, 2093-2107.

Hansen, P. H., 1998, Rank-deficient and discrete ill-posed problems: SIAM.

Hargreaves, N. D., and Calvert, A. J., 1991, Inverse Q filtering by Fourier transform: Geophysics, **56**, 519-527.

Hicks, G.J., 2001, Removing NMO stretch using the Radon and Fourier-Radon transforms: 63rd Mtg.: Eur. Assn. Geosci. Eng., Session: A-18.

Houck, R. T., 2002, Quantifying the uncertainty in an AVO interpretation: Geophysics, **67**, 117-125.

Huber, P.J, 1977, Robust Statistical Procedures: SIAM.

Hunt, L., Trickett, S., Levesque, D., McKenny, P., and Jamieson, S., 2003, The effect of stretch-free stacking on a clastic exploration play in Alberta, Canada, 73rd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 321-324.

Jin, S., Cambois, G., and Vuillermoz, C., 2000, Shear-wave velocity and density estimation from PS-wave AVO analysis: Application to an OBS dataset from the North Sea: Geophysics, **65**, 1446-1454.

Kelly, M. and Skidmore, C., 2001, Non-linear AVO equations and their use in 3-parameter inversions: 71st Ann. Internat. Mtg: Soc. of Expl. Geophys., 255-256.

Kennett, B.L.N., 1984, Seismic wave propagation in stratified media: Cambridge University Press.

Kuehl, H., and Sacchi, M., 2002, Robust AVP estimation using least-squares wave-equation migration, 72nd Ann. Internat. Mtg: Soc. of Expl. Geophys., 281-284.

Larsen, J., Margrave, G., and Lu, H. -X., 1999, AVO analysis by simultaneous P-P and P-S weighted stacking applied to 3-C–3-D seismic data, 69th Ann. Internat. Mtg: Soc. of Expl. Geophys., 721-724.

Landro, M., 2001, Discrimination between pressure and fluid saturation changes from time-lapse seismic data: Geophysics, **66**, 836-844.

Levy, S., and Fullagar, P. K., 1981, Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution: Geophysics, **46**, 1235-1243.

Li, Y., Goodway, B., and Downton, J., 2003, Recent applications of AVO to carbonate reservoirs in the Western Canadian Sedimentary Basin: The Leading Edge, **22**, no. 7, 670-674.

Lin, T. L., and Phair, R., 1993, AVO tuning: 63rd Ann. Internat. Mtg: Soc. of Expl. Geophys., 727-730.

Lindseth, R. O., 1979, Synthetic sonic logs - A process for stratigraphic interpretation: Geophysics, **44**, 3-26.

Lines, L. R. and Treitel, S., 1985, Wavelets well-logs and Wiener filters: First Break, **3**, no. 8, 9-14.

Lortzer, G.J.M., and Berkhout, A.J., 1993, Linearized AVO inversion of multicomponent seismic data *in* Castagna, J.P., and Backus, M. M., Eds., Offset-dependent reflectivity–theory and practice of AVO analysis: Soc. Expl. Geophys., 303-313.

Loures, L., and Moraes, F., 2002, Reservoir porosity inference, 72nd Ann. Internat. Mtg: Soc. of Expl. Geophys., 1809-1812.

Marfurt, K. J., Schneider, R. V., and Mueller, M. C., 1996, Pitfalls of using conventional and discrete Radon transforms on poorly sample data: Geophysics, **61**, 1467-1482.

Mavko, G., and Jizba, D., 1991, Estimating grain-scale fluid effects on velocity dispersion in rocks: Geophysics, **56**, 1940-1949.

Mavko, G., Mukerji, T., and Dvorkin, J., 1998, The rock physics handbook: Cambridge University Press.

Mazzotti, A., 1990, Prestack amplitude analysis methodology and its application to seismic bright spots in the Po valley, Italy: Geophysics, **55**, 157-166.

Menke, W., 1984, Geophysical data analysis: discrete inverse theory: Academic Press Inc.

Mosher, C. C., Keho, T. H., Weglein, A. B. and Foster, D. J., 1996, The impact of migration on AVO: Geophysics, **61**, 1603-1615.

Nocedal, J., and Wright, S.J., 1999, Numerical Optimization: Springer-Verlag.

O'Doherty, R. F., and Anstey, N. A., 1971, Reflections on amplitudes: Geophys. Prosp., **19**, 430-458.

Oldenburg, D. W., Scheuer, T. and Levy, S., 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, **48**, 1318-1337.

Ostrander, W. J., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles-of-incidence: Geophysics, **49**, 1637-1648.

Peterson, R. A., Fillippone, W. R., and Coker, F. B., 1955, The synthesis of seismograms from well-log data: Geophysics, **20**, 516-538.

Potter, C. C., Dey, A.K., and Stewart, R..R., 1998, Density prediction using P- and S-wave sonic velocities: Geotriad, 1998 CSPG, CSEG, CWLS Joint convention.

Richards, P. G., and Frasier, C. W., 1976, Scattering of elastic waves from depth-dependent inhomogeneities: Geophysics, **41**, 441-458.

Rickett, J., and Sava, P., 2001, Offset and angle domain common-image gathers for shot-profile migration: 71st Ann. Internat. Mtg: Soc. of Expl. Geophys., 1115-1118.

Riedel, M., Dosso, S. E., and Beran, L., 2003, Uncertainty estimation for amplitude variation with offset (AVO) inversion: Geophysics, **68**, 1485-1496.

Ronen, S., and Liner, C. L., 2000, Least-squares DMO and migration: Geophysics, **65**, 1364-1371.

Ross, C. P., and Beale, P. L., 1994, Seismic offset balancing: Geophysics, **59**, 93-101.

Ruger, A., 2002, Reflection coefficients and azimuthal AVO Analysis in anisotropic media: SEG geophysical monograph series number 10: Soc. Expl. Geophys.

Rupert, G. B., and Chun, J. H., 1975, The block move sum normal moveout correction: Geophysics, **40**, 17-24.

Rutherford, S.,R., and Williams, R.,H., 1989, Amplitude-versus-offset variations in gas sands: Geophysics, **54**, 680-688.

Sacchi, M. D., 1997, Reweighting strategies in seismic deconvolution: Geophys. J. Int., **129**, 651-656

Sacchi, M. D., 1999, Statistical and transform methods for seismic signal processing: Department of Physics, University of Alberta (available at http://cm-gw.phys.ualberta.ca/~sacchi/saig/publ.html).

Sacchi, M. D. and Ulrych, T.D., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169-1177.

Scales, J. A., and Gersztenkorn, A., 1988, Robust methods in inverse theory: Inverse Problems, **4**, 1071–1091.

Sen, M., and Stoffa P.L., 1995, Global optimization methods in geophysical inversion: Elsevier Science Publishers.

Shapiro, S.A., and Hubral, P., 1999, Elastic Waves in Random media: Springer-Verlag.

Shen, F., Sierra, J., Burns, D.R., and Toksoz, M. N., 2002, Azimuthal offset-dependent attributes applied to fracture detection in a carbonate reservoir: Geophysics, **67**, 355-364.

Sheriff, R. E., 1975, Factors affecting seismic amplitudes: Geophys. Prosp., **23**, 125-138.

Sheriff, R. E., and Geldart, L.P., 1982, Exploration seismology: Volume 1: Cambridge University Press.

Shewchuck, J., 1994, An introduction to the conjugate gradient method without the agonizing pain: School of Computer Science. Carnegie Mellon University. Pittsburg. (available at http://www-2.cs.cmu.edu/~jrs/jrspapers.html).

Shuey, R.,T., 1985, A simplification of the Zoeppritz equations: Geophysics, **50**, 609-614.

Simmons, J. L., Jr. and Backus, M. M., 1994, AVO modeling and the locally converted shear wave: Geophysics, **59**, 1237-1248.

Simmons, J. L., Jr. and Backus, M. M., 1996, Waveform-based AVO inversion and AVO prediction error: Geophysics, **61**, 1575-1588.

Sivia, D.S., 1996, Data Analysis, A Bayesian tutorial: Oxford University Press.

Smith, G.,C., and Gidlow, P.M., 1987, Weighted stacking for rock property estimation and detection of gas: Geophysical Prospecting, **35**, 993-1014.

Smith, G., and Gidlow, P., 2003, The fluid factor angle and the crossplot angle: 73rd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 185-188.

Soroka, W.L., Fitch, T.J., Van Sickle, K.H., and North, P.D., 2002, Successful production application of 3-D amplitude variation with offset: The lessons learned: Geophysics, **67**, 379-390.

Spagnolini, U., 1994, Compound events decomposition and the interaction between AVO and velocity information: Geophys. Prosp., **42**, 241-259.

Spratt, S., 1987, Effect of normal moveout errors on amplitude versus offset-derived shear reflectivity, 57th Ann. Internat. Mtg: Soc. of Expl. Geophys.

Spratt R.S., Goins N.R., and Fitch T.J., 1993, Pseudo-shear–The analysis of AVO, *in* Castagna, J.P., and Backus, M. M., Eds., Offset-dependent reflectivity–theory and practice of AVO analysis: Soc. Expl. Geophys., 37-56.

Swan, H. W., 1993, Properties of direct AVO hydrocarbon indicators, *in* Castagna, J.P., and Backus, M. M., Eds., Offset-dependent reflectivity–theory and practice of AVO analysis: Soc. Expl. Geophys., 78-92.

Swan, H. W., 2001, Velocities from amplitude variations with offset: Geophysics, **66**, 1735-1743.

Takahashi, I., 2000, Quantifying information and uncertainty of rock property estimation from seismic data: Ph.D. Thesis, Standford University.

Taner, M. T. and Koehler, F., 1969, Velocity spectra - Digital computer derivation and applications of velocity functions: Geophysics, **34**, 859-881.

Taner, M. T. and Koehler, F., 1981, Surface consistent corrections: Geophysics, **46**, 17-22.

Taner, M. T., and Sheriff, R. E., 1977, Application of Amplitude, Frequency and Other Attributes to Stratigraphic and Hydrocarbon Determination *in* Payton, C. E., Seismic Stratigraphy – application to hydrocarbon exploration, Memoir 26: A.A.P.G., 302-327.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893-1903.

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

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954-1966.

Thomsen, L., 1988, Reflection seismology over azimuthally anisotropic media: Geophysics, **53**, 304-313.

Thore, P. D., de Bazelaire, E. and Ray, M. P., 1994, Three-parameter equation: An efficient tool to enhance the stack: Geophysics, **59**, 297-308.

Todoeschuck, J. P., Jensen, O. G. and Labonte, S., 1990, Gaussian scaling noise model of seismic reflection sequences: Evidence from well logs: Geophysics, **55**, 480-484.

Trad, D., 2002, Implementations and applications of the sparse Radon transform: Ph.D. Thesis, University of British Columbia.

Trickett, S., 2003, Stretch-free stacking: 73rd Ann. Internat. Mtg.: Soc. of Expl. Geophys., 2008-2011.

Tura, A. C. and Lumley, D., 1999, Estimating pressure and saturation changes from time-lapse AVO data: 69th Ann. Internat. Mtg: Soc. of Expl. Geophys., 1655-1658.

Tsvankin, I., 2001, Seismic signatures and analysis of reflection data in anisotropic media: Elsevier Science Publishers.

Ursenbach, C., 2002, Optimal Zoeppritz Approximations: 72nd Ann. Internat. Mtg: Soc. of Expl. Geophys., 1897-1900.

Ursenbach, C., 2004, A non-linear, three-parameter AVO method that can be solved non-iteratively: 2004 CSEG Convention Expanded Abstracts

Ursin, B., 1983, Review of elastic and electromagnetic wave propagation in horizontally layered media: Geophysics, **48**, 1063-1081.

Ursin, B., 1990, Offset dependent geometrical spreading in a layered medium: Geophysics, **55**, 492-496.

Ursin, B. and Stovas, A., 2002, Reflection and transmission responses of a layered isotropic viscoelastic medium: Geophysics, **67**, 307-323.

Van Koughnet, R. W., Skidmore, C. M., Kelly, M. C., and Lindsay, R., 2003, Prospecting with the density cube: The Leading Edge, **22**, no. 10, 1038-1045.

van Riel, P. and Berkhout, A. J., 1985, Resolution in seismic trace inversion by parameter estimation: Geophysics, **50**, 1440-1455.

Verm, R. and Hilterman, F., 1995, Lithology color-coded seismic sections: The calibration of AVO crossplotting to rock properties: The Leading Edge, **14**, no. 08, 847-853.

Verschuur, D. J., Berkhout, A. J. and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166-1177.

Walden, A. T., 1991, Making AVO sections more robust: Geophys. Prosp., **39**, 915-942.

Walden, A. T. and Hosken, J. W. J., 1986, The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution: Geophys. Prosp., **34**, 1038-1066.

Walden, A. T. and Hosken, J. W. J., 1988, Tutorial - Choosing the averaging interval when calculating primary reflection coefficients from well-logs: Geophys. Prosp., **36**, 799-824.

Walden, A.,T., and White, R.,E., 1984, On errors of fit and accuracy in matching synthetic seismograms and seismic traces; Geophysical Prospecting, **32**, 871-891

Wapenaar, C.P.A., and Berkhout, A.J., 1989, Elastic wave field extrapolation: Elsevier Science Publishers.

Westergren, R., 1999, Mathematics handbook for Science and Engineering: Springer-Verlag.

White, R. E., 1980, Partial coherence matching of synthetic-seismograms with seismic traces: Geophys. Prosp., **28**, 333-358.

Wiggins, R., Kenny, G.S., and Mclure, D.D., 1983, A method for determining and displaying the shear-velocity reflectivities for a geologic formation: European Patent Application 0113944.

Wijngarrden, A. J., and Berkhout, A. J., 1996, Resolution analysis on AVO attributes: SEG International Exposition and 66 annual meeting, 1711-1714.

Winslow, N.W., 2000, Joint Inversion using the convolutional model: Ph.D. Thesis, Rice University.

Appendix A

Gain corrections

A.1 Geometrical spreading correction

Ursin (1990), following Červeny (1987), shows for horizontally layered media, where the source and receiver are at the same level, that the geometrical spreading factor L is

$$L(\tau, p) = \frac{\cos\theta_0}{\alpha_0}\sqrt{\frac{h}{p\left|\frac{d\tau}{dp}\right|}}, \tag{A.1}$$

where θ_0 is the angle of emergence, α_0 is the interval P-wave velocity of the 1^{st} layer, h is the offset, τ is the offset-dependent travel time and p is the horizontal slowness. Note that this correction is a function of two way travel time and the ray parameter p. Implicitly, then, it is a function of the interval velocity field. This can be more easily seen if equation (1.14) is used to generate the ray parameter from the stacking velocity V_{stak}. In this case, the geometric spreading factor is

$$L = \frac{\tau^2 V_{stak}^2}{t_0\alpha_0}\sqrt{1 - \frac{\alpha_0^2 h^2}{\tau^2 V_{stak}^4}}, \tag{A.2}$$

where t_0 is the zero offset travel time. Note that in similar fashion to section 1.2.2 higher order approximations can also be calculated. Both equation (A.1) and (A.2) are a function of $\alpha(x, z)$. To illustrate the influence of this, Figure A.4 shows the geometric spreading correction based on equation A.1 at various CMP locations for the Halfway example in section 1.3.4. The interval velocity field used in the ray tracing is shown in Figure A.1. Figure A.2 shows the angle of incidence generated by the ray tracing for the same CMP gathers shown in Figure A.4. Figure A.3 shows

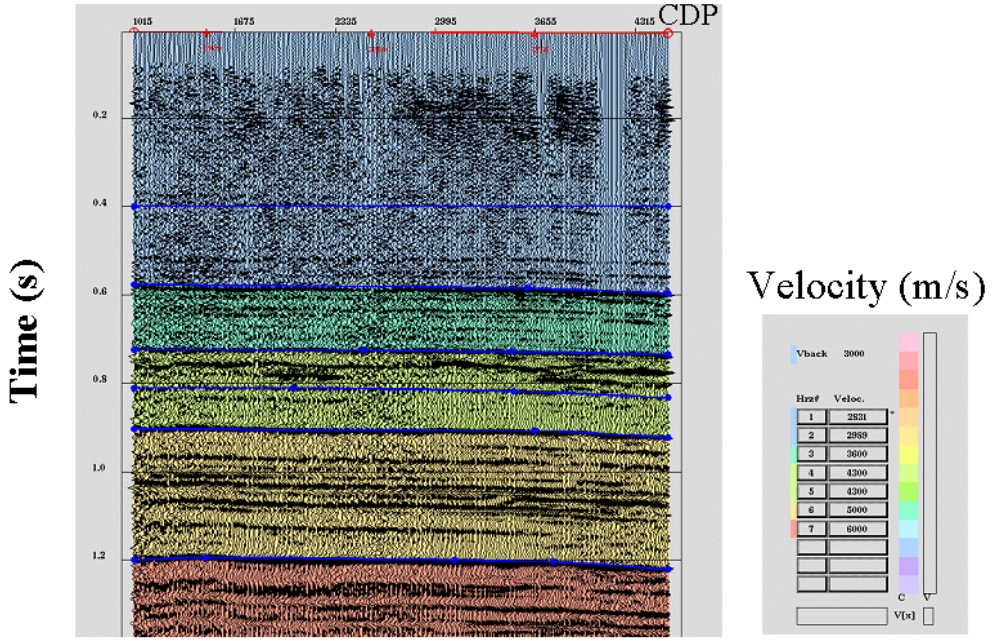

Figure A.1: Interval velocity model used for ray tracing and geometrical spreading correction for Halfway seismic line.

the geometrical divergence correction based on a t_0^2 correction after NMO correction. The (offset-variant) amplitude correction for a specific time sample is graphed at the top of both Figure A.4 and A.3. Note that the correction based on equation (A.1) scales the far offset significantly more than the one based on the t_0^2 correction. Further, the correction based on equation (A.1) includes the effects of refraction due to the interval velocity's variation with depth. Due to subtle changes in the velocity field the gain correction based on equation (A.1) changes spatially.

The line shown in Figure A.1 roughly honors the assumption of horizontal layering. For more complex cases, Červeny (2001) describes more complex corrections, but these really should be applied as part of a prestack migration.

A.2 Angle of emergence and free-surface correction

For land data, P-wave single component geophones measure only the vertical component of plane waves, but the plane waves usually arrive at the surface at some oblique angle. This can be corrected by an angle of emergence correction, which is simply a cosine correction. In actuality the correction is more complex since the

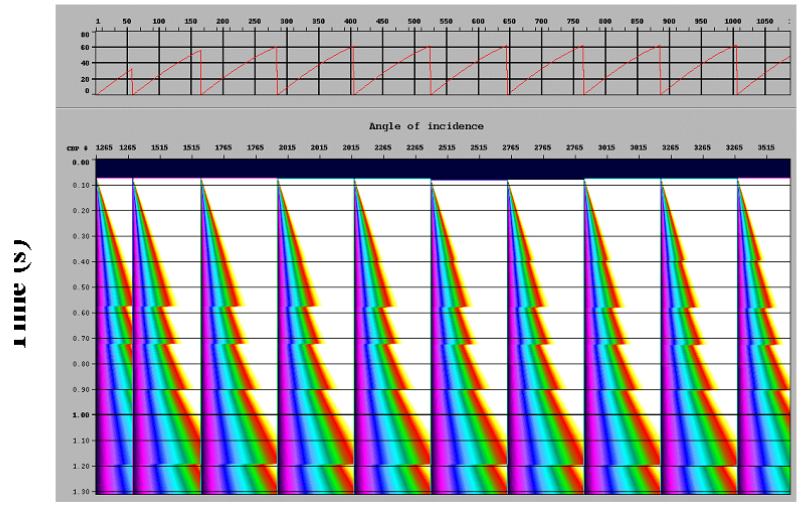

Figure A.2: Angle of incidence displayed for specific CMP gathers for the Halfway line. The data are first sorted by CMP and then by offset. Note the angle of inidence is graphed overhead for the zone of interest.

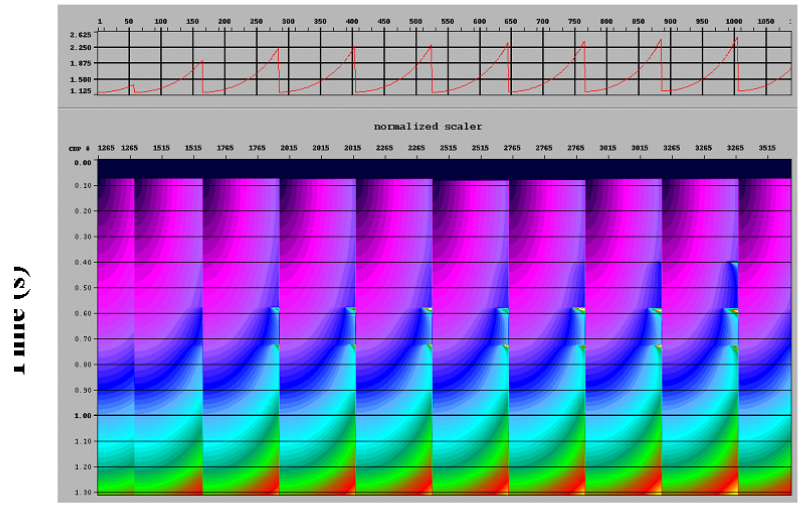

Figure A.3: Geometrical spreading correction for specific CMP gathers for the Halfway line. The data are first sorted by CMP and then by offset Note the scaling is graphed overhead for the zone of interest and note the lateral change in scaling due to lateral changes in velocity.

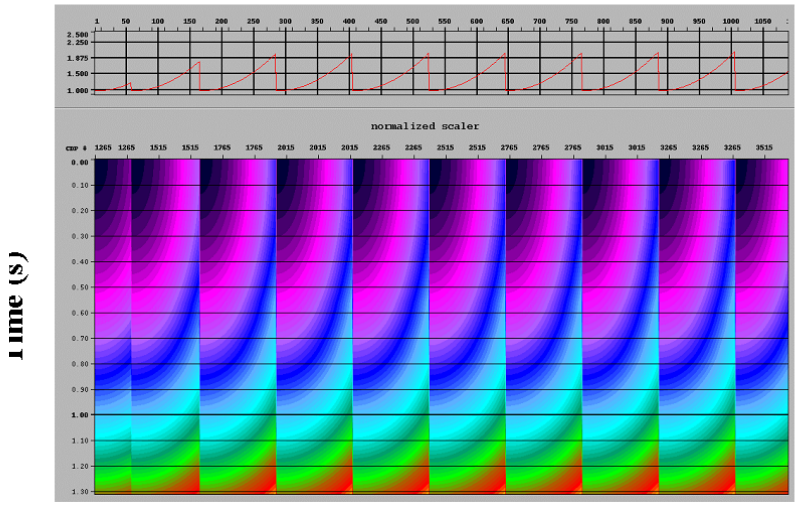

Figure A.4: Gain correction based on t^2 correction for specific CMP gathers for the Halfway line. The data are first sorted by CMP and then by offset.

motion is measured at a free surface. At the free surface the vertical motion Z that the geophone measures is the vector sum of the upgoing P-wave at incidence angle θ with unit amplitude, the downgoing reflected P-wave at angle θ with amplitude $-R_p$ and the downgoing reflected S-wave at angle ϕ with amplitude R_{sv}. In the case where the positive direction of the z-axis is chosen positive upwards

$$Z = \cos\theta - R_p \cos\theta + R_{sv} \sin\phi. \tag{A.3}$$

Using Snell's law, $\frac{\sin\theta}{\alpha_0} = \frac{\sin\phi}{\beta_0}$, equation (A.3) becomes

$$Z = (1 - R_p)\cos\theta + \frac{\beta_0}{\alpha_0} R_{sv} \sin\theta. \tag{A.4}$$

Using the P-wave reflection coefficient on the free surface (Aki and Richards (1980), equation 5.27)

$$R_p = \frac{-\left(\frac{1}{\beta_0^2} - 2p^2\right)^2 + 4p^2 \frac{\cos\theta}{\alpha_0}\frac{\cos\phi}{\beta_0}}{\left(\frac{1}{\beta_0^2} - 2p^2\right)^2 + 4p^2 \frac{\cos\theta}{\alpha_0}\frac{\cos\phi}{\beta_0}}, \tag{A.5}$$

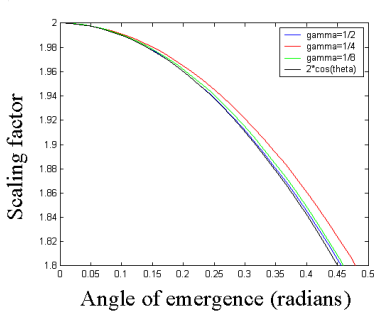

Figure A.5: Angle of emergence correction (black) plus free surface correction for different Vs/Vp ratios.

and the converted wave reflectivity R_{sv} at the free surface (Aki and Richards (1980), equation 5.28)

$$R_{SV} = \frac{4\left(\frac{\alpha_0}{\beta_0}\right) p \frac{\cos\theta}{\alpha_0} \left(\frac{1}{\beta_0^2} - 2p^2\right)}{\left(\frac{1}{\beta_0^2} - 2p^2\right)^2 + 4p^2 \frac{\cos\theta}{\alpha_0} \frac{\cos\phi}{\beta_0}}, \tag{A.6}$$

equation (A.3) becomes

$$Z = \frac{2\left(1 - 2\gamma^2 \sin^2\theta\right)\cos\theta}{\left(1 - 2\gamma^2 \sin^2\theta\right)^2 + 4\gamma^3 \sin^2\theta \cos\theta \cos\phi}. \tag{A.7}$$

Upon the substitution of several trigonometric identities this simplifies to

$$Z = \frac{2\cos\left(2\phi\right)\cos\theta}{\left(\cos 2\phi\right)^2 + 4\gamma \sin^2\phi \cos\theta \cos\phi}. \tag{A.8}$$

Equation (A.8) is a function of γ and the angle of emergence. Figure 41 shows that in practice γ has little influence on the correction. Further, Figure A.5 shows that including the free-surface effect makes little difference when compared to just compensating for a P-wave incident at an angle θ at the free surface.

A.3 Array correction

From Sheriff and Geldart (1982), equation 5.1 the 2D array response A is

$$A(f,p) = \frac{\sin(N\pi f \Delta x p)}{N \sin(\pi f \Delta x p)} \tag{A.9}$$

where N is the number of geophones, f is the frequency, Δx is the geophone spacing and p is the horizontal slowness. Using the ray parameter generated from the ray tracing a frequency-dependent inverse operator can be design to apply to the seismic data as a processing step. This may be used to model both source and receiver arrays.

Appendix B

Transform matrices between reflectivity attributes

This Appendix summarizes the linearized approximations connecting various reflectivity attributes

B.1 Relationship between Shuey and velocity reflectivity

The transform matrix linking the "Shuey" parameters and the Aki and Richards parameters follow from the definitions (2.5), (2.10) and (2.9) resulting in

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -8\bar{\gamma}^2 & -4\bar{\gamma}^2 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{B.1}$$

The inverse transform of this is

$$\begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -\frac{1}{2} & -\frac{1}{8\bar{\gamma}^2} & \frac{1}{8}\frac{4\bar{\gamma}^2+1}{\bar{\gamma}^2} \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix}. \tag{B.2}$$

B.2 Relationship between Impedance and velocity reflectivity

P-wave impedance is a function of P-wave velocity and density

$$\bar{I}_p = \bar{\alpha}\bar{\rho}. \tag{B.3}$$

If the total differential

$$d\bar{I}_p = \frac{d\bar{I}_p}{\partial\bar{\alpha}} d\bar{\alpha} + \frac{\partial\bar{I}_p}{\partial\bar{\rho}} d\bar{\rho}, \tag{B.4}$$

is taken

$$d\bar{I}_p = \bar{\rho} d\bar{\alpha} + \bar{\alpha} d\bar{\rho}, \tag{B.5}$$

and then divided by the average impedance the relationship between velocity and impedance reflectivity attributes is established as

$$\frac{\Delta I_p}{\bar{I}_p} = \frac{\Delta\alpha}{\bar{\alpha}} + \frac{\Delta\rho}{\bar{\rho}}. \tag{B.6}$$

Defining the P-wave velocity reflectivity as $R_a = \frac{1}{2}\frac{\Delta\alpha}{\bar{\alpha}}$ and the density reflectivity as $R_d = \frac{1}{2}\frac{\Delta\rho}{\bar{\rho}}$ then the P-wave impedance reflectivity $R_p = \frac{1}{2}\frac{\Delta I_p}{\bar{I}_p}$ is

$$R_p = R_\alpha + R_d. \tag{B.7}$$

Likewise the relationship between S-wave impedance reflectivity $R_s = \frac{1}{2}\frac{\Delta I_s}{\bar{I}_s}$ and S-wave velocity reflectivity $R_\beta = \frac{1}{2}\frac{\Delta\beta}{\bar{\beta}}$ is

$$R_s = R_\beta + R_d. \tag{B.8}$$

The forward transform linking the velocity reflectivity to impedance reflectivity is

$$\begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}, \tag{B.9}$$

while the inverse transform is

$$\begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix}. \tag{B.10}$$

B.3 Relationship between β/α ratio and velocity reflectivity

In a similar fashion to the P-wave impedance, taking the differential of $\bar{\gamma} = \frac{\bar{\beta}}{\bar{\alpha}}$ and then dividing by $\bar{\gamma}$ results in the relationship

$$\frac{d\bar{\gamma}}{\bar{\gamma}} = \frac{d\bar{\beta}}{\bar{\beta}} - \frac{d\bar{\alpha}}{\bar{\alpha}}, \tag{B.11}$$

Defining the γ reflectivity attribute in a similar fashion to the impedance reflectivity $R_\gamma = \frac{1}{2}\frac{\Delta\gamma}{\bar{\gamma}}$ then

$$R_\gamma = R_\beta - R_\alpha. \tag{B.12}$$

B.4 Relationship between delta-Poisson ratio and velocity reflectivity

The Pseudo-Poisson ratio reflectivity is defined by Shuey (1985) as

$$R_\sigma = \frac{\Delta\nu}{(1-\bar{\nu})^2}, \tag{B.13}$$

and is used by Verm and Hilterman (1995) as a lithology indicator. I call this delta-Poisson ratio reflectivity to distinguish this from the Poisson ratio reflectivity (Section B.5) since the denominator is not the Poisson ratio as the other reflectivity attributes discussed so far. It may be derived from the relationship between the γ ratio and the Poisson ratio

$$\bar{q} = \bar{\gamma}^2 = \frac{\frac{1}{2} - \bar{\nu}}{1-\bar{\nu}}. \tag{B.14}$$

Take the differential of $\bar{q}$ with respect to $\bar{v}$

$$d\bar{q} = \frac{\partial \bar{q}}{\partial \bar{\nu}} d\bar{\nu}. \tag{B.15}$$

On the left hand side write $d\bar{q}$ in terms of $\bar{\gamma}$ and on the right hand side expand $\bar{q}$ in terms of $\bar{\nu}$

$$2\bar{\gamma}d\bar{\gamma} = \frac{\partial}{\partial\bar{\nu}}\left[\left(\frac{1}{2} - \bar{\nu}\right)(1-\bar{\nu})^{-1}\right] d\bar{\nu}. \tag{B.16}$$

Simplifying

$$2\bar{\gamma}d\bar{\gamma} = \left[\frac{\left(-\frac{1}{2}\right)}{(1-\bar{\nu})^2}\right] d\bar{\nu}, \tag{B.17}$$

and rearranging, results in

$$-4\bar{\gamma}^2\frac{d\bar{\gamma}}{\bar{\gamma}} = \frac{d\bar{\nu}}{(1-\bar{\nu})^2}. \tag{B.18}$$

This combined with equation (B.11) results in a relationship between the pseudo-Poisson reflectivity and the velocity reflectivity

$$\frac{\Delta\nu}{(1-\bar{\nu})^2} = 4\bar{\gamma}^2\left[\frac{d\bar{\alpha}}{\bar{\alpha}} - \frac{d\bar{\beta}}{\bar{\beta}}\right], \tag{B.19}$$

or

$$R_\sigma = 8\bar{\gamma}^2\left[R_\alpha - R_\beta\right]. \tag{B.20}$$

B.5 Relationship between Poisson ratio and velocity reflectivity

The Poisson reflectivity attribute may be derived for a definition more consistent with the other definitions in this appendix

$$R_\nu = \frac{1}{2}\frac{\Delta\nu}{\bar{\nu}}. \tag{B.21}$$

To do this, note that equation (B.14) may be rearranged as

$$\bar{\nu} = \frac{\frac{1}{2}-\bar{\gamma}^2}{1-\bar{\gamma}^2}. \tag{B.22}$$

Then if equation (B.17) is divided by $\bar{\nu}$

$$\frac{\bar{\gamma}d\bar{\gamma}}{\bar{\nu}} = -\frac{1}{4}\left[\frac{1}{(1-\bar{\nu})^2}\right]\frac{d\bar{\nu}}{\bar{\nu}}, \tag{B.23}$$

and the definition (equation B.22) is substituted into the result

$$-4\frac{\bar{\gamma}d\bar{\gamma}}{\frac{\frac{1}{2}-\bar{\gamma}^2}{(1-\bar{\gamma}^2)}} = \left[\frac{1}{\left(1-\frac{\frac{1}{2}-\bar{\gamma}^2}{(1-\bar{\gamma}^2)}\right)^2}\right]\frac{d\bar{\nu}}{\bar{\nu}}. \tag{B.24}$$

Upon simplifying

$$\frac{d\bar{\nu}}{\bar{\nu}} = \frac{1}{\frac{3}{2}-\bar{\gamma}^2-\frac{1}{2\bar{\gamma}^2}}\left[\frac{d\alpha}{\alpha} - \frac{d\beta}{\beta}\right], \tag{B.25}$$

this results in the Poisson reflectivity attribute

$$R_\nu = \frac{1}{\frac{3}{2} - \bar{\gamma}^2 - \frac{1}{2\bar{\gamma}^2}} \left[R_\alpha - R_\beta \right]. \tag{B.26}$$

B.6 Relationship between Lamé reflectivity and $\lambda\rho$, λ, $\mu\rho$, μ reflectivity

Both $\lambda\rho$ and $\mu\rho$ can be calculated as linearized reflectivity attributes. Defining the variable $\bar{M}$ such that

$$\bar{M} = \bar{\mu}\bar{\rho} = \bar{I}_s^2, \tag{B.27}$$

then the derivative of this with respect to $\bar{I}_s$ is

$$\frac{d\bar{M}}{d\bar{I}_s} = \frac{d}{d\bar{I}_s}\bar{I}_s^2 = 2\bar{I}_s = 2\frac{\bar{I}_s^2}{\bar{I}_s} = 2\frac{\bar{M}}{\bar{I}_s}, \tag{B.28}$$

or

$$\frac{d\bar{M}}{\bar{M}} = 2\frac{d\bar{I}_s}{\bar{I}_s}. \tag{B.29}$$

Following the convention of reflectivity attributes, the $\mu\rho$ reflectivity is

$$R_{\mu\rho} = \frac{1}{2}\frac{\Delta M}{\bar{M}}, \tag{B.30}$$

so that

$$R_{\mu\rho} = 2R_s. \tag{B.31}$$

Defining $L = \lambda\rho$, then

$$\bar{L} = \bar{\rho}\bar{\lambda} = \bar{I}_p^2 - 2\bar{I}_s^2, \tag{B.32}$$

and the total differential of this is

$$d\bar{L} = \frac{d\bar{I}_p^2}{d\bar{I}_p}d\bar{I}_p - 2\frac{d\bar{I}_s^2}{d\bar{I}_s}d\bar{I}_s = 2\bar{I}_p d\bar{I}_p - 4\bar{I}_s d\bar{I}_s = 2\bar{I}_p^2\frac{d\bar{I}_p}{\bar{I}_p} - 4\bar{I}_s^2\frac{d\bar{I}_s}{\bar{I}_s}, \tag{B.33}$$

or after dividing equation (B.33) by L

$$\frac{d\bar{L}}{\bar{L}} = \frac{2\bar{I}_p^2}{\bar{I}_p^2 - 2\bar{I}_s^2}\frac{d\bar{I}_p}{\bar{I}_p} - \frac{4\bar{I}_s^2}{\bar{I}_p^2 - 2\bar{I}_s^2}\frac{d\bar{I}_s}{\bar{I}_s} = \frac{1}{\frac{\bar{I}_p^2 - 2\bar{I}_s^2}{2\bar{I}_p^2}}\frac{d\bar{I}_p}{\bar{I}_p} - \frac{1}{\frac{\bar{I}_p^2 - 2\bar{I}_s^2}{4\bar{I}_s^2}}\frac{d\bar{I}_s}{\bar{I}_s} = \frac{1}{\frac{1}{2} - \bar{\gamma}^2}\frac{d\bar{I}_p}{\bar{I}_p} - \frac{1}{\frac{1}{4\bar{\gamma}^2} - \frac{1}{2}}\frac{d\bar{I}_s}{\bar{I}_s}. \tag{B.34}$$

Following the convention of reflectivity attributes, the $\lambda\rho$ reflectivity is

$$R_{\lambda\rho} = \frac{1}{2}\frac{\Delta L}{\bar{L}},$$

or

$$R_{\lambda\rho} = \frac{1}{\frac{1}{2} - \bar{\gamma}^2}R_p - \frac{1}{\frac{1}{4\bar{\gamma}^2} - \frac{1}{2}}R_s. \tag{B.35}$$

Writing this in matrix notation

$$\begin{bmatrix} R_{\lambda\rho} \\ R_{\mu\rho} \\ R_d \end{bmatrix} = \begin{bmatrix} -\frac{2}{-1+2\bar{\gamma}^2} & \frac{4\bar{\gamma}^2}{-1+2\bar{\gamma}^2} & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix}, \tag{B.36}$$

and transforming to velocity reflectivity using equation (B.9) results in

$$\begin{bmatrix} R_{\lambda\rho} \\ R_{\mu\rho} \\ R_d \end{bmatrix} = \begin{bmatrix} -\frac{2}{-1+2\bar{\gamma}^2} & \frac{4\bar{\gamma}^2}{-1+2\bar{\gamma}^2} & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{B.37}$$

The transformation from $\lambda\rho$ and $\mu\rho$ reflectivity to λ and μ reflectivity is

$$\begin{bmatrix} R_{\lambda\rho} \\ R_{\mu\rho} \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix}. \tag{B.38}$$

Transforming equation (B.37) to that of λ and μ reflectivity

$$\begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -\frac{2}{-1+2\bar{\gamma}^2} & \frac{4\bar{\gamma}^2}{-1+2\bar{\gamma}^2} & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}, \tag{B.39}$$

results in

$$\begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} \frac{2}{1-2\bar{\gamma}^2} & -\frac{4\bar{\gamma}^2}{1-2\bar{\gamma}^2} & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{B.40}$$

The inverse of this transform matrix is

$$\begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - \bar{\gamma}^2 & \bar{\gamma}^2 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix}. \tag{B.41}$$

Substituting equation (B.41) into the Aki and Richards equation (1.6)

$$\begin{bmatrix} R(\bar{\theta}_m) \end{bmatrix} = \begin{bmatrix} \sec^2\bar{\theta}_m & -8\bar{\gamma}^2\sin^2\bar{\theta}_m & \left(1-4\bar{\gamma}^2\sin^2\bar{\theta}_m\right) \end{bmatrix} \begin{bmatrix} \frac{1}{2} - \bar{\gamma}^2 & \bar{\gamma}^2 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix}, \tag{B.42}$$

results in

$$\begin{bmatrix} R(\bar{\theta}_m) \end{bmatrix} = \begin{bmatrix} \left(\frac{1}{2} - \bar{\gamma}^2\right)\sec^2\bar{\theta}_m & \bar{\gamma}^2\left(\sec^2\bar{\theta}_m - 4\sin^2\bar{\theta}_m\right) & 1 - \frac{1}{2}\sec^2\bar{\theta}_m \end{bmatrix} \begin{bmatrix} R_\lambda \\ R_\mu \\ R_d \end{bmatrix}. \tag{B.43}$$

which is equivalent to Gray et al. (1999) equation (2).

B.7 Relationship between bulk modulus and velocity reflectivity

The bulk modulus expressed in terms of velocity and density is

$$\bar{K} = \bar{\alpha}^2\bar{\rho} - \frac{4}{3}\bar{\beta}^2\bar{\rho}. \tag{B.44}$$

Substituting the differentials of

$$d\left(\bar{\alpha}^2\bar{\rho}\right) = \frac{d\left(\bar{\alpha}^2\bar{\rho}\right)}{d\bar{\alpha}}d\bar{\alpha} + \frac{d\left(\bar{\alpha}^2\bar{\rho}\right)}{d\bar{\rho}}d\bar{\rho} = 2\bar{\alpha}\bar{\rho}d\bar{\alpha} + \bar{\alpha}^2 d\bar{\rho}, \tag{B.45}$$

and

$$d\left(\bar{\beta}^2\bar{\rho}\right) = \frac{d\left(\bar{\beta}^2\bar{\rho}\right)}{d\bar{\beta}}d\bar{\beta} + \frac{d\left(\bar{\beta}^2\bar{\rho}\right)}{d\bar{\rho}}d\bar{\rho} = 2\bar{\beta}\bar{\rho}d\bar{\beta} + \bar{\beta}^2 d\bar{\rho}, \tag{B.46}$$

into the differential of equation (B.44) results in

$$d\bar{K} = 2\bar{\alpha}\bar{\rho}d\bar{\alpha} + \bar{\alpha}^2 d\bar{\rho} - \frac{8}{3}\bar{\beta}\bar{\rho}d\bar{\beta} - \frac{4}{3}\bar{\beta}^2 d\bar{\rho}. \tag{B.47}$$

Rearranging and dividing by the bulk modulus as defined in equation (B.44) results in

$$\frac{d\bar{K}}{\bar{K}} = \frac{2\bar{\alpha}^2\bar{\rho}}{\bar{\alpha}^2\bar{\rho} - \frac{4}{3}\bar{\beta}^2\bar{\rho}}\frac{d\bar{\alpha}}{\bar{\alpha}} - \frac{8}{3}\frac{\bar{\beta}^2\bar{\rho}}{\bar{\alpha}^2\bar{\rho} - \frac{4}{3}\bar{\beta}^2\bar{\rho}}\frac{d\bar{\beta}}{\bar{\beta}} + \frac{\left(\bar{\alpha}^2\bar{\rho} - \frac{4}{3}\bar{\beta}^2\bar{\rho}\right)}{\bar{\alpha}^2\bar{\rho} - \frac{4}{3}\bar{\beta}^2\bar{\rho}}\frac{d\bar{\rho}}{\bar{\rho}}. \tag{B.48}$$

Writing this in terms of $\bar{\gamma}$ and defining the Bulk modulus reflectivity attribute as

$$R_K = \frac{1}{2}\frac{\Delta K}{\bar{K}},$$

results in

$$R_K = \frac{6}{3 - 4\bar{\gamma}^2}R_\alpha - \frac{8\bar{\gamma}^2}{3 - 4\bar{\gamma}^2}R_\beta + R_d. \tag{B.49}$$

The relation between Bulk modulus, μ reflectivity and velocity reflectivity is

$$\begin{bmatrix} R_K \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} \frac{6}{3-4\bar{\gamma}^2.} & -\frac{8\bar{\gamma}^2}{3-4\bar{\gamma}^2} & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{B.50}$$

The inverse of this transform matrix is

$$\begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix} = \begin{bmatrix} \frac{1}{2} - \frac{2}{3}\bar{\gamma}^2 & \frac{2}{3}\bar{\gamma}^2 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_K \\ R_\mu \\ R_d \end{bmatrix}. \tag{B.51}$$

Substituting equation (B.51) into the Aki and Richards equation (1.6)

$$\begin{bmatrix} R(\bar{\theta}_m) \end{bmatrix} = \begin{bmatrix} \sec^2\bar{\theta}_m & -8\bar{\gamma}^2\sin^2\bar{\theta}_m & \left(1 - 4\bar{\gamma}^2\sin^2\bar{\theta}_m\right) \end{bmatrix} \begin{bmatrix} \frac{1}{2} - \frac{2}{3}\bar{\gamma}^2 & \frac{2}{3}\bar{\gamma}^2 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_K \\ R_\mu \\ R_d \end{bmatrix}, \tag{B.52}$$

results in

$$\left[\ R(\bar{\theta}_m)\ \right] = \left[\ \left(\tfrac{1}{2} - \tfrac{2}{3}\bar{\gamma}^2\right)\sec^2\bar{\theta}_m \quad \bar{\gamma}^2\left(\tfrac{2}{3}\sec^2\bar{\theta}_m - 4\sin^2\bar{\theta}_m\right) \quad 1 - \tfrac{1}{2}\sec^2\bar{\theta}_m\ \right] \begin{bmatrix} R_K \\ R_\mu \\ R_d \end{bmatrix}, \tag{B.53}$$

which is equivalent to Gray et al. (1999) equation (1).

B.8 Summary of transforms

Summarizing these relationships results in

$$\begin{bmatrix} A \\ B \\ C \\ R_p \\ R_s \\ R_\gamma \\ R_\sigma \\ R_\nu \\ R_{\lambda\rho} \\ R_{\mu\rho} \\ R_k \\ R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -8\bar{\gamma}^2 & -4\bar{\gamma}^2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \\ 8\bar{\gamma}^2 & -8\bar{\gamma}^2 & 0 \\ \frac{1}{\left(\frac{3}{2}-\bar{\gamma}^2-\frac{1}{2\bar{\gamma}^2}\right)} & \frac{-1}{\left(\frac{3}{2}-\bar{\gamma}^2-\frac{1}{2\bar{\gamma}^2}\right)} & 0 \\ \frac{2}{1-2\bar{\gamma}^2} & 4\frac{\bar{\gamma}^2}{2\bar{\gamma}^2-1} & 2 \\ 0 & 2 & 2 \\ \frac{6}{3-4\bar{\gamma}^2.} & -\frac{8\bar{\gamma}^2}{3-4\bar{\gamma}^2} & 1 \\ \frac{2}{1-2\bar{\gamma}^2} & -\frac{4\bar{\gamma}^2}{1-2\bar{\gamma}^2} & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_\alpha \\ R_\beta \\ R_d \end{bmatrix}. \tag{B.54}$$

Transform matrix (B.54) is written in terms of velocity and density reflectivity. Writing this in terms of impedance and density reflectivity the transform matrix is

$$\begin{bmatrix} A \\ B \\ C \\ R_\alpha \\ R_\beta \\ R_\gamma \\ R_\sigma \\ R_\nu \\ R_{\lambda\rho} \\ R_{\mu\rho} \\ R_k \\ R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -8\bar{\gamma}^2 & 4\bar{\gamma}^2 - 1 \\ 1 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \\ 8\bar{\gamma}^2 & -8\bar{\gamma}^2 & 0 \\ \frac{1}{-\frac{1}{2\bar{\gamma}^2}-\bar{\gamma}^2+\frac{3}{2}} & -\frac{1}{-\frac{1}{2\bar{\gamma}^2}-\bar{\gamma}^2+\frac{3}{2}} & 0 \\ \frac{2}{-2\bar{\gamma}^2+1} & 4\frac{\bar{\gamma}^2}{2\bar{\gamma}^2-1} & -\frac{2}{-2\bar{\gamma}^2+1} - \frac{4\bar{\gamma}^2}{2\bar{\gamma}^2-1} + 2 \\ 0 & 2 & 0 \\ \frac{6}{-4\bar{\gamma}^2+3} & -8\frac{\bar{\gamma}^2}{-4\bar{\gamma}^2+3} & -\frac{6}{-4\bar{\gamma}^2+3} + \frac{8\bar{\gamma}^2}{-4\bar{\gamma}^2+3} + 1 \\ \frac{2}{-2\bar{\gamma}^2+1} & -4\frac{\bar{\gamma}^2}{-2\bar{\gamma}^2+1} & -\frac{2}{-2\bar{\gamma}^2+1} + \frac{4\bar{\gamma}^2}{-2\bar{\gamma}^2+1} + 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix}. \quad \text{(B.55)}$$

B.9 Two-term relationships

Often, for stability reasons, only two parameters are inverted for, such as the P-wave and S-wave impedance reflectivity as is done in the two-term Gidlow equation. If one assumes the Gardner relationship (equation 2.16) holds then

$$\begin{bmatrix} R_p \\ R_s \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{g}{1+g} & 0 \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}, \quad \text{(B.56)}$$

and upon substitution equation (B.55) becomes

$$\begin{bmatrix} A \\ B \\ C \\ R_\alpha \\ R_\beta \\ R_\gamma \\ R_\sigma \\ R_\nu \\ R_{\lambda\rho} \\ R_{\mu\rho} \\ R_k \\ R_\lambda \\ R_\mu \\ R_d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{g}{1+g}\left(4\bar{\gamma}^2-1\right)+1 & -8\bar{\gamma}^2 \\ 1 & 0 \\ -\frac{g}{1+g}+1 & 0 \\ -\frac{g}{1+g} & 1 \\ -1 & 1 \\ 8\bar{\gamma}^2 & -8\bar{\gamma}^2 \\ \frac{1}{-\frac{1}{2\bar{\gamma}^2}-\bar{\gamma}^2+\frac{3}{2}} & \frac{2\bar{\gamma}^2}{-3\bar{\gamma}^2+2\bar{\gamma}^4+1} \\ -\frac{2}{2\bar{\gamma}^2-1} & \frac{4\bar{\gamma}^2}{2\bar{\gamma}^2-1} \\ 0 & 2 \\ -\frac{g}{1+g}-\frac{6}{4\bar{\gamma}^2-3} & \frac{8\bar{\gamma}^2}{4\bar{\gamma}^2-3} \\ -\frac{g}{1+g}-\frac{2}{2\bar{\gamma}^2-1} & \frac{4\bar{\gamma}^2}{2\bar{\gamma}^2-1} \\ -\frac{g}{1+g} & 2 \\ \frac{g}{1+g} & 0 \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}. \tag{B.57}$$

Note that for the case that $\bar{\gamma} = \frac{1}{2}$

$$B = R_p - 2R_s, \tag{B.58}$$

which is equivalent to equation (2.11) in the text. The more general relationship for arbitrary $\bar{\gamma}$ and g is

$$B = \left(\frac{g}{1+g}\left(4\bar{\gamma}^2-1\right)+1\right)R_p - 8\bar{\gamma}^2 R_s, \tag{B.59}$$

or written in terms of S-wave impedance reflectivity

$$R_s = \frac{1}{8\bar{\gamma}^2}\left[\left(\frac{g}{1+g}\left(4\bar{\gamma}^2-1\right)+1\right)A - B\right]. \tag{B.60}$$

The transform matrix (equation B.57) is only approximate since the empirical Gardner relationship is used to reduce the number of parameters from three to two. If another AVO inversion such as the two-term Shuey equation along with the constraint $C = 0$ is used to estimate the parameters the use of different constraints introduces different biases into the parameter estimates. For example, if the two term Shuey

equation is used to estimate the intercept A and gradient B using angles out to $36\,^{\circ}$, one would obtain a biased estimate of the gradient since the third term in the Shuey equation is ignored. Thus, transforming A and B to S-wave impedance reflectivity using equation (B.60) results in a biased estimate different than the estimate from the Gidlow equation. Using a large range of angles, the two-term Gidlow equation gives better estimates of the S-wave impedance reflectivity since it is a better approximation to the Zoeppritz equations at these angles than the two-term Shuey equation.

Appendix C

Relationship between Parameter and Data covariance matrices

C.1 Definition of a covariance matrix

Let there be N random variables x_n whose joint distribution is a multivariate Gaussian distribution. A particular realization is written succinctly in vector form denoted as $\mathbf{x}$. The mean of each variable is written as the vector $\mathbf{x}_o$. The probability of any particular realization of $\mathbf{x}$ is

$$\Pr(\mathbf{x}) \propto \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{x}_o)^T \mathbf{C}_x^{-1}(\mathbf{x}-\mathbf{x}_o)\right], \tag{C.1}$$

where the covariance matrix

$$\mathbf{C}_x = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{12} & \sigma_{22}^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1N} & \sigma_{2N} & \cdots & \sigma_{NN}^2 \end{bmatrix}, \tag{C.2}$$

describes the distribution. The covariance matrix is symmetric and positive definite implying real eigenvalues. The diagonal elements of the covariance matrix represent the variance of each random variable. The variance is a measure of the dispersion or spread of the distribution. The off-diagonal terms are the covariances which measure the degree of correlation between variables.

C.2 Estimating the covariance matrix from sample statistics

The elements of the covariance matrix (equation C.2) represent idealized quantities. Sample statistics can be used to estimate the covariance matrix. Let u_{ik} represent the k^{th} sample of the i^{th} variable. The notation u is used instead of x to indicate that the variable is now a sampled quantity rather than an idealized variable. Assuming that there are K samples of each random variable, then these realizations can be summarized in matrix form

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1K} \\ u_{21} & u_{22} & \cdots & u_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \cdots & u_{NK} \end{bmatrix}, \tag{C.3}$$

where each row represents a different variable and each column a different sample. The matrix $\mathbf{U}$ can also be written as a series of column vectors $\mathbf{u}_k$

$$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_K], \tag{C.4}$$

where each vector represents a different realization. Note bold face variables are vectors while italicized variables are elements. The elements of each of the column vectors $\mathbf{u}_k$ represent the different variables for that sample. The sample mean is the column vector

$$\langle \mathbf{u} \rangle = [\langle \mathbf{u}_1 \rangle, \langle \mathbf{u}_2 \rangle, ..., \langle \mathbf{u}_K \rangle]^T. \tag{C.5}$$

The deviation from the mean for each sample vector is

$$\delta \mathbf{u}_k = \mathbf{u}_k - \langle \mathbf{u} \rangle, \tag{C.6}$$

or in matrix form

$$\delta \mathbf{U} = [\delta \mathbf{u}_1, \delta \mathbf{u}_2, ..., \delta \mathbf{u}_K]. \tag{C.7}$$

The sample covariance matrix is defined as the outer product of $\delta \mathbf{U}$ with itself

$$\hat{\mathbf{C}}_u = \frac{1}{K} \delta \mathbf{U} \delta \mathbf{U}^T = \frac{1}{K} [\mathbf{U} - \langle \mathbf{U} \rangle] [\mathbf{U} - \langle \mathbf{U} \rangle]^T, \tag{C.8}$$

which in matrix form is

$$\hat{\mathbf{C}}_u = \frac{1}{(K-1)} \begin{bmatrix} \delta u_{11} & \delta u_{12} & \cdots & \delta u_{1K} \\ \delta u_{21} & \delta u_{22} & \cdots & \delta u_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \delta u_{N1} & \delta u_{N2} & \cdots & \delta u_{NK} \end{bmatrix} \begin{bmatrix} \delta u_{11} & \delta u_{21} & \vdots & \delta u_{1N} \\ \delta u_{12} & \delta u_{22} & \vdots & \delta u_{2N} \\ \cdots & \cdots & \ddots & \cdots \\ \delta u_{1K} & \delta u_{2K} & \vdots & \delta u_{NK} \end{bmatrix}. \quad \text{(C.9)}$$

Note the symbol $\hat{}$ is used to denote an estimate, so $\hat{\mathbf{C}}_u$ is an estimate of $\mathbf{C}_u$.

C.3 Estimating the uncertainty in linear inverse problems

If the noise is Gaussian, the parameter uncertainty of a linear inverse problem is described by the covariance matrix (C.2). It is estimated by calculating the misfit covariance matrix using sample statistics. The parameter covariance matrix is then estimated from the misfit covariance matrix using the linear relationship between the data and the model space.

The linear equations are written as

$$\mathbf{G}\mathbf{x} = \mathbf{d}, \quad \text{(C.10)}$$

where $\mathbf{G}$ is the linear operator, $\mathbf{x}$ the unknown model parameter vector and $\mathbf{d}$, the data vector. This can be solved in a least squares fashion

$$\mathbf{x} = \mathbf{F}\mathbf{d}, \quad \text{(C.11)}$$

where for an overdetermined set of equations the least squares operator $\mathbf{F}$ is

$$\mathbf{F} = [\mathbf{G}^T\mathbf{G}]^{-1}\mathbf{G}^T. \quad \text{(C.12)}$$

The exact solution for $\mathbf{x}$ is obtained in the case of noise free data $\mathbf{d}$. However, only an estimate of the model parameters $\hat{\mathbf{x}}$ may be obtained when data have noise $\mathbf{n}$. If the data have noise, then the observed data $\tilde{\mathbf{d}}$ is the sum of the actual data and the noise $\mathbf{n}$ so

$$\tilde{\mathbf{d}} = \mathbf{d} + \mathbf{n}, \quad \text{(C.13)}$$

Thus, only an estimate of the model parameters $\hat{\mathbf{x}}$ may be obtained since

$$\hat{\mathbf{x}} = \mathbf{F}\tilde{\mathbf{d}}. \tag{C.14}$$

C.4 Data covariance matrix

Assuming that for each data variable d_m there are K independent observations of it, each with noise, then following equation (C.13), the k^{th} sample of the m^{th} observed data variable is

$$\tilde{d}_{mk} = d_m + n_{mk}, \qquad m = 1, 2, ..., M, \quad k = 1, 2, ..., K. \tag{C.15}$$

Having assumed the noise is Gaussian, the best estimate of each data variable is its mean

$$\bar{d}_m = \left\langle \tilde{d}_{mk} \right\rangle. \tag{C.16}$$

Substituting this estimate (equation C.16) for the ideal in equation (C.15) each samples noise is estimated as

$$\hat{\varepsilon}_{mk} = \tilde{d}_{mk} - \bar{d}_m, \tag{C.17}$$

which is the deviation around the data variable similar to equation (C.7) which written as a matrix is

$$\hat{\mathbf{E}} = \begin{bmatrix} \tilde{d}_{11} - \bar{d}_1 & \tilde{d}_{12} - \bar{d}_1 & \cdots & \tilde{d}_{1K} - \bar{d}_1 \\ \tilde{d}_{21} - \bar{d}_2 & \tilde{d}_{22} - \bar{d}_2 & \cdots & \tilde{d}_{2K} - \bar{d}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{d}_{M1} - \bar{d}_M & \tilde{d}_{M2} - \bar{d}_M & \cdots & \tilde{d}_{MK} - \bar{d}_M \end{bmatrix}, \tag{C.18}$$

or

$$\mathbf{E} = \left[\tilde{\mathbf{d}}_1 - \bar{\mathbf{d}}, \tilde{\mathbf{d}}_2 - \bar{\mathbf{d}}, ..., \tilde{\mathbf{d}}_K - \bar{\mathbf{d}} \right] = \left[\hat{\boldsymbol{\varepsilon}}_1, \hat{\boldsymbol{\varepsilon}}_2, ..., \hat{\boldsymbol{\varepsilon}}_K \right]. \tag{C.19}$$

From the definition of the sample covariance matrix (equation C.8), the sample data covariance matrix is

$$\hat{\mathbf{C}}_d = \frac{1}{K} \hat{\mathbf{E}} \hat{\mathbf{E}}^T, \tag{C.20}$$

which is an estimate of the misfit covariance matrix

$$\mathbf{C}_d = \begin{bmatrix} \sigma^2_{d_{11}} & \sigma^2_{d_{12}} & \cdots & \sigma^2_{d_{1M}} \\ \sigma^2_{d_{12}} & \sigma^2_{d_{22}} & \cdots & \sigma^2_{d_{2M}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2_{d_{1M}} & \sigma^2_{d_{2M}} & \cdots & \sigma^2_{d_{MM}} \end{bmatrix}. \tag{C.21}$$

As the number of samples get larger ($K \to \infty$) the sample covariance matrix (C.20) becomes closer to covariance matrix (C.21).

C.5 Relationship between the misfit covariance matrix and the parameter covariance matrix

From the linear relationship (equation C.14) and the definition of data error (equation C.13)

$$\hat{\mathbf{x}} = \mathbf{F}\mathbf{d} + \mathbf{F}\mathbf{n} = \mathbf{x} + \mathbf{F}\mathbf{n}, \tag{C.22}$$

the parameter uncertainty $\delta\mathbf{x} = \mathbf{x} - \hat{\mathbf{x}}$ may be estimated using

$$\delta\mathbf{x} = \mathbf{F}\mathbf{n}. \tag{C.23}$$

If the noise $\mathbf{n}$ is known then it would be possible to estimate the parameter uncertainty. Since the actual noise is unknown, its estimate $\boldsymbol{\varepsilon}$ is used (equation C.17), thus

$$\delta\hat{\mathbf{x}} = \mathbf{F}\hat{\boldsymbol{\varepsilon}}. \tag{C.24}$$

For K sample error vectors

$$\left[\delta\hat{\mathbf{x}}_1, \delta\hat{\mathbf{x}}_2, \ldots, \delta\hat{\mathbf{x}}_k\right] = \mathbf{F}\left[\hat{\boldsymbol{\varepsilon}}_1, \hat{\boldsymbol{\varepsilon}}_2, \ldots, \hat{\boldsymbol{\varepsilon}}_k\right], \tag{C.25}$$

or

$$\delta\hat{\mathbf{X}} = \mathbf{F}\hat{\mathbf{E}}, \tag{C.26}$$

where

$$\delta\hat{\mathbf{X}} = \begin{bmatrix} \delta\hat{x}_{11} & \delta\hat{x}_{12} & \cdots & \delta\hat{x}_{1K} \\ \delta\hat{x}_{21} & \delta\hat{x}_{22} & \cdots & \delta\hat{x}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \delta\hat{x}_{N1} & \delta\hat{x}_{N2} & \cdots & \delta\hat{x}_{NK} \end{bmatrix}. \tag{C.27}$$

Again, from the definition of the sample covariance matrix, the sample parameter covariance matrix is

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \frac{1}{K}\left(\delta\hat{\mathbf{X}}\right)\left(\delta\hat{\mathbf{X}}\right)^T, \tag{C.28}$$

which is equivalent to

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \frac{1}{K}\left(\mathbf{F}\hat{\mathbf{E}}\right)\left(\mathbf{F}\hat{\mathbf{E}}\right)^T. \tag{C.29}$$

Upon the substitution of equation (C.20) $K\hat{\mathbf{C}}_d = \hat{\mathbf{E}}\hat{\mathbf{E}}^T$ the sample parameter covariance matrix becomes

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \mathbf{F}\hat{\mathbf{C}}_d\mathbf{F}^T, \tag{C.30}$$

where $\hat{\mathbf{C}}_d$ may be calculated using equation (C.20).

C.6 Simplifying assumptions: uniform noise

Often there are not multiple samples for each data variable. In fact, there is often only one sample. In such cases the data sample covariance matrix is

$$\hat{\mathbf{C}}_d = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1M} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1M} \end{bmatrix}, \tag{C.31}$$

which is rank 1. All the rows or columns of the sample covariance matrix are scalar multiples of each other. Thus, there is only enough information to estimate one parameter in the misfit covariance matrix

$$\hat{\mathbf{C}}_d = \begin{bmatrix} \hat{\sigma}^2_{d_{11}} & \hat{\sigma}^2_{d_{12}} & \cdots & \hat{\sigma}^2_{d_{1M}} \\ \hat{\sigma}^2_{d_{12}} & \hat{\sigma}^2_{d_{22}} & \cdots & \hat{\sigma}^2_{d_{2M}} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}^2_{d_{1M}} & \hat{\sigma}^2_{d_{2M}} & \cdots & \hat{\sigma}^2_{d_{MM}} \end{bmatrix}. \tag{C.32}$$

How then to proceed? Typically it is assumed that each variable's misfit is independent. That means that (1) the covariance cross terms disappear, and (2) that each variable's misfit has the same variance $\hat{\sigma}_N^2$. Under these assumptions the misfit covariance matrix becomes

$$\hat{\mathbf{C}}_d = \begin{bmatrix} \hat{\sigma}_N^2 & 0 & \cdots & 0 \\ 0 & \hat{\sigma}_N^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\sigma}_N^2 \end{bmatrix} = \hat{\sigma}_N^2 \mathbf{I} \tag{C.33}$$

so the parameter covariance matrix (C.30) reduces to

$$\hat{\mathbf{C}}_{\hat{\mathbf{x}}} = \hat{\sigma}_N^2 [\mathbf{G}^T \mathbf{G}]^{-1}. \tag{C.34}$$

Appendix D

AVO least squares solution

D.1 Two-term Shuey equation

In this appendix the least squares inverse of the two-term Shuey equation is developed. For M data points the linear equation is

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix} = \begin{bmatrix} 1 & \sin^2\bar{\theta}_1 \\ 1 & \sin^2\bar{\theta}_2 \\ \vdots & \vdots \\ 1 & \sin^2\bar{\theta}_M \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}, \tag{D.1}$$

which written in matrix form is $\mathbf{d} = \mathbf{G}\mathbf{x}$. The least squares solution of this is $\hat{\mathbf{x}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}$. To calculate this, first $\left[\mathbf{G}^T\mathbf{G}\right]^{-1}$ is calculated. Note

$$\mathbf{G}^T\mathbf{G} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \sin^2\bar{\theta}_1 & \sin^2\bar{\theta}_2 & \cdots & \sin^2\bar{\theta}_M \end{bmatrix} \begin{bmatrix} 1 & \sin^2\bar{\theta}_1 \\ 1 & \sin^2\bar{\theta}_2 \\ \vdots & \vdots \\ 1 & \sin^2\bar{\theta}_M \end{bmatrix}, \tag{D.2}$$

and

$$\mathbf{G}^T\mathbf{G} = \begin{bmatrix} M & \sum_{m=1}^{M}\sin^2\bar{\theta}_m \\ \sum_{m=1}^{M}\sin^2\bar{\theta}_m & \sum_{m=1}^{M}\sin^4\bar{\theta}_m \end{bmatrix}. \tag{D.3}$$

Since the inverse of a generic symmetric 2×2 matrix is

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix}^{-1} = \frac{1}{AC - B^2}\begin{bmatrix} C & -B \\ -B & A \end{bmatrix}, \tag{D.4}$$

then the inverse of equation (D.3) is

$$\left[\mathbf{G}^T\mathbf{G}\right]^{-1} = \frac{1}{D}\begin{bmatrix} \sum_{m=1}^{M}\sin^4\bar{\theta}_m & -\sum_{m=1}^{M}\sin^2\bar{\theta}_m \\ -\sum_{m=1}^{M}\sin^2\bar{\theta}_m & M \end{bmatrix}, \tag{D.5}$$

where

$$D = M\sum_{m=1}^{M}\sin^4\bar{\theta}_m - \left(\sum_{m=1}^{M}\sin^2\bar{\theta}_m\right)^2. \tag{D.6}$$

To complete the solution, note that

$$\mathbf{G}^T\mathbf{d} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \sin^2\bar{\theta}_1 & \sin^2\bar{\theta}_2 & \cdots & \sin^2\bar{\theta}_M \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix}, \tag{D.7}$$

is

$$\mathbf{G}^T\mathbf{d} = \begin{bmatrix} \sum_{j=1}^{M} d_j \\ \sum_{j=1}^{M} d_j\sin^2\bar{\theta}_j \end{bmatrix}. \tag{D.8}$$

Therefore $\mathbf{x} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}$

$$\begin{bmatrix} \hat{A} \\ \hat{B} \end{bmatrix} = \begin{bmatrix} \frac{\left(\sum_{m=1}^{M}\sin^4\bar{\theta}_m\right)\left(\sum_{j=1}^{M}d_j\right)-\left(\sum_{m=1}^{M}\sin^2\bar{\theta}_m\right)\left(\sum_{j=1}^{M}d_j\sin^2\bar{\theta}_j\right)}{D} \\ \frac{M\left(\sum_{j=1}^{M}d_j\sin^2\bar{\theta}_j\right)-\left(\sum_{m=1}^{M}\sin^4\bar{\theta}_m\right)\left(\sum_{j=1}^{M}d_j\right)}{D} \end{bmatrix}. \tag{D.9}$$

The covariance matrix as defined by equation (C.34) for the two term Gidlow equation is

$$\begin{bmatrix} \hat{\sigma}^2_{R_p} & \hat{\sigma}_{R_pR_s} \\ \hat{\sigma}_{R_pR_s} & \hat{\sigma}^2_{R_s} \end{bmatrix} = \frac{\hat{\sigma}^2_N}{D}\begin{bmatrix} \sum_{m=1}^{M}\sin^4\bar{\theta}_m & -\sum_{m=1}^{M}\sin^2\bar{\theta}_m \\ -\sum_{m=1}^{M}\sin^2\bar{\theta}_m & M \end{bmatrix}, \tag{D.10}$$

where D is defined by equation (D.6).

D.2 Two-term Gidlow equation

In this appendix the least squares inverse of the two-term Gidlow et al. (1992) equation is developed. For M data points equation (2.20) is

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix} = \begin{bmatrix} \sec^2\bar{\theta}_1 & -8\bar{\gamma}^2\sin^2\bar{\theta}_1 \\ \sec^2\bar{\theta}_2 & -8\bar{\gamma}^2\sin^2\bar{\theta}_2 \\ \vdots & \vdots \\ \sec^2\bar{\theta}_M & -8\bar{\gamma}^2\sin^2\bar{\theta}_M \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}, \tag{D.11}$$

which written is matrix form is $\mathbf{d} = \mathbf{G}\mathbf{x}$. The least squares solution of this is $\hat{\mathbf{x}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}$. In order to solve this first $\left[\mathbf{G}^T\mathbf{G}\right]^{-1}$ is calculated. Note that

$$\mathbf{G}^T\mathbf{G} = \begin{bmatrix} \sum_{m=1}^{M}\sec^4\bar{\theta}_m & -8\bar{\gamma}^2\sum_{m=1}^{M}\tan^2\bar{\theta}_m \\ -8\bar{\gamma}^2\sum_{m=1}^{M}\tan^2\bar{\theta}_m & 64\bar{\gamma}^4\sum_{m=1}^{M}\sin^4\bar{\theta}_m \end{bmatrix}, \tag{D.12}$$

whose inverse is

$$\left[\mathbf{G}^T\mathbf{G}\right]^{-1} = \frac{1}{D}\begin{bmatrix} \sum_{m=1}^{M}\sin^4\bar{\theta}_m & \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M}\tan^2\bar{\theta}_m \\ \frac{1}{8\bar{\gamma}^2}\sum_{m=1}^{M}\tan^2\bar{\theta}_m & \frac{1}{64\bar{\gamma}^4}\sum_{m=1}^{M}\sec^4\bar{\theta}_m \end{bmatrix}, \tag{D.13}$$

where

$$D = \left(\sum_{j=1}^{M}\sec^4\bar{\theta}_j\right)\sum_{m=1}^{M}\sin^4\bar{\theta}_m - \left(\sum_{m=1}^{M}\tan^2\bar{\theta}_m\right)^2. \tag{D.14}$$

Next

$$\mathbf{G}^T\mathbf{d} = \begin{bmatrix} \sec^2\bar{\theta}_1 & \sec^2\bar{\theta}_2 & \cdots & \sec^2\bar{\theta}_M \\ -8\bar{\gamma}^2\sin^2\bar{\theta}_1 & -8\bar{\gamma}^2\sin^2\bar{\theta}_2 & \cdots & -8\bar{\gamma}^2\sin^2\bar{\theta}_M \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix}, \tag{D.15}$$

is

$$\mathbf{G}^T\mathbf{d} = \begin{bmatrix} \sum_{j=1}^{M} d_j\sec^2\bar{\theta}_j \\ -8\bar{\gamma}^2\sum_{j=1}^{M} d_j\sin^2\bar{\theta}_j \end{bmatrix}. \tag{D.16}$$

Therefore the least squares solution $\hat{\mathbf{x}} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}$ is

$$\begin{bmatrix} \hat{R}_p \\ \hat{R}_s \end{bmatrix} = \begin{bmatrix} \frac{\left(\sum_{m=1}^{M} \sin^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j}{D} \\ \frac{\left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \sec^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j}{8\bar{\gamma}^2 D} \end{bmatrix}. \tag{D.17}$$

Written separately

$$\hat{R}_p = \frac{\left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \sin^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j}{D}, \tag{D.18}$$

and

$$\hat{R}_s = \frac{\left(\sum_{m=1}^{M} \sec^4 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sin^2 \bar{\theta}_j - \left(\sum_{m=1}^{M} \tan^2 \bar{\theta}_m\right) \sum_{j=1}^{M} d_j \sec^2 \bar{\theta}_j}{8\bar{\gamma}^2 D}. \tag{D.19}$$

The covariance matrix as defined by equation (C.34) for the two term Gidlow equation is

$$\begin{bmatrix} \hat{\sigma}^2_{R_p} & \hat{\sigma}_{R_p R_s} \\ \hat{\sigma}_{R_p R_s} & \hat{\sigma}^2_{R_s} \end{bmatrix} = \frac{\hat{\sigma}^2_N}{D} \begin{bmatrix} \sum_{m=1}^{M} \sin^4 \bar{\theta}_m & \frac{1}{8\bar{\gamma}^2} \sum_{m=1}^{M} \tan^2 \bar{\theta}_m \\ \frac{1}{8\bar{\gamma}^2} \sum_{m=1}^{M} \tan^2 \bar{\theta}_m & \frac{1}{64\bar{\gamma}^4} \sum_{m=1}^{M} \sec^4 \bar{\theta}_m \end{bmatrix}. \tag{D.20}$$

Appendix E

Marginalization of noise

Since there is no explicit interest in the variance σ_N in equation (3.47)

$$P(\mathbf{x}, \sigma_N|\mathbf{d}, I) \propto (2\pi)^{-\frac{M}{2}} \sigma_N^{-M} \exp\left(-\frac{1}{2\sigma_N^2}\sum_{m=1}^{M}\left(\sum_{n=1}^{3} G_{mn}x_n - d_m\right)^2\right) \exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{W}_\mathbf{x}\mathbf{x}\right], \tag{E.1}$$

it is marginalized (Sivia, 1996). Marginalization

$$P(X|I) = \int P(X, Y|I)dY, \tag{E.2}$$

is useful in the removal of a variable. This is done by integrating over the range of possible values for that particular variable so equation (E.1) becomes

$$P(\mathbf{x}|\mathbf{d}, I) \propto \int_0^\infty \sigma_N^{-M} \exp\left(-\frac{1}{2\sigma_N^2}\sum_{m=1}^{M}\left(\sum_{n=1}^{3} G_{mn}x_n - d_m\right)^2 - \frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j\right) d\sigma_N. \tag{E.3}$$

After substituting $\sigma_N = \frac{1}{t}$ and $d\sigma_N = -\frac{dt}{t^2}$ equation (E.3) becomes

$$P(\mathbf{x}|\mathbf{d}, I) \propto \int_0^\infty t^{M-2} \exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\left(\sum_{n=1}^{3} G_{mn}x_n - d_m\right)^2 - \frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j\right) dt. \tag{E.4}$$

Moving the prior outside the integral since it has no influence on t results in

$$P(\mathbf{x}|\mathbf{d},I) \propto \exp\left[-\frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j\right] \int_0^\infty t^{M-2} \exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\left(\sum_{n=1}^{3} G_{mn}x_n - d_m\right)^2\right) dt, \tag{E.5}$$

which when integrated by parts iteratively reduces to

$$P(\mathbf{x}|\mathbf{d},I) \propto \exp\left[-\frac{1}{2}\sum_{i}^{3}\sum_{j}^{3} x_i W_{x_{ij}} x_j\right] (\sum_{m=1}^{M}\left(\sum_{n=1}^{3} G_{mn}x_n - d_m\right)^2)^{-\frac{(M-1)}{2}}. \tag{E.6}$$

This follows by noting that if $\chi_m = \sum_{n=1}^{3} G_{mn}x_n - d_m$ then the integral I_1

$$I_1 = \int_0^\infty t^{M-2} \exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2\right) dt. \tag{E.7}$$

may be solved by integrating by parts

$$\int_0^\infty f(t)\, g'(t)\, dt = f(t)\, g(t)]_0^\infty - \int_0^\infty f'(t)\, g(t)\, dt, \tag{E.8}$$

where $g(t)$ has been defined as

$$g(t) = -\left(\sum_{m=1}^{M}\chi_m^2\right)^{-1} \exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2\right), \tag{E.9}$$

implying

$$g'(t) = t\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2), \tag{E.10}$$

and

$$f(t) = t^{M-3} \Rightarrow f'(t) = (M-3)\, t^{M-4}. \tag{E.11}$$

Substituting these values into the left hand side of of (E.8)

$$\int_0^\infty f(t)\, g'(t)\, dt = \int_0^\infty t^{M-3} t\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2) = \int_0^\infty t^{M-2}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt, \tag{E.12}$$

and the right hand side of (E.8) results in

$$\begin{aligned} r.h.s &= -t^{M-3}\left(\sum_{m=1}^{M}\chi_m^2\right)^{-1}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)\Bigg]_{t=0}^{\infty} \\ &\quad + (M-3)\int_0^{\infty} t^{M-4}\left(\sum_{m=1}^{M}\chi_m^2\right)^{-1}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt. \end{aligned} \tag{E.13}$$

The first term on the right hand side is zero where the sketch of argument is that the exponent goes to zero faster than t^M (Sivia, 1996) so

$$\int_0^{\infty} t^{M-2}\exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2\right)dt = (M-3)\left(\sum_{m=1}^{M}\chi_m^2\right)^{-1}\int_0^{\infty} t^{M-4}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt. \tag{E.14}$$

Repeat this process, this time integrating the right hand side by inspection, results in

$$r.h.s = (M-3)(M-5)\left(\sum_{m=1}^{M}\chi_m^2\right)^{-2}\int_0^{\infty} t^{M-6}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt, \tag{E.15}$$

or

$$r.h.s = (M-3)(M-5)\cdots(M-(M-1))\left(\sum_{m=1}^{M}\chi_m^2\right)^{-\frac{M-1}{2}}\int_0^{\infty}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt, \tag{E.16}$$

where $\int_0^{\infty}\exp(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2)dt = c$ so

$$\int_0^{\infty} t^{M-2}\exp\left(-\frac{t^2}{2}\sum_{m=1}^{M}\chi_m^2\right)dt \propto \left(\sum_{m=1}^{M}\chi_m^2\right)^{-\frac{M-1}{2}}, \tag{E.17}$$

which is the student distribution and so upon substitution of $\chi_m = \sum_{n=1}^{3} G_{mn}x_n - d_m$ and the inclusion of the *a priori* constraints the desired result

$$P(\mathbf{x}|\mathbf{d},I) \propto \exp\left[-\frac{1}{2}\sum_{i}^{3}\sum_{j}^{3}x_i W_{x_{ij}}x_j\right]\left(\sum_{m=1}^{M}\left(\sum_{n=1}^{3}G_{mn}x_n - d_m\right)^2\right)^{-\frac{M-1}{2}}, \tag{E.18}$$

is arrived at.

Appendix F

Two-term covariance matrix parameterized in terms of impedance

Gidlow et al. (1992) use a similar relationship to the mudrock line, but parameterized in terms of impedance

$$I_p \approx m_I I_s + b. \tag{F.1}$$

to develop the fluid stack based on impedance reflectivity. In water saturated sediments equation (F.1) written in terms of reflectivity is

$$R_p \approx m_I \bar{\gamma} R_s, \tag{F.2}$$

where $\gamma = \beta/\alpha$. Note that even if R_p is stationary, R_s is most likely non-stationary since $\bar{\gamma}$ usually changes as a function of depth. The variables are transformed

$$\begin{bmatrix} R_p \\ \tilde{R}_s \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \bar{\gamma} \end{bmatrix} \begin{bmatrix} R_p \\ R_s \end{bmatrix}, \tag{F.3}$$

so as to make the transformed variables stationary. Under this transformation equation (F.2) becomes

$$R_p = m_I \tilde{R}_s. \tag{F.4}$$

where $\tilde{R}_s$ is the scaled S-wave impedance reflectivity. The slope may be calculated as a least squares problem assuming that R_p and R'_s are stationary and that the mudrock slope is constant over the inversion window. In this case, equation (F.4) written in matrix notation is

$$\mathbf{r}_p = m_I \tilde{\mathbf{r}}_s, \tag{F.5}$$

for which the least squares solutions is

$$m_I = \frac{\tilde{\mathbf{r}}_s^T \mathbf{r}_p}{\tilde{\mathbf{r}}_s^T \tilde{\mathbf{r}}_s} = \frac{\sigma_{R_p \tilde{R}_s}}{\sigma^2_{\tilde{R}_s}}. \tag{F.6}$$

Here $\sigma_{R_p R'_s}$ is the covariance between the P-wave and S-wave impedance reflectivity and $\sigma^2_{R'_s}$ is the variance of the scaled S-wave impedance reflectivity. If m_I is not constant over the whole inversion window then analysis may be broken into smaller windows for which m_I is constant and then apply the constraints in a time dependent fashion. Rewriting equation (F.6), written in terms of the correlation coefficient (equation 3.11), results in a relationship between the P- and S- wave velocity reflectivity variance

$$\sigma_{\tilde{R}_s} = \frac{r_{R_p \tilde{R}_s}}{m_I} \sigma_{R_p}. \tag{F.7}$$

After substituting equation (3.11) and (F.7) the covariance matrix for the transformed variables $\mathbf{C}_{\mathbf{x}'}$ written in terms of the mudrock slope m_I and the correlation of coefficient $r_{R_p R'_s}$ is

$$\mathbf{C}_{\tilde{\mathbf{x}}} = \begin{bmatrix} \sigma^2_{R_p} & \sigma_{R_p \tilde{R}_s} \\ \sigma_{R_p \tilde{R}_s} & \sigma^2_{\tilde{R}_s} \end{bmatrix} = \sigma^2_{R_p} \begin{bmatrix} 1 & \frac{r^2_{R_p \tilde{R}_s}}{m_I} \\ \frac{r^2_{R_p \tilde{R}_s}}{m_I} & \frac{r^2_{R_p \tilde{R}_s}}{m_I^2} \end{bmatrix}. \tag{F.8}$$

Writing this in terms of the original variables using equation (Chapter 3, equation 3.27) $\mathbf{C}_x = \mathbf{T}^{-1} \mathbf{C}_{\tilde{\mathbf{x}}} \left(\mathbf{T}^{-1}\right)^T$ results in

$$\mathbf{C}_{\mathbf{x}} = \begin{bmatrix} \sigma^2_{R_p} & \sigma_{R_p R_s} \\ \sigma_{R_p R_s} & \sigma^2_{R_s} \end{bmatrix} = \sigma^2_{R_p} \begin{bmatrix} 1 & \frac{r^2_{R_p \tilde{R}_s}}{m_I \bar{\gamma}} \\ \frac{r^2_{R_p \tilde{R}_s}}{m_I \bar{\gamma}} & \frac{r^2_{R_p \tilde{R}_s}}{m_I^2 \bar{\gamma}^2} \end{bmatrix}. \tag{F.9}$$

Appendix G

Norms based on long-tailed distributions

G.1 Derivative of ℓ^p norm

To facilitate calculating the derivative, equation (6.28) is rewritten as

$$J_p = \sum_i \rho_p(u_i), \tag{G.1}$$

where $\rho_p = \frac{1}{p}|u_i|^p$ and$u_i = x_i/\sigma_{x_i}$. Taking the derivative of J_p results in

$$\frac{\partial J_p}{\partial x_n} = 0 = \sum_i \frac{\partial \rho_p(u_i)}{\partial x_n} = \sum_i \frac{\partial \rho_p(u_i)}{\partial u_i}\frac{\partial u_i}{\partial x_n} = \frac{1}{\sigma_{x_n}}\sum_i \frac{\partial \rho_p(u_i)}{\partial u_i}\delta_{in} = \frac{1}{\sigma_{x_n}}\sum_i \psi_p(u_i)\,\delta_{in}, \tag{G.2}$$

where

$$\psi_p(u_i) = \frac{\partial \rho_p(u_i)}{\partial u_i} = \frac{\partial}{\partial u_i}\frac{1}{p}|u_i|^p = |u_i|^{p-1}\,sgn(u_i) = \frac{u_i}{|u_i|}|u_i|^{p-1} = u_i|u_i|^{p-2} \tag{G.3}$$

so with $u_i = x_i/\sigma_{x_i}$

$$\frac{\partial J_p}{\partial x_n} = \frac{1}{\sigma_{x_n}}\sum_i \frac{x_i}{\sigma_{x_i}}\left|\frac{x_i}{\sigma_{x_i}}\right|^{p-2}\delta_{in} = \frac{1}{\sigma_{x_n}}\frac{x_n}{\sigma_{x_n}}\left|\frac{x_n}{\sigma_{x_n}}\right|^{p-2} \tag{G.4}$$

thus

$$\frac{\partial J_p}{\partial x_n} = \frac{1}{\sigma_{x_n}^2}\left|\frac{x_n}{\sigma_{x_n}}\right|^{p-2} x_n \tag{G.5}$$

G.2 Cauchy Norm

This derivation largely follows Sacchi and Ulrych (1995) making modifications needed to handle variables with two distinct variances. The entropy is given by

$$h = -\int p\left(\mathbf{x}'\right) \log\left[p\left(\mathbf{x}'\right)\right] d\mathbf{x}', \tag{G.6}$$

where

$$p\left(\mathbf{x}'\right) = \exp\left(-\lambda_0 - \lambda_1 S\left(\mathbf{x}'\right)\right), \tag{G.7}$$

and where λ_0, λ_1 are Lagrange multipliers. The parameter λ_0 is a normalization factor such that

$$\int p\left(x\right) dx = 1. \tag{G.8}$$

The function $S\left(\mathbf{x}'\right)$ contains prior information about the model. A function with minimum structure

$$S\left(\mathbf{x}'\right) = \sum_{i=1}^{N} \ln\left(x_i x_i^* + b\right), \tag{G.9}$$

fulfills this objective. The parameter b is a small additive perturbation that represents default power in absence of events. This insures that the gradient of $S\left(\mathbf{x}'\right)$ is continuous. The prior probability according to equation (G.7) and (G.9) is

$$P\left(\mathbf{x}'\right) = \exp\left(-\lambda_0 - \lambda_1 \sum_{i=1}^{N} \ln\left(x_i x_i^* + b\right)\right), \tag{G.10}$$

or using the fact that the random variables in this case are real

$$P\left(\mathbf{x}'\right) = \prod_{i}^{N} \frac{\exp\left(-\lambda_0\right)}{\left(\left|x_i\right|^2 + b\right)^{\lambda_1}}. \tag{G.11}$$

For a single element and $\lambda_1 = 1$ this distribution (G.11) becomes

$$P\left(x_i\right) = \frac{e^{-\lambda_0}}{\left(\left|x\right|^2 + b\right)}, \tag{G.12}$$

which has the same functional form as the Cauchy distribution (Westergren, 1998)

$$f\left(x\right) = \frac{a}{\pi\left(x^2 + a^2\right)}. \tag{G.13}$$

In equation (G.11) λ_1 controls the sparseness in the model with large values leading to sharp distributions. At the other extreme, very small values lead to a uniform distribution. Further, when b is large the distribution tends to a normal distribution.

Equation (G.9) is rearranged so that reflectivity is normalized by the coefficient b

$$S\left(\mathbf{x}'\right) = \sum_{i=1}^{N} \ln\left(x_i^2 + b\right) = \sum_{i=1}^{N} \ln\left(b\left(\frac{x_i^2}{b} + 1\right)\right) = \sum_{i=1}^{N} \ln b + \sum_{i=1}^{N}\left(\frac{x_i^2}{b} + 1\right). \quad \text{(G.14)}$$

The coefficient b acts in a similar fashion to variance in that it controls the width or dispersion of the distribution, though strictly speaking the variance is not defined for the Cauchy distribution. The value of b is different for the two variables. For comparison with other methods the dispersion term b is rewritten as $\sigma = \sqrt{\frac{b}{2}}$ and in a similar fashion to the ℓ^1 derivation the following substitution is performed

$$u_i = \begin{cases} \frac{x_i}{\sigma_1} & i \leq L \\ \frac{x_i}{\sigma_2} & i > L \end{cases}, \quad \text{(G.15)}$$

to get

$$S\left(\mathbf{x}'\right) = C + \sum_{i=1}^{N} \ln\left(\frac{u_i^2}{2} + 1\right), \quad \text{(G.16)}$$

using the fact that the term $C = \sum_{i=1}^{N} \ln\left(b\right)$ is a constant. Analogous to the ℓ^p norm derivation, this is rewritten as

$$S\left(\mathbf{x}'\right) = C + \sum_{i=1}^{n_s} \rho_2\left(u\right) \quad \text{(G.17)}$$

where

$$\rho_2\left(u\right) = \ln\left(\frac{u_i^2}{2} + 1\right). \quad \text{(G.18)}$$

In deriving the MAP solution the derivative of the natural logarithm of the prior probability distribution (G.7) is required.

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\partial}{\partial x_n}\left(\lambda_0 + \lambda_1 S\left(\mathbf{x}'\right)\right). \quad \text{(G.19)}$$

Upon substitution of equation (G.17) this results in

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\partial}{\partial x_n}\left[\lambda_1\left(C + \sum_{i=1}^{N}\rho_2\left(u_i\right)\right)\right] = \lambda_1 \sum_{i=1}^{N}\frac{\partial \rho_2\left(u_i\right)}{\partial x_n}, \tag{G.20}$$

where

$$\frac{\partial \rho_2\left(u_i\right)}{\partial x_n} = \frac{\partial \rho_2\left(u_i\right)}{\partial u_i}\frac{\partial u_i}{\partial x_n} = \frac{\partial \rho_2\left(u_i\right)}{\partial u_i}\frac{\partial}{\partial x_n}\left(\frac{x_i}{\sigma_i}\right) = \frac{\partial \rho_2\left(u_i\right)}{\partial u}\frac{1}{\sigma_i}\delta_{in} = \psi_c\left(u_i\right)\frac{1}{\sigma_i}\delta_{in}, \tag{G.21}$$

and

$$\psi_c\left(u_i\right) = \frac{\partial}{\partial u_i}\left(\ln\left(\frac{u_i^2}{2}+1\right)\right) = \frac{u_i}{\left(\frac{u_i^2}{2}+1\right)}. \tag{G.22}$$

Combining these last three results gives

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \lambda_1 \sum_{i=1}^{N}\frac{u_i}{\left(\frac{u_i^2}{2}+1\right)}\frac{1}{\sigma_i}\delta_{in}. \tag{G.23}$$

The substitution of u, (equation G.15) and equation (6.19) results in

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\lambda_1}{\sigma_{r_p}^2}\begin{cases} \frac{1}{\tilde{\sigma}_1^2}\frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2\tilde{\sigma}_1^2}+1\right)}x_n & n \le L \\ \frac{1}{\tilde{\sigma}_2^2}\frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2\tilde{\sigma}_2^2}+1\right)}x_n & n > L \end{cases} \tag{G.24}$$

The substitution for $\sigma_1^2 = \sigma_{r_p}^2\tilde{\sigma}_1^2$ and $\sigma_2^2 = \sigma_{r_p}^2\tilde{\sigma}_2^2$ is not strictly valid since in the Cauchy derivation σ^2 is not variance. However, σ^2 does describe dispersion of the distribution. This is accomplished through the substitution. Further, the relative spread of the two variables is conveyed. Since the scaling factor λ_1 is present only the relative dispersions between the two variables is required which is accomplished through the substitution.

In matrix form equation (G.24) becomes

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\lambda_1}{\sigma_{r_p}^2}\mathbf{Q}_c\mathbf{x} \tag{G.25}$$

where the diagonal elements of matrix $\mathbf{Q}_c$ are defined by

$$Q_{nn} = \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2 \tilde{\sigma}_1^2}+1\right)} & n \leq L \\ \frac{1}{\tilde{\sigma}_2^2} \frac{1}{\left(\frac{x_i^2}{2\sigma_{r_p}^2 \tilde{\sigma}_2^2}+1\right)} & n > L \end{cases} . \tag{G.26}$$

G.3 Huber Norm

In deriving the MAP solution (Appendix H) the gradient of the natural logarithm of the prior probability distribution (equation 6.37)

$$\frac{\partial J_H}{\partial x_n} = \frac{\partial}{\partial x_j} \sum_{i=1}^{N} \ell\left(u_i\right), \tag{G.27}$$

where $u_i = \frac{x_i}{\sigma_{x_i}}$, and

$$\ell\left(u_i\right) = \begin{cases} \frac{1}{2} u_i^2 & |u_i| \leq \epsilon \\ \epsilon \left|u_i\right| - \frac{\epsilon^2}{2} & |u_i| > \epsilon \end{cases} , \tag{G.28}$$

is required. Using the chain rule

$$\frac{\partial J_H}{\partial x_n} = \frac{\partial u_n}{\partial x_n} \frac{\partial}{\partial u_n} \left(\sum_{i=1}^{N} \ell\left(u_i\right) \right) = \frac{\delta_{in}}{\sigma_i} \frac{\partial}{\partial u_n} \sum_{i=1}^{N} \ell\left(u_i\right), \tag{G.29}$$

where

$$\frac{\partial}{\partial u_n} \sum_{i=1}^{N} \ell\left(u_i\right) = \begin{cases} u_n & |u_n| \leq \epsilon \\ \epsilon \; sgn\left(u_n\right) & |u_n| > \epsilon \end{cases} , \tag{G.30}$$

so

$$\frac{\partial J_H}{\partial x_n} = \begin{cases} \frac{1}{\sigma_1} \begin{cases} \frac{x_n}{\sigma_1} & \left|\frac{x_n}{\sigma_1}\right| \leq \epsilon \\ \epsilon \; sgn\left(\frac{x_n}{\sigma_1}\right) & \left|\frac{x_n}{\sigma_1}\right| > \epsilon \end{cases} & n \leq L \\ \frac{1}{\sigma_2} \begin{cases} \frac{x_n}{\sigma_2} & \left|\frac{x_n}{\sigma_2}\right| \leq \epsilon \\ \epsilon \; sgn\left(\frac{x_n}{\sigma_2}\right) & \left|\frac{x_n}{\sigma_2}\right| > \epsilon \end{cases} & n > L \end{cases} . \tag{G.31}$$

Since $sgn\left(\frac{x_n}{\sigma_L}\right) = \frac{x_n}{|x_n|}$

$$\frac{\partial J_H}{\partial x_n} = \begin{cases} \frac{1}{\sigma_1^2} \begin{cases} x_n & \left|\frac{x_n}{\sigma_1}\right| \le \epsilon \\ \epsilon \frac{1}{\frac{|x_n|}{\sigma_1}} x_n & \left|\frac{x_n}{\sigma_1}\right| > \epsilon \end{cases} & n \le L \\ \frac{1}{\sigma_2^2} \begin{cases} x_n & \left|\frac{x_n}{\sigma_2}\right| \le \epsilon \\ \epsilon \frac{1}{\frac{|x_n|}{\sigma_2}} x_n & \left|\frac{x_n}{\sigma_2}\right| > \epsilon \end{cases} & n > L \end{cases} . \tag{G.32}$$

In matrix form

$$\frac{\partial J_q}{\partial x_n} = \frac{1}{\sigma_{r_p}^2} \mathbf{Q}_h \mathbf{x}' \tag{G.33}$$

where $\mathbf{Q}_h$ is a diagonal matrix whose elements are

$$Q_{nn} = \begin{cases} \frac{1}{\tilde{\sigma}_1^2} \begin{cases} 1 & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| \le \epsilon \\ \epsilon \frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_1}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_1}\right| > \epsilon \end{cases} & n \le L \\ \frac{1}{\tilde{\sigma}_2^2} \begin{cases} 1 & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| \le \epsilon \\ \epsilon \frac{1}{\frac{|x_n|}{\sigma_{r_p}\tilde{\sigma}_2}} & \left|\frac{x_n}{\sigma_{r_p}\tilde{\sigma}_2}\right| > \epsilon \end{cases} & n > L \end{cases} , \tag{G.34}$$

having made use of the substitution of equation (6.19), $\sigma_1^2 = \sigma_{r_p}^2 \tilde{\sigma}_1^2$ and $\sigma_2^2 = \sigma_{r_p}^2 \tilde{\sigma}_2^2$.

G.4 Long-tailed misfit weights

To facilitate calculating the derivative, equation (8.28) is rewritten as

$$J_p = \sum_{m=1}^{M} \rho_p\left(\chi_m\right), \tag{G.35}$$

where

$$\rho_p\left(\chi_m\right) = \frac{1}{p}\left|\chi_m\right|^p, \tag{G.36}$$

and

$$\chi_m = \frac{\sum_{n=1}^{N} G_{mn} x_n - d_m}{\xi} = \frac{\varepsilon_m}{\xi}, \tag{G.37}$$

where $\varepsilon_m = \sum_{n=1}^{N} G_{mn}x_n - d_m$.Taking the derivative of J_p results in

$$\frac{\partial J_p}{\partial x_j} = 0 = \sum_{m=1}^{M} \frac{\partial \rho_p(\chi_m)}{\partial x_j} = \sum_{m=1}^{M} \frac{\partial \rho_p(\chi_m)}{\partial \chi_m}\frac{\partial \chi_m}{\partial x_j} = \frac{G_{mj}}{\xi}\sum_{m=1}^{M} \frac{\partial \rho_p(\chi_m)}{\partial \chi_m} = \frac{G_{mj}}{\xi}\sum_{m=1}^{M} \psi_2(\chi_m), \tag{G.38}$$

where

$$\psi_2(\chi_m) = \frac{\partial \rho_p(\chi_m)}{\partial \chi_m} = \frac{\partial}{\partial \chi_m}\frac{1}{p}|\chi_m|^p = |\chi_m|^{p-1} sgn(\chi_m) = \frac{\chi_m}{|\chi_m|}|\chi_m|^{p-1} = \frac{\chi_m}{|\chi_m|^{2-p}}, \tag{G.39}$$

so

$$\frac{\partial J_p}{\partial x_j} = \frac{G_{mj}}{\xi}\sum_{m=1}^{M} \frac{\chi_m}{|\chi_m|^{2-p}}, \tag{G.40}$$

or

$$\frac{\partial J_p}{\partial x_j} = \frac{G_{mj}}{\xi^2}\sum_{m=1} \frac{\sum_{n=1}^{N} G_{mn}x_n - d_m}{\left|\frac{\varepsilon_m}{\xi}\right|^{2-p}}. \tag{G.41}$$

In matrix notation equation (G.41) is

$$\boldsymbol{\nabla} J_p = \frac{\mathbf{G}'^T\mathbf{W}_{\ell_p}\mathbf{G}'\mathbf{x} - \mathbf{G}'^T\mathbf{W}_{\ell_p}\mathbf{d}}{\xi^2}, \tag{G.42}$$

where

$$\mathbf{W}_{\ell_p} = diag\left\{ \begin{array}{cc} \epsilon^{p-2} & \left|\frac{\varepsilon_m}{\xi}\right| \le \epsilon \\ \left|\frac{\varepsilon_m}{\xi}\right|^{p-2} & \left|\frac{\varepsilon_m}{\xi}\right| > \epsilon \end{array} \right\}, \tag{G.43}$$

and ϵ is a small number. The inequality $\left|\frac{\varepsilon_m}{\xi}\right| \le \epsilon$ is introduced to protect against the singularity that arises as $\epsilon = 0$ for $1 \le p < 2$.

Appendix H

AVO waveform optimization problem

Bayes' theorem provides a theoretical framework to make probabilistic estimates of the unknown reflectivity attributes $\mathbf{x}$ from uncertain data and *a priori* information. The resulting probabilistic parameter estimates are called the Posterior Probability Distribution Function (*PPDF*). The *PPDF* written symbolically as $P(\mathbf{x}|\mathbf{d},I)$ indicates the probability of the parameter vector $\mathbf{x}$ given the data vector $\mathbf{d}$ (offset dependent reflectivity) and information I. Bayes' theorem

$$P(\mathbf{x}|\mathbf{d},I) = \frac{P(\mathbf{d}|\mathbf{x},I)\,P(\mathbf{x}|I)}{P(\mathbf{d}|I)}, \tag{H.1}$$

calculates the *PPDF* from the likelihood function $P(\mathbf{d}|\mathbf{x},I)$ and an *a priori* probability function $P(\mathbf{x}|I)$. The denominator $P(\mathbf{d}|I)$ is a normalization function which may be ignored if only the shape of the *PPDF* is of interest so

$$P(\mathbf{x}|\mathbf{d},I) \propto P(\mathbf{d}|\mathbf{x},I)\,P(\mathbf{x}|I)\,. \tag{H.2}$$

The most likely estimate occurs at the maximum of the *PPDF*. The uncertainty of the parameter estimate is proportional to the width of the *PPDF*.

Assuming uniform uncorrelated Gaussian noise (for simplicity), the likelihood function may be derived in a similar fashion as in Chapter 3 resulting in

$$P(\mathbf{d}|\mathbf{x}',\sigma_N,I) \propto (2\pi)^{-\frac{N}{2}}\,\sigma_N^{-N}\exp\left(-\sum_{m=1}^{M}\frac{\left(\sum_{n=1}^{N}G_{mn}x_n - d_m\right)^2}{2\sigma_N^2}\right), \tag{H.3}$$

where the total number of data points is equal to the number of offsets times the

number of time samples (interfaces) $M = K \times L$ and the number of unknown parameters is two times the number of time samples $N = 2L$.

Section 6.2.2 developed constraints from three long-tailed *a priori* probability distributions. This derivation uses the Cauchy distribution, but the other derivations follow in a similar fashion. Following Sacchi and Ulrych (1995), the a priori Cauchy distribution (equation G.7)

$$p\left(\mathbf{x}'\right) = \exp\left(-\lambda_0 - \lambda_1 S\left(\mathbf{x}'\right)\right). \tag{H.4}$$

combined with the likelihood function (H.3) into Bayes' theorem (H.2) results in

$$p\left(\mathbf{x}'|\mathbf{d}\right) \propto \left(2\pi\right)^{-\frac{N_d}{2}} \sigma_N^{-N_d} \exp\left(-\lambda_0 - \lambda_1 S\left(\mathbf{x}'\right)\right) \exp\left(-\frac{1}{2}\frac{\left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)^H \left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)}{\sigma_N^2}\right). \tag{H.5}$$

Upon marginalizing the variance σ_N in a similar fashion to Chapter 3, (Appendix E) leads to

$$p\left(\mathbf{x}'|\mathbf{d}\right) \propto \exp\left(-\lambda_0 - \lambda_1 S\left(\mathbf{x}'\right)\right) \left(\left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)^H \left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)\right)^{-\frac{M-1}{2}}. \tag{H.6}$$

The maximum of the objective function is found when the function is stationary $\frac{\partial P}{\partial x_i} = 0$ and convex. It is easier to determine this by taking the logarithm of the probability distribution. Taking the logarithm of equation (H.6) J results in

$$J \propto \left(-\lambda_0 - \lambda_1 S\left(\mathbf{x}'\right)\right) - \frac{M-1}{2} \log\left(\left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)^H \left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)\right), \tag{H.7}$$

which can be written as the sum of two parts

$$0 = \frac{\partial J}{\partial x_j} = -\frac{\partial J_{cauch}}{\partial x_j} - \frac{\left(M-1\right)}{2}\frac{\partial J_{misfit}}{\partial x_j}, \tag{H.8}$$

where

$$\frac{\partial J_{cauch}}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\lambda_0 + \lambda_1 S\left(\mathbf{x}'\right)\right), \tag{H.9}$$

and

$$\frac{\partial J_{misfit}}{\partial x_j} = \frac{\partial}{\partial x_j} \log\left(\left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)^H \left(\mathbf{G}'\mathbf{x}'-\mathbf{d}\right)\right). \tag{H.10}$$

The partial derivative $\frac{\partial J_{cauch}}{\partial x_j}$ is (equation G.25)

$$\frac{\partial J_{Cauchy}}{\partial x_n} = \frac{\lambda_1}{\sigma_1^2}\mathbf{Q}_c\mathbf{x}', \tag{H.11}$$

where the diagonal elements of matrix $\mathbf{Q}_c$ are defined by equation (G.26). The partial derivative $\frac{\partial J_{misfit}}{\partial x_j}$ is

$$\frac{\partial J_{misfit}}{\partial x_j} = \frac{2\left[\mathbf{G}'^T\mathbf{G}'\mathbf{x}'-\mathbf{G}'^T\mathbf{d}\right]}{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}, \tag{H.12}$$

where $\boldsymbol{\varepsilon} = \mathbf{G}'x-\mathbf{d}$. Thus, upon substitution of equation (H.11) and (H.12) into (H.8) results in

$$0 = -\frac{\lambda_1}{\sigma_1^2}\mathbf{Q}_c\mathbf{x}' - \frac{(M-1)}{2}\frac{2\left[\mathbf{G}'^T\mathbf{G}'\mathbf{x}'-\mathbf{G}'^T\mathbf{d}\right]}{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}, \tag{H.13}$$

or

$$\left[\mathbf{G}'^T\mathbf{G}'+\Theta\mathbf{Q}_c\right]\mathbf{x}' = \mathbf{G}'^T\mathbf{d}, \tag{H.14}$$

where

$$\Theta = \lambda_1\frac{\sigma_N^2}{\sigma_1^2}, \tag{H.15}$$

and

$$\sigma_N^2 = \frac{\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{(M-1)}. \tag{H.16}$$

Appendix I

Change of variables in standard form

The eigenvector decomposition of the covariance matrix is

$$\mathbf{C_x} = \mathbf{V\Sigma V}^{-1} = \mathbf{V\Sigma V}^{T}. \tag{I.1}$$

assuming the eigenvectors are orthonormal. Thus, the inverse of the covariance matrix, $\mathbf{W}_x$ eigenvector decomposition is

$$\mathbf{W_x} = \mathbf{C_x^{-1}} = \mathbf{V\Sigma V}^{-1} = \mathbf{V\Sigma}^{-1/2}\left(\mathbf{\Sigma}^{-1/2}\mathbf{V}^{-1}\right) = \left(\mathbf{\Sigma}^{-1/2}\mathbf{V}^{-1}\right)^{T}\left(\mathbf{\Sigma}^{-1/2}\mathbf{V}^{-1}\right). \tag{I.2}$$

The constrained inverse problem

$$\begin{bmatrix} \mathbf{G}' \\ \sqrt{\Theta}\mathbf{W}_x^{1/2} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}, \tag{I.3}$$

may thus be written as

$$\begin{bmatrix} \mathbf{G}' \\ \sqrt{\Theta}\mathbf{\Sigma}^{-1/2}\mathbf{V}^{-1} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}. \tag{I.4}$$

Using the change of variables

$$\mathbf{x} = \mathbf{V}\mathbf{x}', \tag{I.5}$$

and

$$\mathbf{x}' = \mathbf{V}^{-1}\mathbf{x}, \tag{I.6}$$

equation (I.4) becomes

$$\begin{bmatrix} \mathbf{G}' \\ \sqrt{\Theta}\mathbf{\Sigma}^{-1/2} \end{bmatrix} \mathbf{x}' = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}, \tag{I.7}$$

so the weighting matrix is diagonal.

Note that the covariance matrix

$$\mathbf{C}_x = \frac{\mathbf{X}\mathbf{X}^T}{N-1}, \tag{I.8}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N]$ and $\mathbf{x}_n = [r_p, r_s, r_d]^T$, after the change of variables is diagonal

$$\mathbf{C}_{x'} = \frac{\mathbf{X}'\mathbf{X}'^T}{N-1} = \frac{\mathbf{V}^{-1}\mathbf{X}\mathbf{X}^T\left(\mathbf{V}^{-1}\right)^T}{N-1} = \mathbf{\Sigma}, \tag{I.9}$$

since $\mathbf{C}_x = \frac{\mathbf{X}\mathbf{X}^T}{N-1} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^T$.