Estimation and removal of complex near surface effects in seismic measurements

Estimation and removal of complex near surface effects in seismic measurements

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 10 mei 2007 om 12.30 uur

door Cornelius Okko Henri Hindriks

Mijnbouwkundig ingenieur
geboren te Haarlem.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. A.J. Berkhout

Toegevoegd promotor:
Dr. ir. D.J. Verschuur

Samenstelling Promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. A.J. Berkhout	Technische Universiteit Delft, promotor
Dr. ir. D.J. Verschuur	Technische Universiteit Delft, toegevoegd promotor
Prof. dr. A. Gisolf	Technische Universiteit Delft
Prof. dr. ir. C.P.A. Wapenaar	Technische Universiteit Delft
Prof. dr. D. Gajewski	Universität Hamburg
Prof. dr. J. Virieux	Université de Nice Sophia Antipolis
Dr. J.W. de Maag	Shell International Exploration and Production

ISBN 978-90-8559-292-1

Printed in The Netherlands by: Optima Grafische Communicatie

Aan mijn ouders

Contents

Notation and terminology **ix**

1 Introduction **1**
1.1 Seismic processing and near surface effects 1
1.2 Problem statement 3
1.3 Outline 5

2 Near surface corrections, an overview **9**
2.1 Introduction 9
2.2 Historical background 10
2.3 Datum statics 11
2.4 Near surface model estimation 14
2.4.1 Uphole surveys 14
2.4.2 Refraction analysis 18
2.5 Residual statics 24
2.6 Error effect of statics 26
2.7 Static corrections by focusing 27

3 From static corrections to a wavefield solution **35**
3.1 Introduction 35
3.2 Simple near surface background model 35
3.3 Near surface model with abruptly changing anomalies 39
3.4 Near surface model with smoothly changing anomalies 41
3.5 Conclusions 44

4 Focusing of seismic data **47**
4.1 Introduction 47

4.2 Two-way integral representations . . . 47
4.2.1 The wave equation and reciprocity . . . 48
4.2.2 The Kirchhoff integral . . . 49
4.2.3 The Rayleigh integrals . . . 51
4.3 One-way integral representations . . . 53
4.3.1 One-way reciprocity . . . 53
4.3.2 Forward extrapolation . . . 54
4.3.3 Inverse extrapolation . . . 56
4.4 Physical field measurements, the WRW-matrix representation . . . 57
4.4.1 Derivation of the **WRW**-model . . . 57
4.4.2 Surface free datum response . . . 58
4.4.3 Free surface datum response . . . 60
4.4.4 Two-way target response . . . 60
4.5 Near surface redatuming . . . 64
4.5.1 Derivation of the inverse propagation operators . . . 64
4.5.2 Focusing operators . . . 66
4.6 Focusing operator updating . . . 70
4.6.1 Principle of equal traveltime . . . 70
4.6.2 Fermat modeling . . . 75
4.7 Synthetic Example . . . 79
4.7.1 Operator updating . . . 79
4.7.2 Analysis and processing of final operators . . . 81
4.7.3 Focusing analysis . . . 83
4.8 Conclusions . . . 86

5 Intialization of focusing operators 89
5.1 Introduction . . . 89
5.2 Fresnel zone reconstruction in case of a complex near surface . . . 89
5.3 Near surface impact on wave propagation . . . 92
5.3.1 Domain of smallest operator complexity . . . 94
5.4 Construction of initial focusing operators . . . 97
5.4.1 Forward model and objective . . . 97
5.4.2 Inversion for CSP operators by Simulated Evolution . . . 102

6 Traveltime operator inversion, a tomographic solution 109
6.1 Introduction . . . 109
6.2 Traveltime inversion of CFP operators . . . 110
6.3 Forward modeling of seismic traveltimes . . . 114
6.3.1 Raytracing . . . 115
6.3.2 Shortest path methods . . . 116
6.3.3 Finite difference methods . . . 118
6.4 Parameterization of the velocity model . . . 120
6.5 Comparison of raytracing and grid-based methods . . . 123
6.6 Travel-time inversion for a complex near surface, an example . . . 126

6.7 Conclusions . . . 129

7 True amplitude redatuming 133
7.1 Introduction . . . 133
7.2 Least-squares redatuming . . . 134
7.2.1 Redatuming as two focusing steps . . . 135
7.2.2 Effect of wavefield deficiencies on redatuming . . . 137
7.2.3 Least-squares implementation of redatuming . . . 138
7.2.4 Example of least-squares redatuming . . . 140
7.3 True amplitude redatuming strategy . . . 143
7.3.1 Preprocessing for least-squares redatuming . . . 144
7.3.2 Operator amplitude determination for least-squares redatuming 150
7.4 Example . . . 157
7.5 Conclusions . . . 162

8 Examples 165
8.1 Introduction . . . 165
8.2 The Middle East model . . . 165
8.2.1 CFP processing . . . 167
8.2.2 CFP redatuming . . . 174
8.3 The small wavelength near surface model . . . 179
8.3.1 CFP processing . . . 181
8.3.2 CFP redatuming . . . 184

9 Conclusions and recommendations 189
9.1 Conclusions . . . 189
9.2 Recommendations . . . 191

A Two-way and one-way wavefields 193
A.1 Introduction . . . 193
A.2 Two-way acoustic wave equations . . . 194
A.2.1 Two-way impulse response . . . 194
A.2.2 Reciprocity relations for two-way wavefields . . . 196
A.3 One-way acoustic wave equation . . . 199
A.3.1 From two-way to one-way wavefields . . . 199
A.3.2 One-way reciprocity theorem . . . 202
A.3.3 One-way impulse responses . . . 203

B Bayes based parameter estimation 207
B.1 Introduction . . . 207
B.2 Bayes' rule . . . 207
B.3 Characterization of Probability Density functions . . . 210
B.4 MAP estimator . . . 211
B.4.1 Probability density function . . . 211

B.5 Linear inverse formulation 215
B.5.1 Comparison of performance between hybrid norm and Gaussian norm 217
B.5.2 Krylov subspace methods for linear system inversion 219
B.6 Non-linear optimization 221
B.6.1 Local linearization 222
B.6.2 Error and resolution 224
B.6.3 Scaling 225
B.6.4 Simulated evolution 228
B.6.5 Residual statics example 231
B.7 Sparse operators 233

Bibliography **239**

Summary **247**

Samenvatting **251**

Dankwoord **255**

About the author **257**

Notation and terminology

Notations

In this thesis use is made of continuous and discrete wavefield descriptions.

In case of continous wavefield formulations, the Einstein summation convention is adopted. According to this convention, the lower case latin subscripts $\{k, l, m, n, p, q, r, s\}$ are assigned to the values 1, 2 and 3 unless stated otherwise and the lower case Greek subscripts are assigned to the values 1 and 2. The summation convention applies to repeated subscripts,

$$a_k b_{kl} = \sum_{k=1}^{3} a_k b_{kl}.$$

For discrete formulations, the following conventions are used: Scalar quantities are denoted by normal simbols, vector quantities are denoted by lower case bold symbols, matrix quantities are denoted by upper case bold symbols. Elements of vectors and matrices take subscript indices according to their entry locations, taking into account the convention according to their dimension. Let for instance,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}.$$

Then

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \mathbf{b}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \mathbf{c},$$

and

$$A_{21} = 1 = b_1, \quad A_{12} = 2 = c_1.$$

Complex conjugation is denoted by $\cdot^*$, transposition is denoted by $\cdot^T$ and complex conjugate transposition by $\cdot^H$. For example,

$$\mathbf{P} = \begin{bmatrix} a_{11} + ja_{12} & b_{11} + jb_{12} \\ a_{21} + ja_{22} & b_{21} + jb_{22} \end{bmatrix}, \quad \mathbf{P}^* = \begin{bmatrix} a_{11} - ja_{12} & b_{11} - jb_{12} \\ a_{21} - ja_{22} & b_{21} - jb_{22} \end{bmatrix},$$

and,

$$\mathbf{P}^T = \begin{bmatrix} a_{11} + ja_{12} & a_{21} + ja_{22} \\ b_{11} + jb_{12} & b_{21} + jb_{22} \end{bmatrix}, \quad \mathbf{P}^H = \begin{bmatrix} a_{11} - ja_{12} & a_{21} - ja_{22} \\ b_{11} - jb_{12} & b_{21} - jb_{22} \end{bmatrix}.$$

Transformations

Throughout the thesis extensive use is made of domain transformations.
The temporal Fourier Transform pair is defined as,

$$\begin{aligned} f(\mathbf{x},\omega) &= \mathcal{F}_t^+\{\hat{f}(\mathbf{x},t)\} &&= \int_{-\infty}^{\infty} \hat{f}(\mathbf{x},t)\exp(-j\omega t)\mathrm{d}t, \\ \hat{f}(\mathbf{x},t) &= \mathcal{F}_t^-\{f(\mathbf{x},\omega)\} &&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\mathbf{x},\omega)\exp(-j\omega t)\mathrm{d}t, \end{aligned}$$

in which t resembles time and ω the temporal frequency. The two-dimensional spatial Fourier Transform pair is defined as,

$$\begin{aligned} \bar{f}(\mathbf{k},x_3,\omega) &= \mathcal{F}_x^+\{f(\mathbf{x},\omega)\} &&= \int_{(x_1,x_2)\in\partial\mathbb{R}^2} f(\mathbf{x},\omega)\exp(+jk_\alpha x_\alpha)\mathrm{d}^2x_\alpha, \\ \hat{f}(\mathbf{x},\omega) &= \mathcal{F}_x^-\{\bar{f}(\mathbf{k},x_3,\omega)\} &&= \frac{1}{4\pi^2}\int_{(k_1,k_2)\in\partial\mathbb{R}^2} \bar{f}(\mathbf{k},x_3,\omega)\exp(-jk_\alpha x_\alpha)\mathrm{d}^2k_\alpha, \end{aligned}$$

in which $\mathbf{x}$ resembles the vector describing the spatial coordinates and $\mathbf{k}$ the 2-dimensional vector describing the spatial wavenumbers. In Chapters 4 and 6 use is made of Radon transformed data. The forward Radon Transform is defined as,

$$\tilde{f}(\mathbf{p},x_3,\tau) = \mathcal{R}\{\hat{f}(\mathbf{x},t)\} = \int_{(x_1,x_2)\in\partial\mathbb{R}^2} \hat{f}(\mathbf{x},t=\tau+p_\alpha x_\alpha)\mathrm{d}^2x_\alpha$$

and

$$\mathcal{F}_t^+\{\tilde{f}(\mathbf{p}\omega,x_3,\tau)\} = \bar{f}(\mathbf{k},x_3,\omega),$$

in which τ is the zero-offset intercept time and $\mathbf{p}$ contains the lateral slownesses.

Thus, the following notational conventions are adopted,

domain	time	frequency	wave-number	wave-number frequency	Radon
function	$\hat{f}(x,t)$	$f(x_1,\omega)$	$\hat{\bar{f}}(k_1,t)$	$\bar{f}(k_1,\omega)$	$\tilde{f}(p_1,\tau)$

The same conventions hold for transformed matrix and vector quantities.

Terminology

The following terminology is used:

- **AVP, Amplitude Versus Rayparameter**
- **CDP, Common Depth Point**
- **CFP, Common Focus Point**
- **CMP, Common Mid Point**
- **CSP, Common Surface Point**
- **DTS, Differential Time Shift**
- **Focusing in detection**
- **Focusing in emission**
- **NMO, Normal Move Out**
- $\cdot^{+}$, denotes a downgoing wavefield.
- $\cdot^{-}$, denotes an upgoing wavefield.
- $\cdot_{fr}$, denotes a free surface response.
- $\mathbf{D}$, denotes source or detector response operator.
- $\mathbf{F}$, denotes a spatially discrete inverse wavefield propagator.
- ω, denotes the angular frequency.
- $\mathbf{P}$, denotes the spatially discrete measured seismic data for one frequency.
- S, denotes a source function.
- $\mathbf{W}$, denotes a spatially discrete forward wavefield propagator.
- $\mathbf{x}_{src}$, denotes the source location.

- $\mathbf{x}_{rcv}$, denotes the receiver location.
- $\mathbf{x}_{cfp}$, denotes the common focus point location.
- $\mathbf{X}$, denotes the spatially discrete response of the earth.

1

Introduction

1.1 Seismic processing and near surface effects

The seismic method aims at constructing an image of the subsurface in a non-destructive way; Yilmaz (1987) gives an excellent overview on the traditional seismic processing techniques. The most conventional way of obtaining the data for construction of such an image is placing a sound source just beneath the surface, such as a vibroseis or dynamite at land or an airgun array at sea, and record the signal back into receivers placed at or near the surface as well. The signals that are recorded consist of the direct arrival of the emitted signal as well as reflected and refracted energy which is generated in the subsurface by contrasts in density and wave-propagation velocities. By conducting many of these single source experiments, the source being positioned at different locations, a subsurface coverage is obtained which aims at illuminating the subsurface contrasts from many angles. The constructed image represents the reflection properties of the earth.

The seismic method is often used under the assumption of smoothly varying material properties in the lateral directions. A traditional seismic processing flow consists of the following steps (Yilmaz, 1987):

- Removal of waves that travel along the surface.
- zero-phasing the reflection wavelets.
- removal of multiple energy.
- resorting the data into Common Mid Point gathers.
- Analysis of the propagation velocities.
- Normal Move-Out correction.

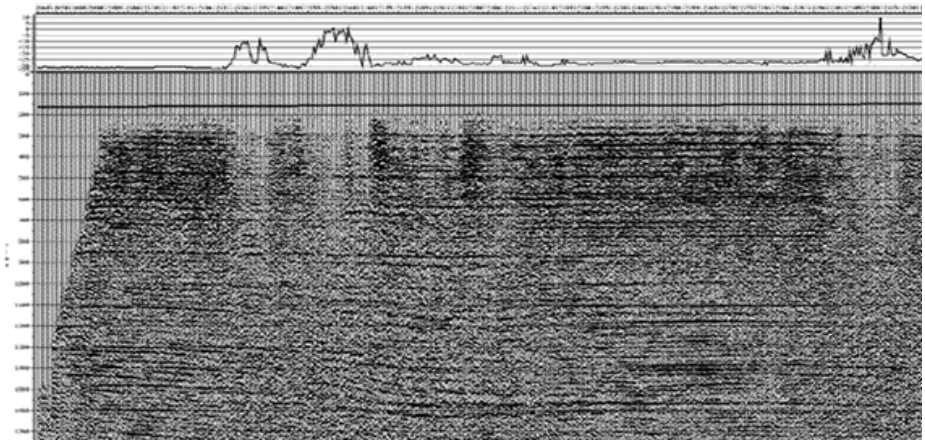

Figure 1.1: *Stack of Middle East land data, courtesy Petroleum Development Oman, contaminated by near surface effects. On top of the stack the applied mean static corrections are displayed. Large static corrections indicate the presence of dunes.*

- Stacking of the data.
- Migrating the data.

In land-data, the near surface part of the earth can severely degrade the quality of the seismic measurements and influence all the subsequent processing steps in a negative sense. Due to lateral variations within the near-surface, the data is distorted close to the source and receivers. If these variations are strong, amplitude changes, time shifts and scattered noise are observed. Especially processes involving velocity analysis, normal move-out correction and migration of the data are degraded. Therefore, a strong need exists to remove the effects of the near surface. Cox (1999) gives an elaborate evaluation on current solutions for complex near surface problems. Commonly used strategies aim at removal of these effects by applying surface-consistent time shifts and amplitude corrections, irrespective of offset and recording time. Hence, the term 'statics' is used. Figure 1.1 shows a typical stack of seismic data acquired in a Middle East desert environment after application of static corrections of which the mean values are displayed on top of the stack. In most areas, the static procedures have indeed enhanced the data, as can be concluded from the continuity of deeper reflections. However, it is clear that the surface-consistent assumptions can break down. Figure 1.1 shows, especially where the static corrections are relatively large due to the presence of sanddunes, dim spots occur and the continuity of reflection events has suffered.

Instead of applying so called surface-consistent static corrections, the most com-

plete corrections will be derived by considering the full propagation characteristics of the near surface; the data can then be corrected through a wave equation based redatuming (Tegtmeier-Last, 2007). This was already recognized by Reshef (1991). His problem did not specifically deal with sub-surface inhomogeneities close to the sources and receivers but with a rugged topography only; Although no true velocity problems were present, the rugged terrain posed similar problems to the subsequent processing steps if it was not dealt with properly. Reshef (1991) redatumed the data to a flat surface through wave-equation redatuming using a simple background velocity. Rajasekaran and McMechan (1995) went one step further and migrated the data directly from a rugged topography after estimating a near surface model first, by using first-break tomography.

Quite often, however, the estimated near surface models are not adequately describing the near surface heterogeneities. In practice, one would like to avoid a model description as long as possible. An opening to conceptually new methodologies was provided by Taner and Berkhout (1997; 1998), who considered the wavefield to heal from near surface inhomogeneities away from source and receiver. Using the CFP technology introduced by Berkhout (1997a), one would in principle be able to obtain the near surface propagation characteristics without explicitly estimating and defining a near surface model. Bakulin and Calvert (2004) use a comparable way of reasoning by directly measuring the near surface propagation characteristics by recording the signals emitted by surface sources at downhole receiver stations. These physically measured propagation characteristics are then used to redatum the data to a sparse redatumed data-set.

In this thesis the CFP method, as introduced by Berkhout (1997a), is used to estimate the propagation characteristics from the surface data and redatum the data to a full data-set underneath the subsurface. By using the CFP methodology, a so called focusing operator can be estimated that contains the propagation characteristics from a single point underneath the weathered layer toward all surface locations. Such an operator can be obtained for multiple subsurface locations.

1.2 Problem statement

Sherriff (1991) states the aim of the statics procedures as follows,

> "... corrections applied to seismic data to compensate for the effects of variations in elevation, weathering thickness, weathering velocity or reference to a datum."

> "Determine the reflection arrival times which would have been observed if all measurements had been made on a (usually) flat datum with no weathering or low velocity material present."

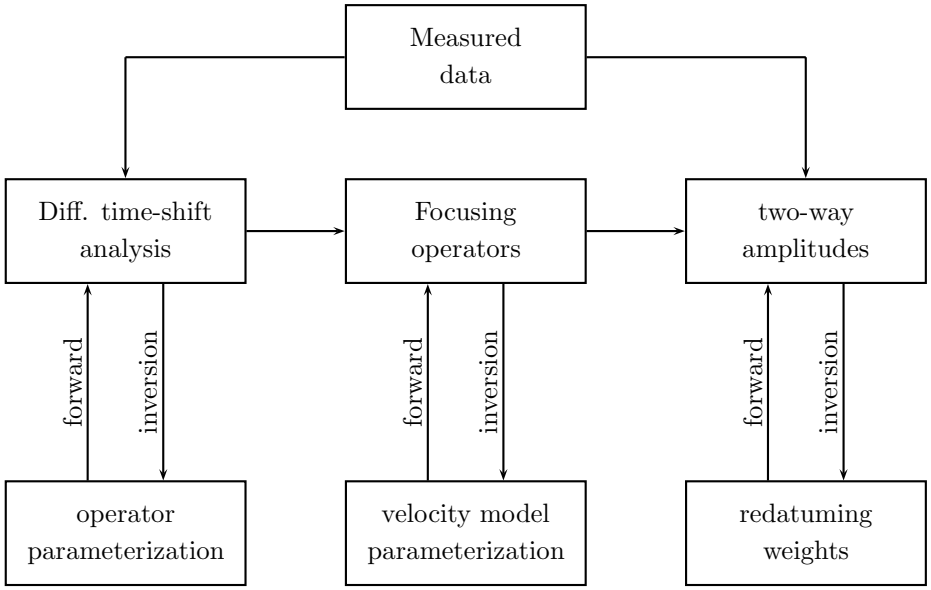

Figure 1.2: *To be able to compute data at a datum underneath the near surface, three steps need to be taken: Operators need to be determined from the measured data by a differential time-shift analysis (see Chapter 4), from the determined operators a velocity model needs to be estimated (see Chapter 6) and redatuming weights need to be determined from the amplitudes of the measured data (see Chapter 7). All steps involve an inversion scheme.*

In this thesis we will take the second quotation in a even broader sense, since near surface effects not only manifest themselves in single time shifts, but in multidimensional operators that include amplitude effects, focusing/defocusing effects, triplication, diffraction patterns, etc. Therefore the aim is redefined as,

> "Determine the recorded reflection energy which would have been observed if all measurements had been made on a datum without the influence of the weathering material present."

This statement implicates an optimization/inversion procedure in which the input is defined to be the recorded data and the output the corrected data and the correction operators.

What is needed for a good inversion result are a forward seismic model, an error criterion, an update algorithm and a quality assessment of the results.

To handle near-surface problems by the CFP method, three stages can be defined as depicted in Figure 1.2:

- Determination of focusing operators (one-way Green's functions).
- Estimation of an effective velocity model.
- Computation of redatuming results.

At each stage we are looking for a set of parameters that describe smooth traveltime operators, a velocity model that is consistent with these operators and redatuming weights respectively. Determination of the involved parameters requires solving an inverse problem.

For the parameterization of the traveltimes of the focusing operators and the redatuming weights, use will be made of the concept of wavefield healing as described by Taner and Berkhout (1998). This concept is depicted in Figure 1.3, which shows the wavefield recorded at multiple surface locations and excited by a pointsource located underneath the near surface in Figure 1.3A and the wavefield recorded at multiple receiver locations underneath the near surface and excited by a pointsource at the surface in Figure 1.3B.

Clearly, the wavefield suffers from the near surface inhomogeneities close at the surface but recovers from those inhomogeneities in the deeper subsurface. This concept is not violating reciprocity as reciprocity relates a single detector with a single source only. For instance the centre traces are exactly the same for the two records, as displayed in Figure 1.3C. In a sense, the concept of wavefront healing allows a parametrization of smooth operators describing the propagation characteristics from a single surface location through the near surface into the deeper subsurface. This property will be exploited in this thesis.

1.3 Outline

In Chapter 2 an extensive overview is given of existing and commonly used methods to compute static corrections in the situation of near surface problems. The chapter ends with the proposal of a new method to determine surface consistent time shifts.

In Chapter 3 the shortcomings of using statics is demonstrated. The chapter also describes how redatuming can improve the full processing cycle.

In Chapter 4 the acoustic wave-equations are revisited, leading to the so called WRW-model. The WRW-model elegantly describes the reflection measurements, leading to a formulation of, among others, surface-related multiple prediction. More important for this thesis, it is the basis of a redatuming scheme. The WRW-model provides a number of criteria which can be used to estimate operators that describe

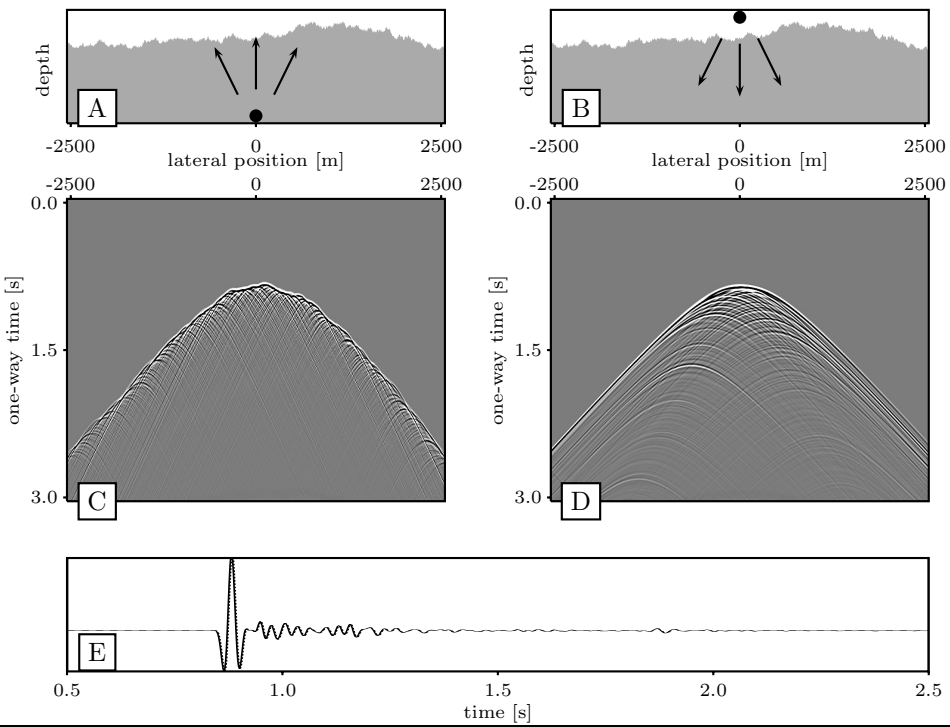

Figure 1.3: *The shotgather of (C) is obtained by recording the wavefield excited by a single source, placed underneath a complex near surface, at multiple receiver locations placed at the surface, as shown in (A). The shotgather of (D) is obtained by recording the wavefield excited by a single source, placed at the surface, at multiple receiver locations placed underneath the same complex near surface as shown in (B). The two middle traces overly each other exactly, as shown in (E).*

the near surface characteristics without the explicit need of a near surface model. These will be described in Chapter 4 as well.

In Chapter 5 use is made of the healing concept as illustrated by Taner and Berkhout (1998) to parametrize the operators that describe the near surface propagation. The chapter shows how, by using a nonlinear optimization scheme, initial operators are found which already adequately contain the effects of near surface anomalies.

Once the final operators have been determined, these operators can be used to estimate a near surface model as described in Chapter 6. It should be emphasized that the model is consistent with the traveltimes of the operators. Besides an effective velocity profile of the near surface, the model estimation also provides the depth of

the new datum.

In Chapter 7 it is shown how the amplitudes of the one-way wave propagation operators can be derived from the traveltime operators in combination with the measured data. These amplitudes will then be used as weights within a weighted least-squares redatuming scheme to determine the data at a level underneath the complex near surface (the new datum).

Chapter 8 illustrates the workflow on two examples. One example demonstrates the applicability on low frequency complexities, such as sand-dunes and wadis, which are typical features within desert operations. The other example contains higher frequency complexities. This example is part of a large model which is used to test the performance of static algorithms available on the market (courtesy BP).

Finally in Chapter 9 conclusions are drawn and recommendations for further research are given.

2

Near surface corrections, an overview

2.1 Introduction

The seismic method, aimed at the exploration and production of Hydrocarbons, is designed to image contrast in elastic properties within the earth under the assumption that these contrasts are embedded within a generally smoothly changing earth. As the seismic method comes down to the determination of the acoustic impulse response of the earth, the design is generally based on a proper discretization of the signal measurements. This is needed both from a practical point of view as well as an economic point of view. Practical in the sense that it is not possible for instance to have a continuous spatial measurement of the reflected wavefield. Economic in the sense that a very dense set of recorders would lead to an enormous amount of data and a tremendous effort in placing (land) or towing (sea) the equipment.

As the earth is a dispersive medium with a general frequency recovery of about 100 Hz for the aimed targets and with acoustic compressional velocities in the range of 1500 to 5000 m/s, in the light of measurement properties for signal processing, we can suffice with a spatial distribution of sources and receivers with an approximate interval of 20 to 50 m. and a time sampling rate of 1 to 4 ms. However, on land the near surface can be very rugose due to the presence of mountains, sand dunes, carsts, man-made obstacles, etc. such that the assumption of a smoothly varying medium and the assumed range of compressional velocities brakes down, which has a detrimental effect on the recovery of the seismic impulse response of the earth. In other words, the setup of the seismic experiment is just not equipped for a proper measurement of the near surface effects. Moreover, these so-called near surface effects can occur on all scales. Within high frequency laboratory experiments

on rock-samples for instance, the effects can be caused by small scale impurities.

As these near surface effects mainly appear close to source and receiver, in practice they are accounted for by estimating correction factors (mainly in phase, but also in amplitude) for each surface location such that the coherency between subsequent spatial measurements is optimized. As these correction factors are applied on a complete trace, independent of the measuring time, they are also known as *static correction* or shortly *statics*.

In this chapter an overview is given on currently used methods to estimate these corrections, followed by an analysis on when these methods will break down. A large part of the following sections summarizes the work of Cox (1999), who gives a very extensive and elaborate discussion on the many aspects of statics.

2.2 Historical background

As the historical overviews found in Marsden (1993a; 1993b; 1993c) show, the problems and difficulties of complex near surface heterogeneities were already recognized in the early days of seismic exploration. In his paper on the invention of the reflection seismograph, Karcher (1987) describes his involvement in the first reflection seismic experiments of which the log-book notifies on December 1-23, 1921 the "Experiments to determine methods to make time corrections for weathered sections of soil". Weatherby (1971) points out that already in the early work "additional shots were always taken in shallow shot holes near the recording spread. These gave in effect a short refraction profile which was solved as a simple two-layer case." From here on it became common practice to compute weathered layer effects by refraction methods. Due to the small amount of data (24-trace, 6-fold) it was possible to perform manual refraction analysis on refraction surveys. Also the so-called uphole surveys (see section 2.4.1) became a popular technique to define near surface anomalies.

By 1970, seismic processing has become more and more computer intensive, which has its effect on statics computations as well. Due to the evolution of the CDP technique (Mayne, 1956) since the 1950's, the redundancy of CDP data is used to compute residual statics (Taner et al., 1974; Booker et al., 1976). As the number of channels and fold increased, and surface sources where introduced such as vibroseis, manual refraction analysis and uphole surveys became impractical and economically unfeasible. Thus the statistical approach of most residual static procedures became common practice to solve all near surface problems. As mentioned by Schneider (1971) statistical methods however need to be employed carefully when S/N ratios are low, spatial frequencies are low and if the deviations are larger than half the period of the dominant reflection event. In those cases for instance cycle skipping can easily occur. Moreover low spatial frequencies in the near surface can introduce

artificial structures. Also mentioned by Schneider (1971) in his review on the development of the seismic method is his opinion that the general near surface problem is understood, the problem is to solve it economically.

With the further growth of the number of recording channels, the decrease of the group intervals and the increase of the array lengths in the 1980's, refractions on reflection records show less attenuation and full analysis of refraction energy became possible, giving rise to the revival of the original datum/field statics with a higher accuracy.

2.3 Datum statics

Datum statics are surface consistent statics computed from an estimated near surface model. The aim of the datum statics is to synthesize the sources and receivers along a datum, which does not suffer from near surface effects, as if the data was originally recorded along this datum. Sometimes datum statics are referred to as field statics, as for instance topology effects can readily be accounted for during acquisition under certain assumptions of the near surface velocities.

The main assumption made is that due to large velocity contrasts, raypaths travel vertically through the near surface. The principle of the datum statics is illustrated by Figure 2.1. Traveltime corrections are in principle computed for each station (source and receiver) from acquisition surface to a chosen datum. The most logical choice for the datum, would be a flat datum underneath the near surface. However, any other convenient datum could be possible as well, such as depicted in Figure 2.2. In case of Figure 2.2 the total static corrections, T_A, T_B, T_C for the three points A,B,C would be,

$$T_A = -t_{Aw} - t_{Ad}, \tag{2.3.1}$$

$$T_B = -t_{Bw} + t_{Bd}, \tag{2.3.2}$$

$$T_C = -t_{Cw} + t_{Cd}, \tag{2.3.3}$$

in which t_{Aw} is the weathering correction for location A and t_{Ad} is the elevation correction for location A. The weathering correction corrects for the thickness and velocity through the overburden to the base of the overburden layer, A_b. The elevation correction corrects for the distance between the base of the overburden and the chosen datum, $\|A_d - A_b\|$, and a known subsurface velocity.

In general, the datum can be arbitrarily chosen such that it minimizes the effects of near surface heterogeneities and anomalies when a simple time shift is used for each trace to virtually place shot and receiver downward to the datum.

It can be observed in Figure 2.1 that, although the datum corrections aim to correct the source and receiver time such that the source and receivers are synthetically

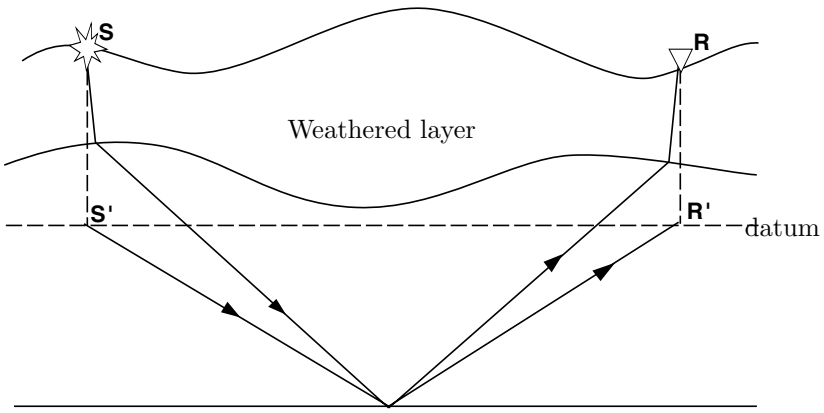

Figure 2.1: *Datum statics correct traveltimes such that the source location S and receiver location R are virtually placed along a user defined datum at respectively location S' and R'. These corrections are generally done in a surface-consistent way, which means that the corrections are independent of the offset and depth. The figure clearly shows that the correction also implies a change of raypath, which is often neglected in practice.*

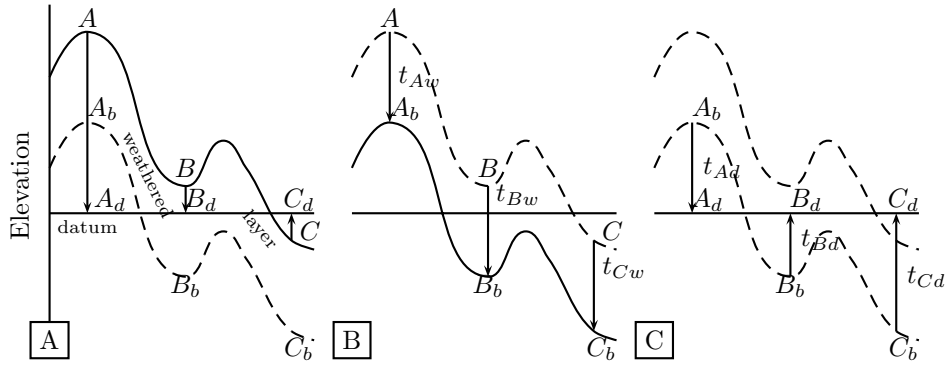

Figure 2.2: *Computation of static correction T_A, T_B and T_C for an arbitrary datum; A) Datum statics tend to replace the velocity of the weathered layer with the velocity of the layer underneath the datum. B) For that purpose, first times from the specific source and receiver locations are corrected to the base of the weathered layer with the weathered layer velocity. C) These corrections are followed by a time correction from the base of the weathered layer to the datum, using the base-layer velocity.*

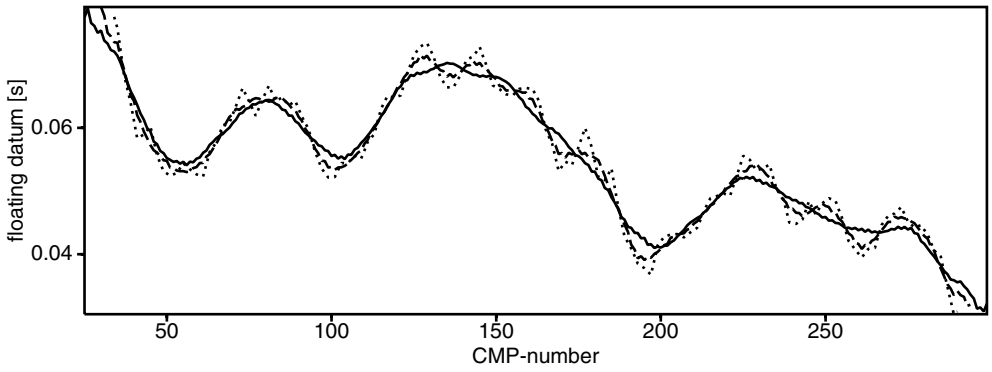

Figure 2.3: *The floating time-datum depends on the fold of the CMP gathers. As for each CMP, the floating datum is chosen to be the mean static within the CMP gather, the floating datum is basically a filtered version of the static corrections. The dotted line equals to the computed single fold static at each CMP. The dashed line depicts the floating datum for 24-fold data, the solid line depicts the floating datum for 48-fold data.*

positioned at the new datum, the true raypaths and the assumed raypaths differ. For that purpose the floating datum was introduced. Applying datum corrections as computed by Equation 2.3.3 does not take into account hyperbolic move-out, and relies on almost vertical raypaths between surface and datum. If significant move-out is expected or the replacement velocity is wrongly estimated, false structure can be introduced for deeper reflections. The floating datum concept (Rogers (1981)) tends to minimize such type of artefacts. In Profeta et al. (1995) an elaborate discussion can be found on the importance of static corrections to be minimal.

The floating datum can be thought of as an intermediate datum within the CMP process. It serves as a time datum, minimizing the necessary static shifts for each CMP gather by eliminating the mean static shift for each CMP gather. In this way the applied statics are minimal and circumvent stacking velocity corrections beforehand. The resulting time shift from the floating datum to the actual datum is determined by the mean static shift for each CMP gather.

The floating datum correction is applied prior to CMP processing. After application of the floating datum correction, velocity analysis can be performed, followed by stacking. Stacking velocities can then be used to refine the mean datum correction. The stacked section is finally repositioned to the desired datum by applying the mean datum correction or a time migration.

The floating datum computation is equivalent to CMP averaging of the computed datum correction. Figure 2.3 shows the computed datum static corrections for CMP

sorted traces and the resulting mean datum corrections which are equivalent to the time-difference between the floating datum and the desired datum. The static correction are shown for increasing fold (1, 24 and 48) with a constant spatial sampling distance. For higher folds, the floating datum becomes more and more a low frequency result of the computed datum correction which is displayed in Figure 2.3 as well.

As the floating datum correction applies minimum static correction to the traces, followed by a mean correction after stacking, it is only applicable in case the seismic data is subject to poststack migration. These days processing is more and more focused on prestack migration, for which the floating datum principle no longer gives an advantage.

2.4 Near surface model estimation

As described in the previous section, field statics and datum statics (including the elevation statics), fully depend on the estimation of a near surface model in terms of thickness and velocity, as well as a subsurface velocity, just below the base of the overburden layers. Any errors made in the estimation are suppressed by the introduction of the floating datum concept, which allows for instance a replacement velocity update by using stacking velocities. Further improvements over the datum correction and the vertical raypath assumption is covered by residual dynamic and static correction, subject to the next section. The process flow of conventional static computation and application is displayed in Figure 2.4. In this section the conventional near surface model determination is covered, which is mainly done by conducting uphole surveys and refraction analysis.

■ 2.4.1 Uphole surveys

Already in 1930's, use is made of measurements of the first arrival time from a buried dynamite source to the surface, to obtain an estimate of average near surface velocity. These measurements are referred to as *production upholes*. Nowadays, uphole surveys are conducted either with a source placed in a deep shot-hole and receivers placed at the surface in a walk-away configuration (see Figure 2.5A) or with a source placed at the surface and a hydrophone string hung from the borehole (see Figure 2.5B), as described by, for instance, Parry and Lawton (1994). Both configurations carry similarities with Vertical Seismic Profiling recordings as conducted within production seismics. Depending on the complexity of the problem, uphole times are used to compute a near surface model by the following methods.

Zero offset corrected traveltime interpretation

In Figure 2.6 the uphole traveltime interpretation is illustrated. First we consider the simple horizontally layered velocity model as depicted in Figure 2.6A. The recorded

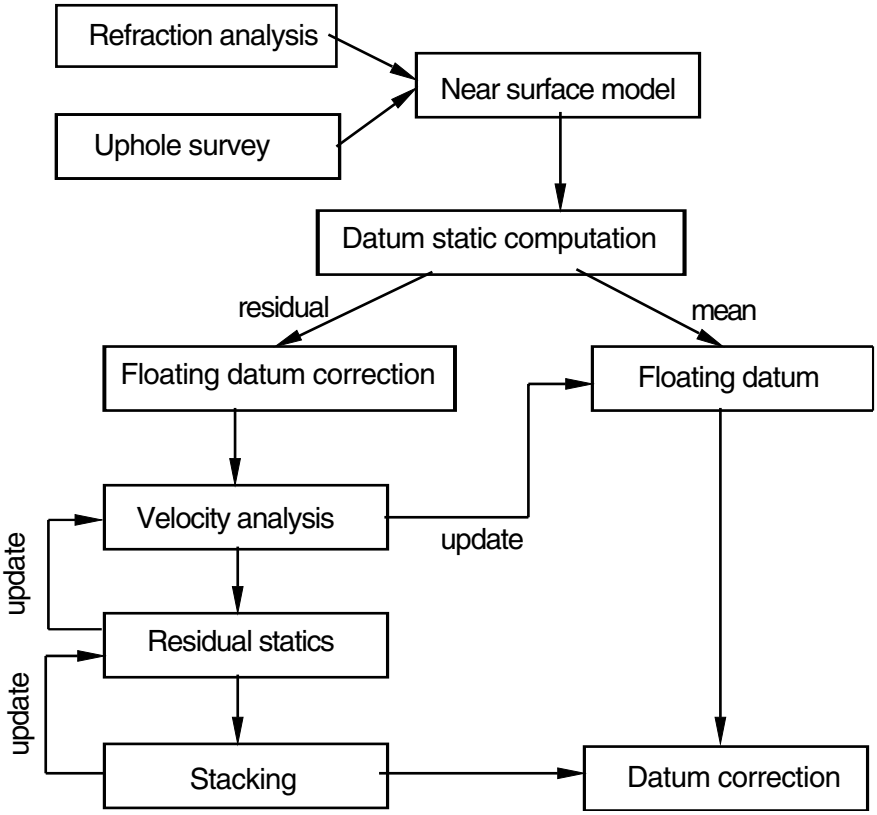

Figure 2.4: *Processing scheme used for static corrections. Refraction analysis and uphole surveys are used to compute an initial near surface model from which datum statics are computed. The floating datum will be constructed, such that the datum statics to be applied are minimal within each CMP gather. An iterative process is then used to determine residual statics and stacking velocities. The resulting velocities can be used to update the floating datum corrections. After convergence of both datum statics and residual statics, the data can be stacked and corrected for the user defined datum, by applying the floating datum times.*

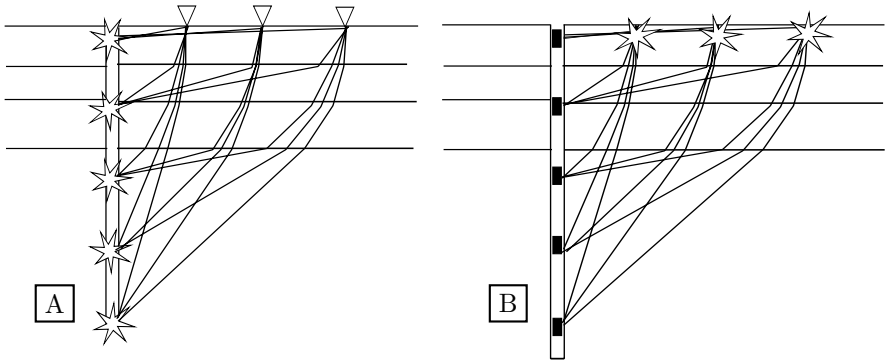

Figure 2.5: *Uphole surveys are conducted within the field, by either placing a series of sources (A) or a series of receivers (B) in a constructed downhole. Their purpose is to measure the first arrivals between the downhole stations and the respectively receivers (A) and sources (B) placed in a walk-away configuration at the surface.*

uphole times of Figure 2.6C, as measured in the horizontally layered model of Figure 2.6A, are corrected for move-out by a simple parabolic correction,

$$T(x_3) = t(\mathbf{x}^S, \mathbf{x}^R) \frac{x_3^S - x_3^R}{\sqrt{|\mathbf{x}^S - \mathbf{x}^R|^2}}, \tag{2.4.1}$$

in which $T(x_3)$ is the move-out corrected traveltime as function of depth x_3 for the measured traveltime $t(\mathbf{x}^S, \mathbf{x}^R)$ between source $\mathbf{x}^S$ and receiver $\mathbf{x}^R$ as displayed by Figure 2.6D. Then the corrected traveltimes are interpreted as the response of a layered model, for which each layer velocity is described by the interpreted slope and depth as depicted by Figure 2.6E.

If the model becomes more complex (Figure 2.6B) then especially for the velocities the results deteriorate from the correct results as displayed by Figure 2.6F, G and H. The process of move-out correction and velocity computation can be subject to an iterative system, in which the applied move-out corrections are enhanced by using the found velocities of the previous iteration in Dix' equation for stacking velocities (Dix, 1955).

Wavefront diagrams

The wavefront diagram was introduced by Meissner (1961)(see also Sherriff, 1991), for which a horizontally layered model is assumed for each uphole. For each source/receiver pair, the measured uphole time is displayed at downhole depth and the offset. Under the assumption of a horizontally layered model, each raypath is mirrored over its

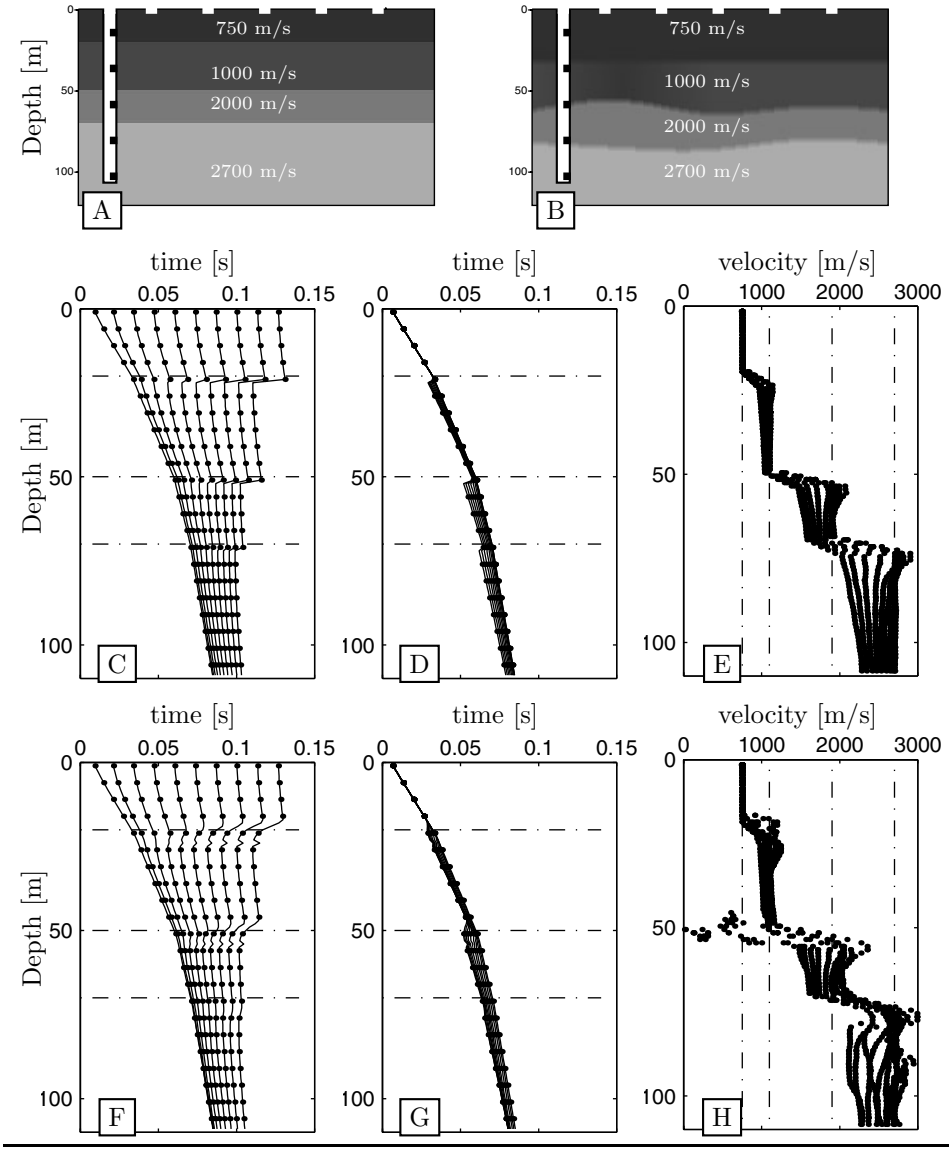

Figure 2.6: *Panel (A) displays a horizontally layered model, including an uphole-survey design. Panel (B) displays the same survey design, placed in a laterally varying earth. First arrivals (horizontally) are measured at 22 downhole receivers with a spacing of 5m, for 10 different shot locations, for which panel (C) displays the measurements for model (A) and panel (F) the results for model (B). After applying spherical divergence corrections (D and G), the 4 intervals are recognized. By taking the derivative of the corrected arrival-time curves (E and H), the interval velocities are retrieved. The correct interval-velocities are indicated by the dotted lines in (E) and (H).*

midpoint, as if it is originating from a source at depth=0 in the borehole. In this way wavefronts are formed by drawing isochrones, virtually excited by a source at depth=0 and offset=0. A near surface model is constructed by determining interval velocities from the gradients of the isochrones as depicted in Figure 2.7.

Again, when the near surface becomes more complex, the raypaths from source to receiver are no longer symmetric with respect to the wave path between virtual source and virtual receiver and the method slowly breaks down for the larger off-sets as can be observed in Figure 2.7F. The concept has proved its usability to determine for instance near surface anisotropy as well.

Tomography

If near surface variations are considerable, the assumption of a horizontally layered earth does not longer hold and results obtained from for instance the wavefront diagram method diverge from the correct model. Tomographic inversion can give adequate but cumbersome results as shown by for instance Nolet (1987) and Stewart (1991).

The tomographic inversion approach can be applied to a broader ranch of measurements such as hole-to-hole analysis as described by for instance Macrides et al. (1988) and Harlan (1990). Since tomographic inversion will also be addressed in the remainder of this thesis, the reader is referred to Chapter 6 for a more detailed description on tomographic inversion.

■ 2.4.2 Refraction analysis

Besides the uphole survey, which is expensive in terms of manpower and acquisition, use is made of first arrival times or refraction analysis within the production seismics as well to determine either a near surface model or equivalent weathering layer delay-times. In general, data is required from so called reversed profiles to compute these corrections. Reversed profiles assure the availability of reciprocal times as depicted in the general set-up of a refraction survey in Figure 2.8. When refracted energy from different sources is measured along a profile in a split spread at several receiver positions, as depicted by Figure 2.8A and 2.8B, the underlying subsurface has a so-called multi-fold coverage. Due to the multi-fold character of the seismic reflection survey, these reciprocal raypaths are implicitly available within the gathered data and can be recovered by combinations of experiments as depicted by Figure 2.8E.

Graphical methods

Graphical methods find their base in the ABC-method (Edge and Laby, 1931; Bahorich et al., 1982). The ABC method determines for a surface location B, in between two locations A and C (see Figure 2.8E), the weathered layer delay-time as follows.

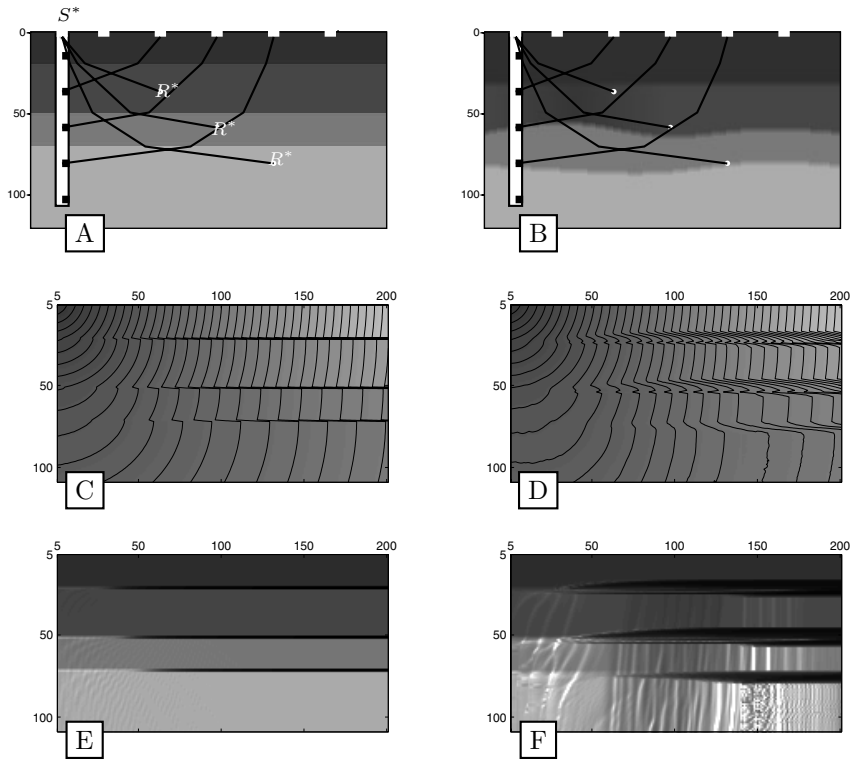

Figure 2.7: *From first arrival times of uphole surveys, wavefront diagrams can be constructed. Under the assumption of a horizontally layered earth, a virtual shot and virtual receivers can be interpreted from the uphole design (panel (A) and (B)). By reordering the recorded times in analogy with the constructed virtual acquisition of the top figures, wavefront diagrams are constructed for the lateral invariant model if panel (C) and the lateral varying model in panel (D). The black lines depict isochrones. The norm of the divergence of the wavefront diagram traveltimes reflect the interval velocities (panel (E) and (F)). The velocities are correctly estimated in comparison with the true model if velocities are not laterally varying, see panel (E). If the velocities are laterally varying as shown in the panel (B) then the resulting velocities will deteriorate as depicted by panel (F).*

For the measured traveltimes $t_{..}$ between subsequent stations we can write,

$$t_{AB} = \frac{l_{AD}}{c_1} + \frac{l_{DE}}{c_2} + \frac{l_{BE}}{c_1}, \tag{2.4.2}$$

$$t_{BC} = \frac{l_{CG}}{c_1} + \frac{l_{GF}}{c_2} + \frac{l_{BF}}{c_1}, \tag{2.4.3}$$

$$t_{CA} = \frac{l_{AD}}{c_1} + \frac{l_{DG}}{c_2} + \frac{l_{CG}}{c_1}, \tag{2.4.4}$$

in which $l_{...}$ refers to a certain path length, c_1 to the weathered layer velocity and c_2 to the sub-weathering velocity. The subscripts refer to segments between the locations as depicted in Figure 2.8E. Then a combination of the relationships from Equation 2.4.4 leads to,

$$t_{AB} + t_{BC} - t_{CA} = \frac{l_{BE} + l_{BF}}{c_1} - \frac{l_{EF}}{c_2}. \tag{2.4.5}$$

Under the assumption of a high velocity contrast, l_{EF} will almost vanish and l_{BE} will almost equal l_{BF}. Thus the weathered layer delay-time for station B can be described as,

$$t_{w,B} = \frac{1}{2}(t_{AB} + t_{BC} - t_{CA}), \tag{2.4.6}$$

$$\approx \frac{l_{BE}}{c_1}. \tag{2.4.7}$$

Hagedoorn (1959) extended the ABC-method to the so-called Plus-Minus method. In this method the delay time computed through Equation 2.4.7 is called the Plus-time and an extra time, the Minus-time is computed according to,

$$t^{-}_{w,B} = \frac{1}{2}(t_{AB} - t_{BC} + t_{CA}), \tag{2.4.8}$$

$$\approx \frac{l_{AD}}{c_1} + \frac{l_{DE}}{c_2}. \tag{2.4.9}$$

By moving station B in between stations A and C, the first term on the right-hand side is constant, while the second term described a line with a slope equal to the slowness of the refractor.

Wavefront and Raypath methods

From a reciprocal refraction analysis survey (measurements along the line between two source locations), time-distance curves can be constructed. These time-distance curves describe the intercept time of a refracted wave-front with the receivers at the sources. From these time-distance curves at fixed Δt the distance along the line can

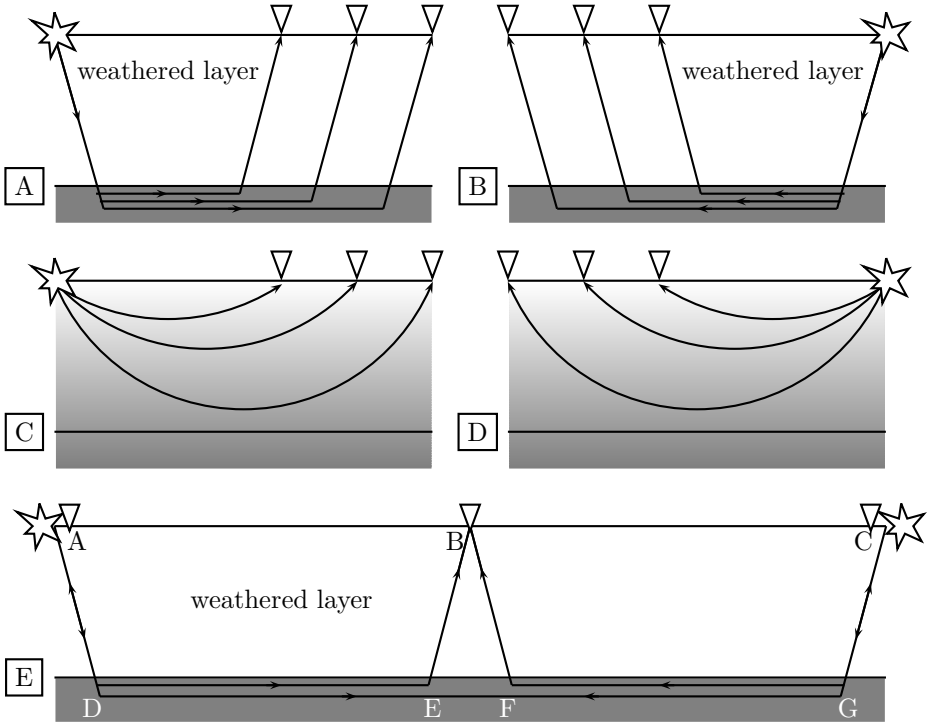

Figure 2.8: *First break times are used for near surface profiling. A) First break times interpreted as refracted ray-times. B) Equivalent reverse profile of A). C) First breaks interpreted as turning rays. D) Equivalent reverse profile of C). E) Schematic of the principle of reverse profile first break times analysis.*

be deducted. These emergence locations are plotted in the graph, for both sources. The wavefront emerging from the location at $N\Delta t$ is constructed by taking the envelope of the arcs with radius $nc_1\Delta t$ centering around the emergence location $n\Delta t$ for $n < N$. This principle is shown in Figure 2.9. By constructing the wavefronts for both sources, the depth of the refractor is found at the points where the sum of the wavefront times of two crossing wavefronts equal to the reciprocal time between the two sources.

For multi-layer problems, this procedure can be repeated along the constructed first refractor, with the time-distance curve for the refraction times of the second refractor and so on. Hagedoorn (1959) showed that the Plus-Minus method resembles the wavefront method. The Plus-times and Minus-times are resembled within the wavefront methods as so called Plus-lines and Minus-lines, as indicated in Figure 2.9. The distance between minus-lines equals both the geometrical distance along

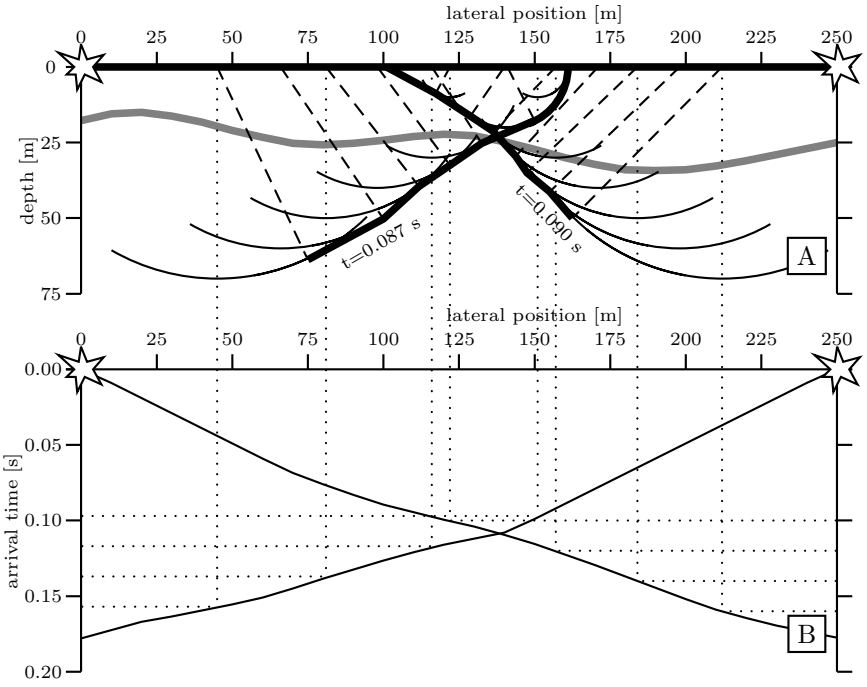

Figure 2.9: *Within the wavefront technique, wavefronts are constructed for two reciprocal refraction surveys. From the measured time-distance curves (B), emergence-locations are defined at fixed Δt (A) for both profiles. From the emergence locations, circular wavefronts are constructed by the aid of an assumed near surface velocity. The envelope of the constructed wavefronts determine the total refracted wavefronts emerging from the two sources. At the point where the two wavefronts are crossing, the boundary of the near surface layer is found. The traveltimes of the two refracted wavefront is chosen, such that the sum is equal to the reciprocal traveltime.*

the recording line, as well as $c_2 * \Delta t$ from which the refractor velocity can be retrieved. Two Plus-lines lie Δt times apart, thus, taking the difference between the time of the Plus-time line at the refractor and the Plus-line at the surface, gives the Plus-time for the particular surface location. From a more general point of view, the wavefront method can be looked upon as a special case of migration of refraction energy.

Numerical methods

The concepts of the graphical methods also have found their ways in the 1990's into numerical and computational methods. Hill (1987) and Taner et al. (1988) implemented the wavefront technology into a downward continuation method applied

in the linear Radon domain, which was eventually implemented in 3D by Taner et al. (1988). As explained in the previous section, the refractor is imaged at locations where constructed wavefronts add to the reciprocal arrival time. Hill (1987) formulated the following imaging condition in 2D,

$$I(x,z) = \int_{\partial\mathbb{R}} \hat{P}(\mathbf{x}^A_{src}, x, z, t) \int_{\partial\mathbb{R}} \hat{P}(\mathbf{x}^B_{src}, x, z, t_R - t)\mathrm{d}t, \tag{2.4.10}$$

in which t_R is the reciprocal time between source (receiver) location A and source (receiver) location B. The image will be constructed since the integral will only have a non-zero contribution when depth continued wavefronts $\int \hat{P}(\mathbf{x}^A_{src}, x, z, t)$ and $\int \hat{P}(\mathbf{x}^B_{src}, x, z, t_R - t)$ coincide. Use is made of the linear Radon transform, to decompose the recorded refraction wavefields $\hat{P}(\mathbf{x}^A_{src}, \mathbf{x}_{rcv}, t)$ and $\hat{P}(\mathbf{x}^B_{src}, \mathbf{x}_{rcv}, t)$ into plane waves and extrapolate them into depth using an assumed velocity function, $c(z)$,

$$P(\mathbf{x}^A_{src}, x, z, \omega) = \frac{1}{2\pi} \int \bar{P}(\mathbf{x}^A_{src}, k_x, z, \omega) \exp\left(-j(k_z - k_x x)\right)\mathrm{d}k_x, \tag{2.4.11}$$

with

$$k_z^2 = \frac{\omega^2}{c^2(z)} - k_x^2. \tag{2.4.12}$$

Instead of implementing wavefront construction algorithms, also full tomography algorithms have been applied to refraction times. Hampson and Russell (1984) introduced a general linearized inversion method applied to refraction data. Their forward model consists of modeling refractors along with refraction traveltimes as calculated according to Figure 2.9 . The method is linearized through perturbation theory. For further details on tomography and perturbation theory the reader is referred to Chapter 5 and references cited there.

Instead of basing the tomography on a layered model, the method can also be extended to a model under the assumption of a gradually increasing velocity as function of depth for each lateral position and use ray-tracing algorithms to model first break times. Since the velocity will increase with depth, modeled first-break times originate from turning rays as demonstrated by figure 2.8C and D. Therefore, these methods are generally referred to as turning ray tomography (Zhu et al., 1992; Epili et al., 2001). A further extension to turning ray tomography and wavefront methods was introduced by Gonzalez et al. (1999), who recognized that at the turning point, the ray-parameter is actually equal to the underlying velocity. In Gonzalez et al. (1999), Osypov (2000) and Osypov (2001) this method is explained in more detail. A remaining problem of most refraction analysis models, as pointed out by Docherty (1992), is the fact that long wavelength ambiguities remain due to small aperture ranges. They claim that through an Eigenvalue analysis of the measured refraction times it can be shown that velocity-depth trade-offs have little effect on the observations. Sensitivity studies of static solutions related to the offset range or aperture can be found in Saleh (1994).

2.5 Residual statics

Datum statics computed from refraction analysis is based on the assumption of vertical raypaths through the near surface. Besides errors introduced by this assumption the resulting corrections are dictated by the resolution of the interpretation. Residual static computations are estimated to remove these residual short wavelength errors and are by no means meant to compute a full statics solution. Accurate datum statics are demanded within conventional processing to remove the longer wavelength static anomalies. Residual statics time-shifts are in most methods based on the basic traveltime equation,

$$T_{ijk} = G_k + S_i + R_j + M_k X_{ij}^2 + N, \tag{2.5.1}$$

in which T_{ijk} is the total synthesized zero-offset time for the trace produced by source i and receiver j, within CMP gather k after NMO-correction. G_k is the structural or geological term representing the two-way reflection time from the reflector to the datum, S_i is the residual source static, R_j is the residual receiver static, $M_k X_{ij}^2$ is a residual move-out term depending on the move-out coefficient M_k and the trace offset X_{ij}. N is a noise component. The main assumptions within Equation 2.5.1 are that the arrival times within an NMO-corrected CMP gather are constant for all offsets. Note that this does not allow for dip within the subsurface and that source and receiver statics are offset and time independent from which the expression *surface consistent* relates. A modified version was presented by Larner et al. (1979), including an extra crossline dip term and making the expression time dependent,

$$T_{ijkh} = G_{kh} + S_i + R_j + M_{kh} X_{ij}^2 + D_{kh} Y_{ij} + N, \tag{2.5.2}$$

in which D_{kh} is the crossline coefficient and Y_{ij} the perpendicular distance from the CMP to the effective profile line for trace ij. The modified traveltime function of Equation 2.5.2 is time dependent through the subscript h allowing to analyze residual move-out crossline dip and reflection times for each separate observed reflection. Thus Equation 2.5.2 is equipped to handle crooked line recordings adequately through the crossline coefficient. Note that the source and receiver statics are not time-dependent (through definition). The modified traveltime-equation of Equation 2.5.2 can be rewritten in matrix notation as,

$$\mathbf{t} = \mathbf{Am} + \mathbf{N}, \tag{2.5.3}$$

in which $\mathbf{t}$ is a vector with observed time-shifts, $\mathbf{A}$ is a model matrix containing the known coefficients of Equation 2.5.2 and $\mathbf{m}$ containing the unknown parameters: source and receiver statics, move-out coefficient and crossline coefficient. The noise is contained within $\mathbf{N}$. By minimizing the error/noise term,

$$\mathbf{N} = \mathbf{t} - \mathbf{Am}, \tag{2.5.4}$$

through using Bayes' theorem and probability theory, a least-squares solution for the unknown parameters is obtained. For detailed discussion on inversion the reader

is referred to Appendix B.

The major effort and time-consuming part for residual statics computations lies within the determination of the individual input times for the traces. Picking all input times from the prestack shot gathers by hand would be a cumbersome task. Time-picking is generally automated by cross-correlating individual traces $\hat{P}_i$ and $\hat{P}_j$ over time within a window (t_1, t_2) around the expected reflection energy,

$$\hat{\phi}_{ij}(\tau) = \frac{\int_{t1}^{t2} \hat{P}_i(t)\hat{P}_j(t+\tau)dt}{\sqrt{\int_{t1}^{t2} \hat{P}_i(t)\hat{P}_i(t+\tau)\mathrm{d}t \int_{t1}^{t2} \hat{P}_j(t)\hat{P}_j(t+\tau)\mathrm{d}t}}. \tag{2.5.5}$$

The maximum from the cross-correlation determines the differential time-shift between the two traces.

The result strongly depends on the quality of the input traces. To enhance signal to noise ratio's a pilot or model trace can be constructed. Within a CMP gather the initial pilot trace is the trace resulting from stacking the gather, which is correlated with the individual traces within the CMP gather to obtain the individual time-shifts. A drawback is that when static shifts are considerably, the quality of the pilot trace is degraded significantly (which is explained in more detail in Section 2.6) and hence the detection of time-shifts will become difficult. To overcome these problems, the pilot trace can be updated iteratively by constructing the pilot trace from the traces corrected by the found time-shifts to date.

Alternatively, Ronen and Claerbout (1985) introduced the concept of stack power optimization. The stack power of two traces $\hat{P}_i$ and $\hat{P}_j$ with differential time shift Δt reads,

$$E(\Delta t) = \int (\hat{P}_i(t-\Delta t) + \hat{P}_j(t))^2 \mathrm{d}t, \tag{2.5.6}$$

$$= \int (\hat{P}_i(t-\Delta t)^2 + \hat{P}_j(t)^2)\mathrm{d}t + 2\int \hat{P}_i(t-\Delta t)\hat{P}_j(t)\mathrm{d}t. \tag{2.5.7}$$

The first term in equation(2.5.7) is simply the stacked power of both traces individually, independent of the time shift, while the second term represents the cross-correlation between the two traces. By making two supertraces, one trace with concatenated traces from an individual source location and one trace from concatenated stacked CMP traces, excluding the source traces, the maximum power indicates the source-static. This procedure is repeated for all source and receiver locations.

A drawback of the correlation methods is that the resulting picks can be trapped within a local maximum of the computed correlation, when for example a large time window needs to be defined if large time-shifts are expected. This feature is known as cycle skipping. For that purpose several authors have introduced global optimization techniques such as Monte Carlo (Rothman, 1986), genetic algorithms (Dubose,

1993; Vasudeva et al., 1991) and simulated annealing (Sen and Stoffa, 1991).

In Chapter 4 an algorithm is discussed to find imaging operators that bares a close resemblance to the detection of residual statics by using a genetic algorithm; the reader can find an example on using genetic algorithms for residual static detection in Appendix B.

A generalization of the outlined residual statics method has also been used by Hatherly et al. (1994) to determine refraction statics from the refracted energy of linear move-out corrected data.

2.6 Error effect of statics

Residual statics cause errors with a statistical nature opposite to errors introduced by erroneous datuming assumptions. Within any stacking procedure, whether part of CMP processing or implicitly within imaging procedures, residual statics degrade the resulting image severely. For that matter consider the result of stacking N move-out corrected traces within a CMP-gather with fold N,

$$\hat{P}^S(\mathbf{x}^{cmp}, t) = \frac{1}{N}\sum_{i=1}^{N} \hat{P}(\mathbf{x}^{cmp}, \mathbf{x}_i^{off}, t), \tag{2.6.1}$$

in which $\mathbf{x}^{CMP}$ is the CMP-position under consideration and $\mathbf{x}^{off}$ are the offset values for the collected traces having similar a CMP location. When the traces contributing to the stacked trace are corrected for are perfectly corrected for move-out, spherical divergence and other amplitude effects, then, in the noise free situation, the result equals an average over N similar traces. If each trace would have an individual static shift, the stacking result can be written as,

$$\hat{P}^S(\mathbf{x}^{cmp}, t) = \frac{1}{N}\sum_{i=1}^{N} \hat{P}(\mathbf{x}^{cmp}, \mathbf{0}, t + \delta t_i), \tag{2.6.2}$$

with δt_i representing the individual static time shifts. If the traces within the CMP-gather contain uniform distributed residual statics with zero mean and maximum Δt, then the result reads after sorting the traces to their static shifts,

$$\hat{P}^S(\mathbf{x}^{cmp}, t) = \frac{1}{N}\sum_{i=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{P}(\mathbf{x}^{cmp}, \mathbf{0}, t + i\frac{2\Delta t}{N}). \tag{2.6.3}$$

which is a convolution in time with the boxcarr function $g(t)$, defined by

$$\hat{g}(t) = \hat{u}(t + \Delta t) - \hat{u}(t - \Delta t), \tag{2.6.4}$$

in which $u(t)$ resembles the Heaviside function. In the temporal frequency domain, Equation 2.6.3 can be written as the product in time of the spectrum of the zero-offset trace with a sinc-function $G(\omega)$, determined by the static distribution,

$$G(\omega) = \frac{\sin(\omega \Delta t)}{N\omega}, \tag{2.6.5}$$

which includes notches at every $\frac{1}{2\Delta t}$Hz interval. Thus for statics uniform distributed between $-8ms$ and $8ms$ the first notch occurs at 62.5 Hz, which is already within the dynamic range of the current seismic experiments. Figure 2.10a displays the stacked results for a 64-fold CMP-stack over a set of Ricker wavelets, with maximum static time-shift varying between 4 ms and 64 ms in steps of 4 ms (equal to the time sampling interval). Figure 2.10b shows the frequency attenuation in dB with respect to the static free result.

In general the static distribution will be normally distributed, for which the derivation will be far more cumbersome, but Figure 2.10c illustrates the effect for a 64-fold CMP gather with statics normally distributed with a mean of zero and a standard deviation between 4 ms and 64 ms in steps of 4ms. Clearly, the resulting spectrum of Figure 2.10d is severely degraded toward the noise level of -30 dB, which is quite common within land seismics at 20-30 Hz for a statics distribution with standard deviation of 8 ms.

2.7 Static corrections by focusing

In this section two alternatives are described to compute surface consistent static shifts from the data by combining and correlating shotrecords. Such a correlation is closely related to the so-called WRW-model, which is described in detail in Chapter 4. In this chapter we will use the focusing methodology derived from the WRW-model without detailed explanation and will refer to Chapter 4 when necessary.

The described methods are not based on a model and can therefor be considered data-driven approaches to static time-shift determination. In addition the methods could also provide initial focusing velocities.

The outcome of the proposed alternatives can later on be used to initialize parameters for the so called focusing operators as will be suggested in Chapter 5.

Focusing shot records

By using each shot-gather as a "focusing" operator on the full set of shot-gathers, an image is constructed, which images the shot-gather in itself. By taking the i-th shot-gather the image for this particular shot-gather is constructed by,

$$\mathbf{s}_i = \mathbf{p}_i^H(z_0, z_0)\mathbf{P}(z_0, z_0). \tag{2.7.6}$$

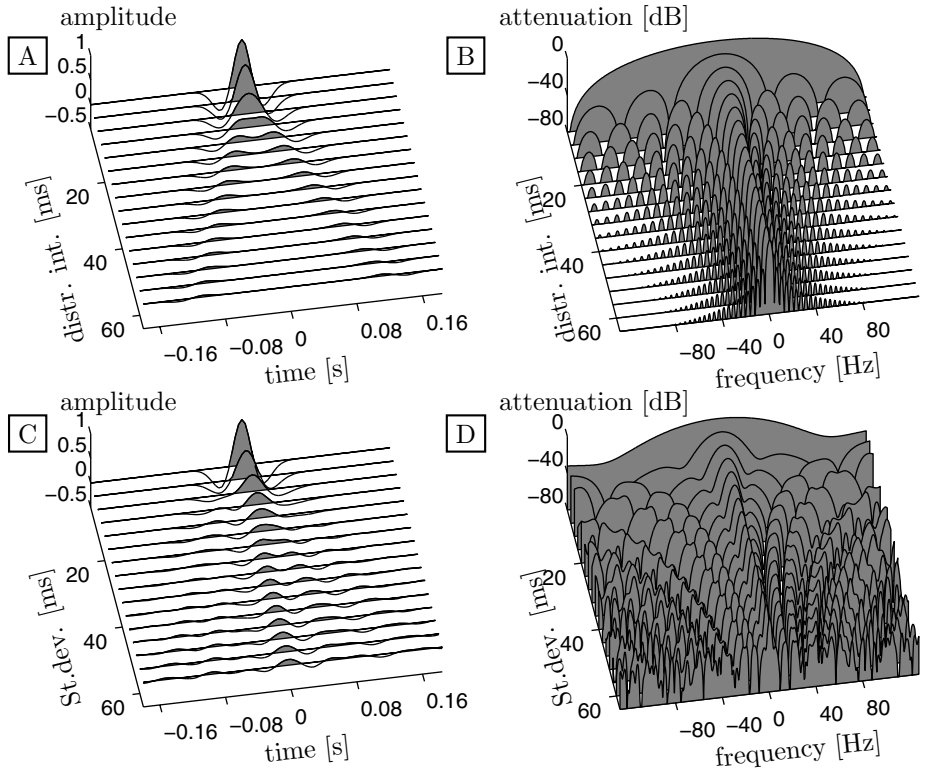

Figure 2.10: *Unresolved residual statics can severely degrade the stacked result in for instance CMP processing. In figure A) the stacked result is shown for a 64-fold CMP gather, for different amount of residual statics ranging from a uniform distributed statics between -4ms and +4 ms to -60ms and 60 ms. The stack degrades severely. The figure B) shows the attenuation curves of the stacked result with respect to a static free result. The statics introduce clearly recognizable notches. Figure C) shows the stacked result for the same 64-fold data, but here a more realistic normally distributed static was applied with a standard deviation ranging from 4ms to 64ms. Again a clear degradation is observed in both the stacked result as well as in the attenuation curves D).*

Here $\mathbf{p}_i$ resembles one specific shot-gather,

$$\mathbf{p}_i(z_0, z_0) = \mathbf{P}(z_0, z_0)\vec{I}_i, \tag{2.7.7a}$$

$$= \mathbf{W}^-(z_0, z_d)\mathbf{R}(z_d)\mathbf{w}_i^+(z_d, z_0)S(\omega), \tag{2.7.7b}$$

with $\vec{I}_i$ as defined in Equation 4.6.53. Then the focusing definition of Equation 2.7.6 can be written as,

$$\mathbf{s}_i = \left[\mathbf{w}_i^+(z_0,z_d)\right]^H \mathbf{R}^H(z_d) \left[\mathbf{W}^-(z_0,z_d)\right]^H \mathbf{W}^-(z_0,z_d)\mathbf{R}(z_d)\mathbf{W}^+(x_d,z_0)S^H(\omega)S(\omega), \tag{2.7.8a}$$

$$\mathbf{s}_i = \left[\mathbf{w}_i^-(z_0,z_d)\right]^* \mathbf{R}^H(z_d)\mathbf{R}(z_d)\mathbf{W}^+(z_d,z_0)S^H(\omega)S(\omega), \tag{2.7.8b}$$

in which it was assumed that sources and receivers are placed at the same surface locations such that,

$$\mathbf{W}^+(z_d,z_0) = \left[\mathbf{W}^-(z_0,z_d)\right]^T, \tag{2.7.9}$$

and the matched filter inverse propagators read according to equation 4.5.27 and 4.5.28,

$$\mathbf{F}^+(z_0,z_d) = \left[\mathbf{W}^-(z_d,z_0)\right]^*, \tag{2.7.10a}$$

$$\mathbf{F}^-(z_d,z_0) = \left[\mathbf{W}^+(z_0,z_d)\right]^*. \tag{2.7.10b}$$

Apart from the reflection operators, Equation 2.7.8b can be recognized as the expression for a second focusing step which will render a focused image. Although an image is created, this procedure is not directly applicable for image analysis. In principle the correlations should perfectly image all energy at $(t=0,\mathbf{x}_i)$. for all reflections, when cross-terms are handled properly within the correlation. All other $\mathbf{x}$ locations will render zero energy through deconstructive interference. On the other hand, as all energy collapses in the same point and no use is made of model parameters, either physically or intrinsically, no extra information is obtained.

The procedure will become more interesting, when opposite to the focusing principle analysis, the move-out times of the shot records are perturbed, as to defocus energy. In the extreme case of a perfect move-out correction with respect to the background medium excluding the static near surface effects, semi-plane waves are created with their source and receiver static timeshifts still included. Then correlating the corrected shot-gathers, by using Equation 2.7.6, is interpreted as stacking flat events, for which the receiver statics have canceled out, with a time-delay equal to the relative difference in the source statics. Note indeed that no static perturbations are observed at all within the correlation prior to stacking the traces.

If an optimal move-out would have been applied, maximum energy is rendered within the resulting correlation-gathers for all pseudo offsets (offset between source i and source j). In Figure 2.11 results are shown in the time domain of two shot-gathers correlated in time for different applied move-outs. One trace in the image-gather of Equation 2.7.6 is obtained by stacking the results. Figure 2.11A shows the result without applying move-out. Residual hyperbolic move-out is recognized, which will render maximum energy within the Fresnel-zone, which is absent here. The other figures show results after applying move-out with a move-out velocity of

1000,1500,1800,2500 and 3000 m/s. Move-out with 1800 m/s (Figure 2.11C) renders almost flat events, which will result in a maximum energy when stacked. In Figure 2.12 the results are shown when the complete focusing procedure is carried out with one shot-record on the whole set of shot-records with the same set of move-out velocities applied. When no move-out is applied, Figure 2.12A, the result shows the imaged energy at time equal to zero and the location of the shot-gather used for focusing. Maximum energy throughout the-gather is obtained for a move-out velocity of 1800 m/s, Figure 2.12C. The times where maximum energy occurs is equal to the relative difference in source statics.

Shot-gather cross correlation

Alternatively initial time-shifts could be derived through a spatial correlation between the shotgathers in the frequency domain,

$$\mathbf{s}_{ij} = \mathbf{p}_i(x)^* * \mathbf{p}_j(-x) \tag{2.7.11}$$

of which the result gives the relative spatial shift and time shift between shot i and shot j. Since the shot-gathers have common receiver statics, which cancel in the correlation, the result depends along the time-axis on the difference in source statics and along the spatial axis on structural effects. These structural effects are dominated mainly by the difference in source location for small pseudo offsets. By stacking the result along the spatial axis, these effects are averaged, resulting in a single trace with maximum energy at the relative time-difference of the source statics.

Both methods can be applied by considering common receiver-gathers to analyze the receiver statics as well. The cross-correlation method is sensitive to lateral changes in the subsurface. Therefore in this thesis the focusing method is preferred. Although no further research was done on this procedure, the method can be subjected to an optimization procedure in order to determine stacking velocities by power optimization.

By applying the proposed methods on all source-gathers a time-shift surface is found with relative static shifts for all sources. Consequently, a time-shift surface can be determined for all relative receiver statics as well, see Figure 2.13.

The obtained time-shift surface, $\mathbf{A}$, contains combinations of source-statics through,

$$A_{ij} = s_i - s_j, \tag{2.7.12}$$

in which s_i resembles the source static shift for source i. Then the static shifts are computed through,

$$\sum_{i=1}^{N} \frac{A_{ij}}{N} = s_j - \bar{s}, \tag{2.7.13}$$

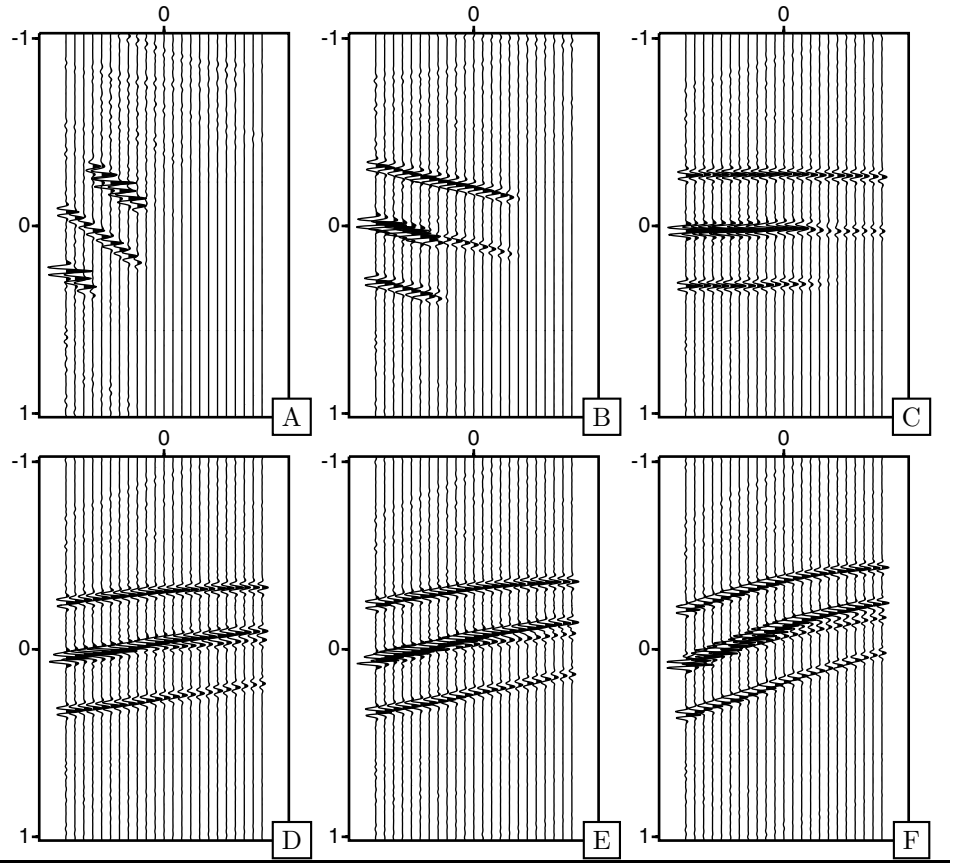

Figure 2.11: *The NMO velocity determines the alignment of correlated shot-gathers before stacking. Horizontal alignment determines the optimal NMO velocity. In (A), no NMO correction was applied, in (B) a too small correction was applied, meaning the use of a too high NMO velocity. In (C) a perfect correction was applied, resulting in perfect alignment of the correlated events. (D) to (F) are the results of applying NMO overcorrection, using too low velocities.*

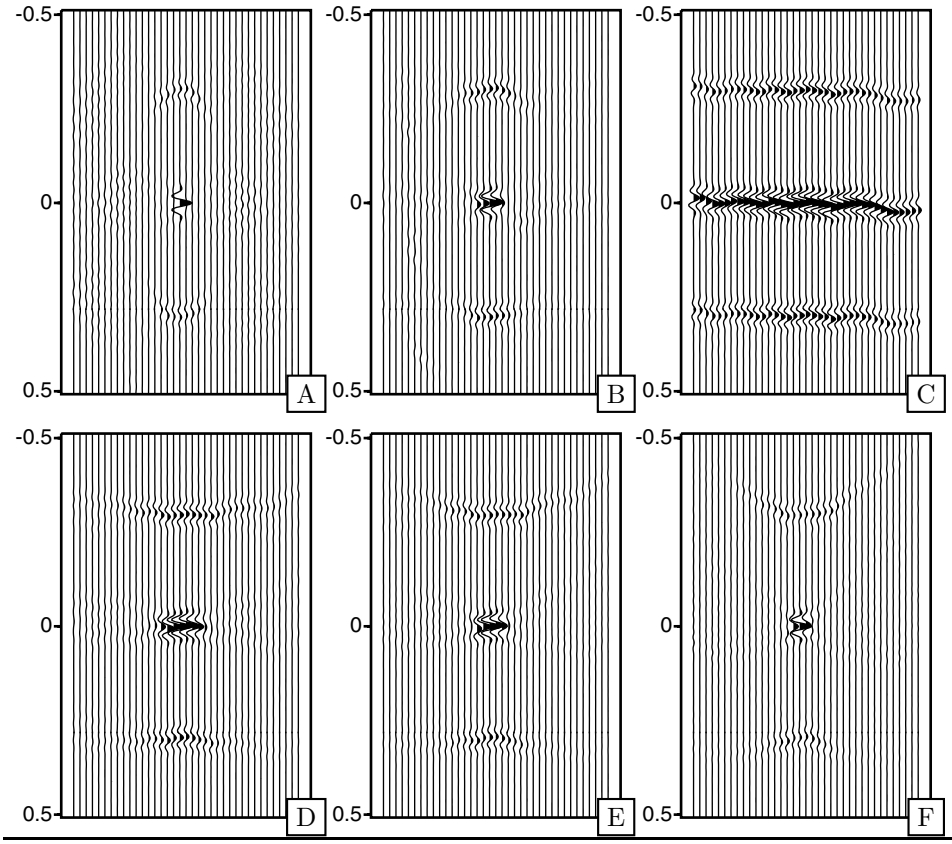

Figure 2.12: *Stack resulting by stacking individually correlated shot-gathers with one particular reference shot-gather. In figure A, no NMO correction was applied prior to correlation. Note that only the at the shot-location of the reference shot-gather a perfect correlation is recognized. In figure B a too small correction was applied, in figure C a perfect correction was applied. The final stacks clearly shows the near-surface behavior. Figure D to F are resulting after application of a too large NMO correction.*

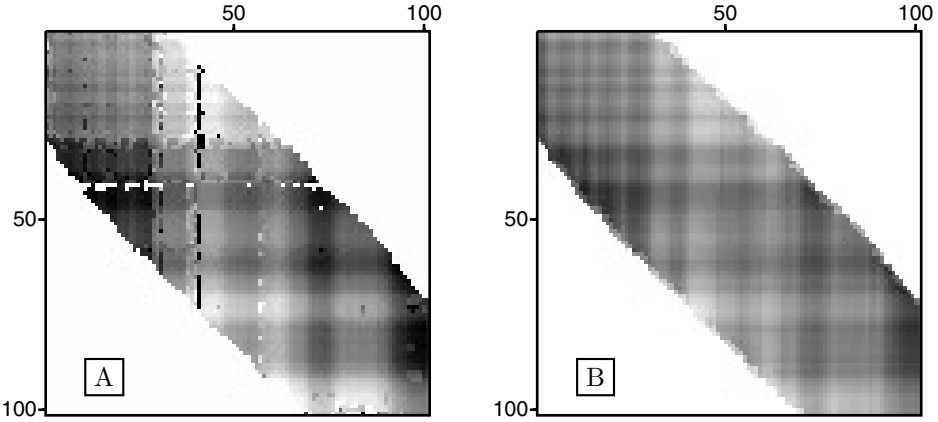

Figure 2.13: *The time-shift surface for cross-correlated shot-records (A), and for focused shot-gathers after applying a correct move-out correction (B).*

in which the average static, $\bar{s}$, is still unknown. However, assuming a priori $\bar{s} = 0$ renders the a priori statics to be $\hat{s}_j = \sum_{i=1}^{N} \frac{A_{ij}}{N}$. Then the average static distribution would read,

$$\sum_{j=1}^{N} \sum_{i=1}^{N} \frac{A_{ij}}{N^2} = 0, \tag{2.7.14}$$

which confirms the a priori estimate on the average static distribution. By applying the found statics distribution and the optimized NMO-velocity, the a priori zero-offset times can easily be windowed after which the NMO-velocity in combination with the zero-offset times will supply the a priori parameters for a global inversion algorithm.

Figure 2.14 displays the estimated statics for both methods. The initial static time-shifts were determined for a full range of shotgathers, modeled over the model of Figure 5.8C. Clearly both methods almost produce the same results. The focused shot-gather method is preferred since it provides a means to estimate an optimal stacking velocity and renders a cleaner and easier to determine traveltime surface.

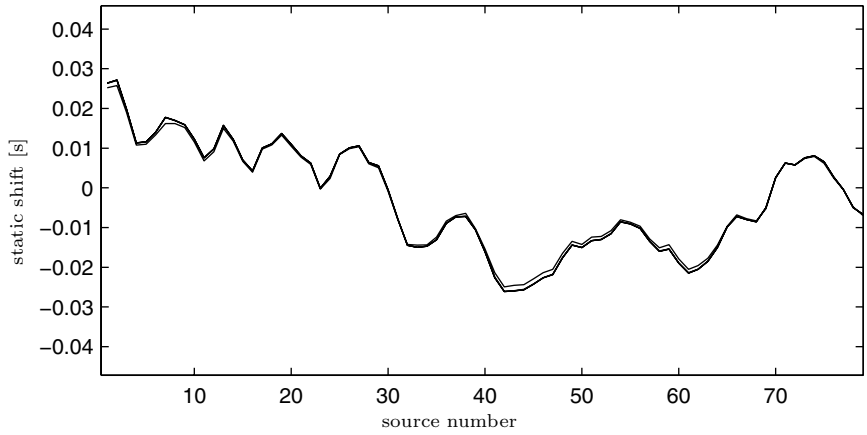

Figure 2.14: *Source statics derived from linear inversion of shot-gather cross-correlations and focusing shot-gathers for data simulated in the subsurface model of Figure 4.8C. Usage of the traveltime surfaces of both methods almost render the same results as both lines are almost exactly overlaying.*

3

From static corrections to a wavefield solution

3.1 Introduction

Even with perfectly applied static corrections, errors will still be introduced within the image result. This is due to the assumptions made. In this chapter, the statics solution is assessed with respect to imaging results and compared with results obtained through a correct wavefield handling. First, the impact of static corrections is analyzed on a simple near surface model, then more complex near surface models are discussed.

3.2 Simple near surface background model

The first example, Figure 3.1A, displays a low velocity layer (750 m/s) on top of a high velocity layer (2000 m/s). Within the high velocity layer a diffractor was modeled, for which Figure 3.1B displays one shot-record. For the near surface layer static corrections are computed that replace the near surface with the velocity of the subsurface. Then the resulting data was imaged with a phase-shift migration algorithm, leading to the image of Figure 3.1C. As an alternative, shots and receivers are redatumed to the boundary between subsurface and near surface and migrated from the datum as depicted by Figure 3.1D. For a more detailed description of the redatuming procedure the reader is referred to Chapter 7 of this thesis. Clearly the statics solution of Figure 3.1C tends to smear the point refractor. The underlying reason is that the current statics methods do not account for the change in raypaths when the near-surface velocity is replaced and for the fact that for larger offsets the rays through the near-surface are no longer vertical (see also Section 2.2). In Figure 3.2 the same experiment has been conducted on a dipping reflector of finite

length, for which the model is displayed in Figure 3.2A and one of the shot-records is shown in Figure 3.2B. Again, the recorded data was statically corrected followed by a phase-shift migration (Figure 3.2C) and compared with the migration result after redatuming, Figure 3.2D. The correct reflector position is indicated by the line in both figures. For this example, imaging after applying static corrections results in mispositioning of the reflector. Although these artefacts have a close resemblance to the use of erroneous migration velocities, it should be noted here that the errors introduced by statics are irrecoverable by adjusting the migration velocity.

How these effects, as demonstrated on the simple near surface cases, manifest themselves within more realistic processing is depicted by the following examples. Data was gathered over a model containing a stack of horizontal layers and forward extrapolated for both sources and receivers through two types of near surface models, a near surface model which is abruptly changing in the lateral direction and a near surface model which is more smoothly changing in the lateral direction. In this way the direct reflection response of the near surface layer is absent, but all propagation effects are included. The stack of horizontal layers was modeled by using a randomly changing density log and a velocity profile for which the stacking velocity from the datum level downward is linearly changing from 1500 m/s to 2000 m/s.

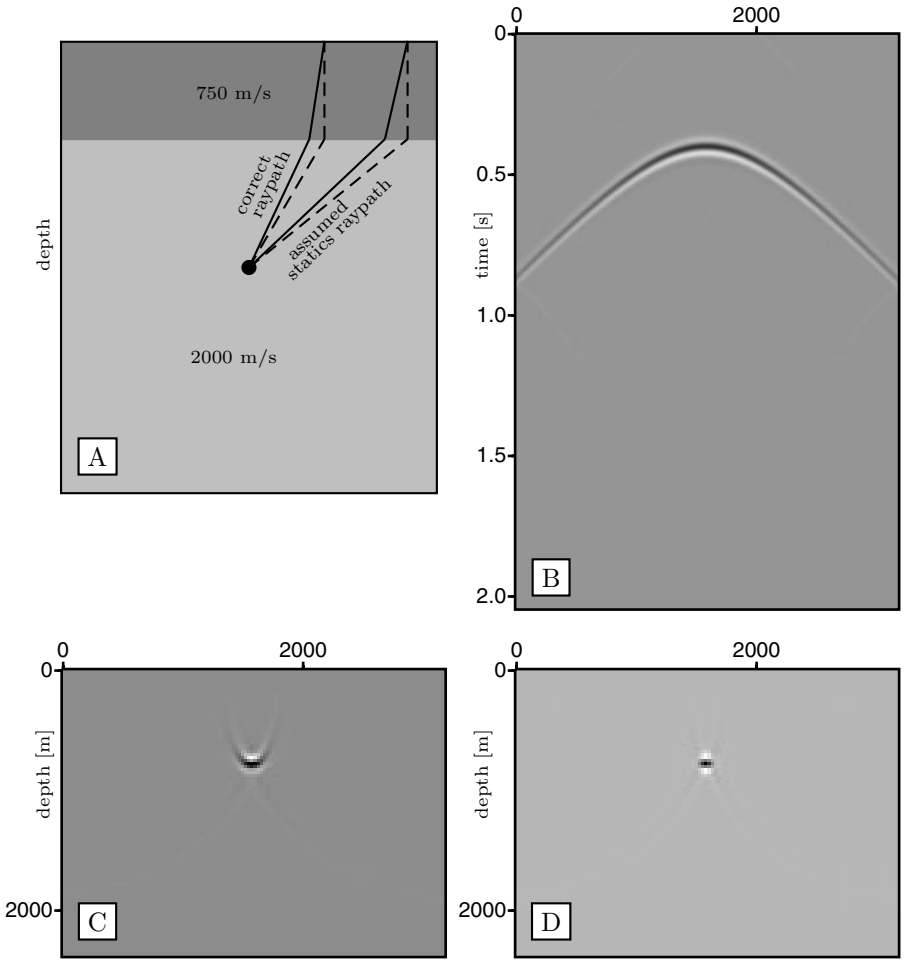

Figure 3.1: *The shot-record of (B) is recorded for a model containing a simple near surface layer, with a diffractor in the deeper subsurface, as displayed in (A). Static corrections to replace the near surface layer results into a smeared final image (C), while imaging after redatuming results in a perfectly imaged diffractor (D).*

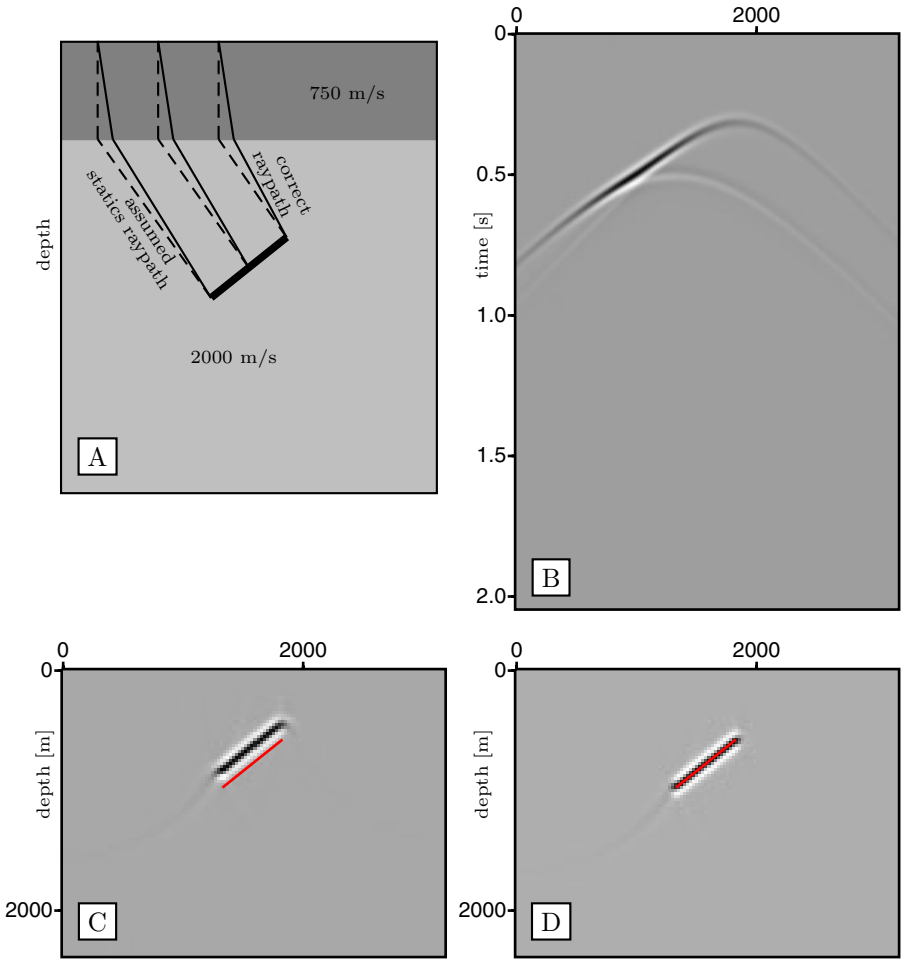

Figure 3.2: *The shot-record of (B) is recorded for a model containing a simple near surface layer, with a finite dipping reflector in the deeper subsurface, as displayed in (A). Static corrections to replace the near surface layer results into an incorrectly placed image with an erroneous dip(C), while imaging after redatuming results in a correctly imaged reflector (D).*

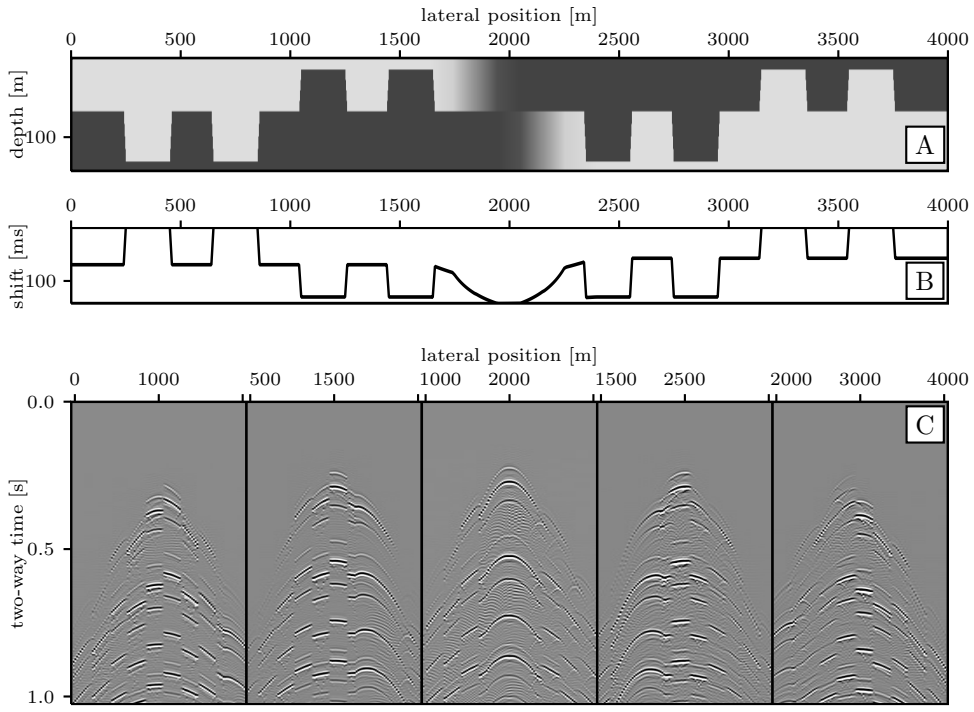

Figure 3.3: *The shot-records of (C) are the response of a stack of twelve horizontal layers, forward propagated through the near surface model of Figure (A). Dark color reflects a high velocity, light color a low velocity. The corresponding datum statics are depicted in Figure (B). The shots are placed at x =1000m, 1500m, 2000m, 2500m and 3000m respectively.*

3.3 Near surface model with abruptly changing anomalies

The first example to consider is a stacked layer configuration underneath a near surface which is abruptly changing laterally. Figure 3.3 depicts the modeled near surface layer; the near surface model is symmetrically built with on the left hand side a low velocity layer over a high velocity layer and on the right hand side a high velocity layer over a low velocity layer. Figure 3.3B displays the computed datum statics according to the near surface model. As a consequence of the generated model, the datum corrections are symmetric around the central location of the model. Thus we also expect the shot-records to behave more or less symmetrically. The shot records of Figure 3.3C, taken at shot locations 1000m, 1500m, 2000m, 2500m and 3000m respectively, show indeed similar effects in a mirrored sense. Next, the computed statics will be applied to the generated shot-records. The resulting shot-records of Figure 3.4A are input to a stacking velocity analysis and a subsequent NMO cor-

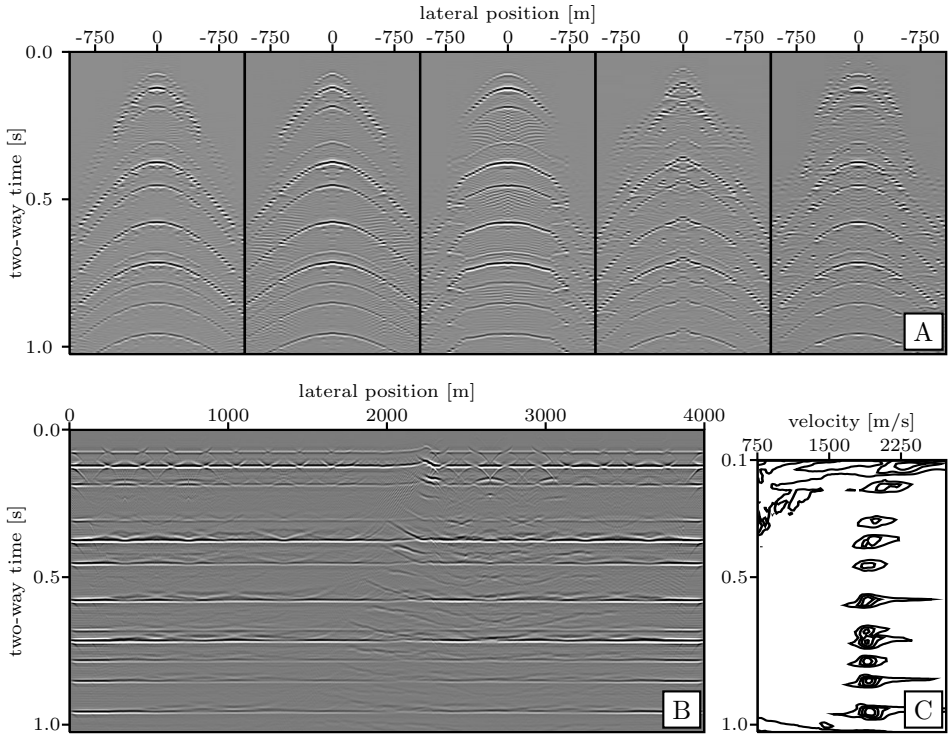

Figure 3.4: *(A) shows 5 CMP-gathers after application of datum corrections to the shot-gathers of Figure 3.3. (B) shows the resulting stack after applying NMO corrections. (C) shows the average semblance after static corrections.*

rection procedure. Figure 3.4B shows the stacked result after applying both the datum statics as well as the NMO-corrections. The results show artefacts which are due to imperfect datum statics. Residual statics might improve a little. However, note that the semblance for velocity analysis of Figure 3.4C, after applying datum statics shows very high velocities in the top of the model, which contradicts the true velocity profile. Even for the deepest events the velocity was estimated too high. The stack obtained after applying conventional statics will be compared with the stacked resulted when inverse wavefield extrapolation is applied to the modeled shot-records for both receivers as well as sources, taking into account the near surface anomalies. In fact, a full redatuming has taken place of sources and receivers to a level underneath the near surface. The operators used for wavefield extrapolation were modeled as single arrival time operators, not taking into account triplications and amplitude effects. They are related to a datum just below the near surface.

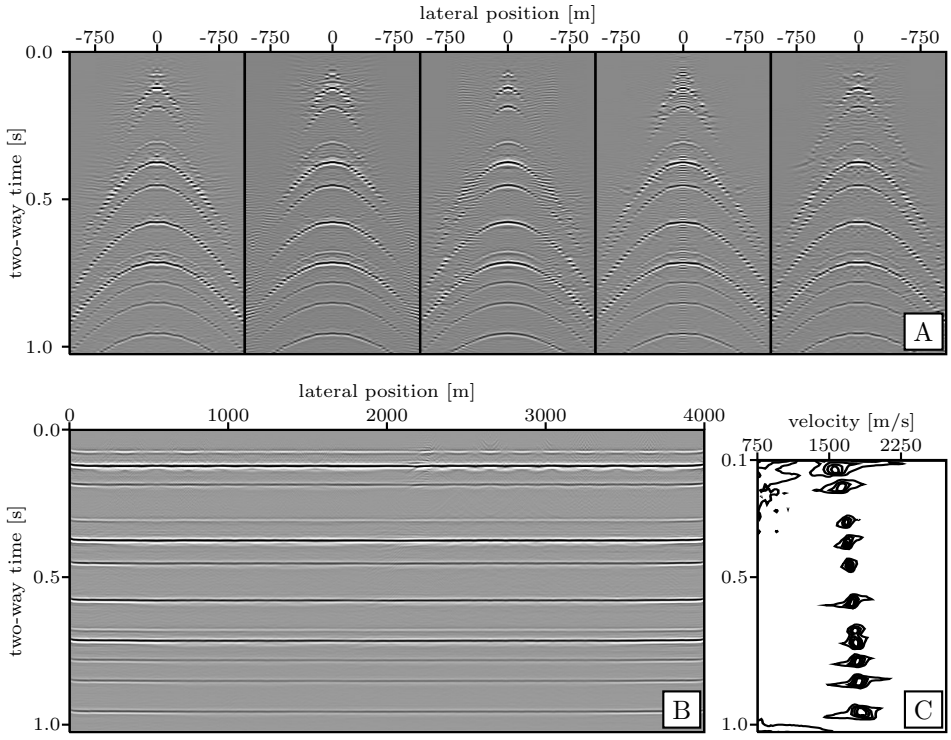

Figure 3.5: *(A) shows 5 CMP-gathers after inverse extrapolation of the shot-gathers of Figure 3.3. (B) shows the resulting stack after applying NMO corrections. (C) shows the average semblance after inverse extrapolation.*

Figure 3.5A shows the same set of CMP gathers as the gathers of Figure 3.4B, but this time collected after inverse extrapolation with the single arrival-time operators. The resulting stack of Figure 3.5B hardly shows any artefacts and the semblance plot of Figure 3.5C correctly reflects the true stacking velocity profile.

3.4 Near surface model with smoothly changing anomalies

The second example is again an analysis of the same stacked layer configuration. In this case the near surface layer has a more smooth character as depicted in Figure 3.6A. This sinusoidal behavior will give rise to triplications, i.e. multi-valued arrivals, in the data. Again Figure 3.6B shows the computed datum statics and Figure 3.6C 5 shot-records taken at identical lateral positions as the records of Figure 3.3C. Note that the shot-records no longer show symmetrical features. Due to

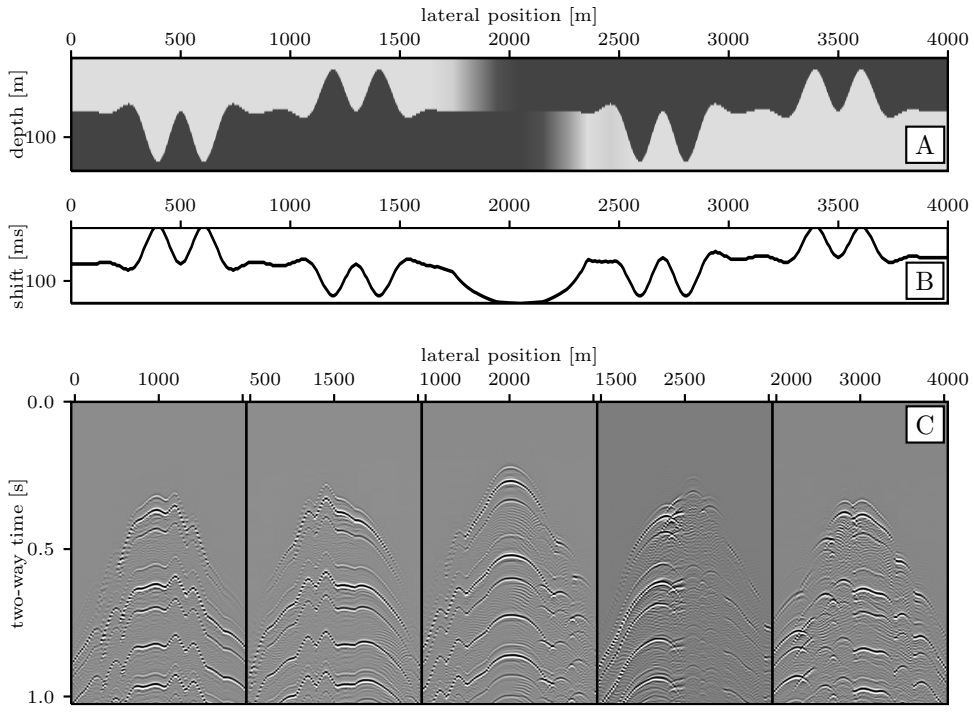

Figure 3.6: *The shot-records of (C) are the response of a stack of horizontal layers, forward propagated through the near surface model of (A). Dark color reflects a high velocity, light color a low velocity. The corresponding datum statics are depicted in (B). The shot-gathers are placed symmetrically around x =1000m, 1500m, 2000m, 2500m and 3000m respectively.*

the high velocity layer over a low velocity layer, triplications occur which are far more persistent than for the area of the model where a low velocity layer overlies a high velocity layer. As for the CMP gathers after datum correction in Figure 3.7A, especially the corrections for the CMP-gathers at the left part of the model are far from correct. Events are hardly recognized, which will not be improved by residual corrections due to the introduction of among others, cycle skipping. One of the assumptions of conventional static methods is that rays through the near surface refracted toward the normal of the interface between near surface and deeper layers and therefore the rays through the near surface are assumed to be independent of offset and reflection depth. This holds to some extend for a low velocity surface, but breaks down completely for a high velocity near surface layer (for instance a permafrost layer). The resulting stack of Figure 3.7B is severely degraded and the velocity semblance of Figure 3.7C again shows high stacking velocities in the top of the model. The redatumed shot-gathers of Figure 3.8A show good hyperbolic

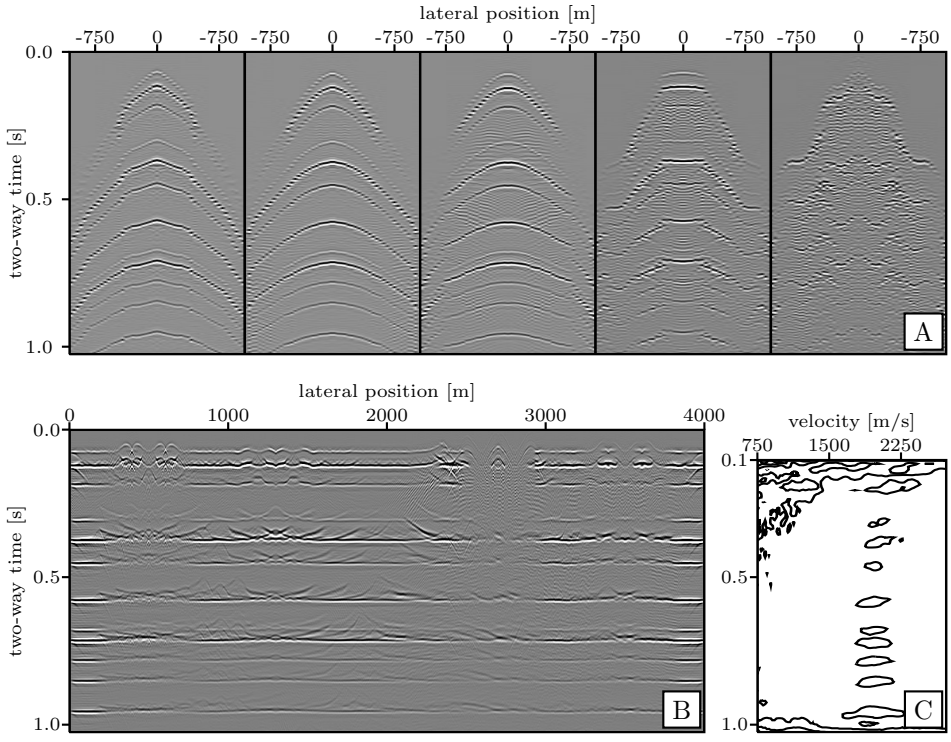

Figure 3.7: *(A) shows 5 CMP-gathers after application of datum corrections to the shot-gathers of Figure 3.6. (B) shows the resulting stack after applying NMO corrections. (C) shows the average semblance after static corrections.*

move-out and correlated energy is still fairly well recognized within the part of the model where a high velocity layer is overlaying a low velocity layer (at the right side of the model). Also, the stack of Figure 3.8B shows far less artefacts and again the semblance plot of Figure 3.8C resembles well the correct velocity semblance.

The redatuming operators where modeled with a first arrival traveltime algorithm, while, due to triplications, the high energy arrivals are in general not the first arrivals at the right-hand side part of the model. This explains the degraded results at the right hand side of the stack. In the 5th CMP-gather this effect is well visible. The reconstructed reflection events are not nicely aligned along a hyperbola as would be expected. This is not a fundamental issue, but is caused by the limited capabilities of the eikonal solver. The eikonal solver, used to model the operators, did not correctly compute the arrival times of the most energetic event when a rapidly varying

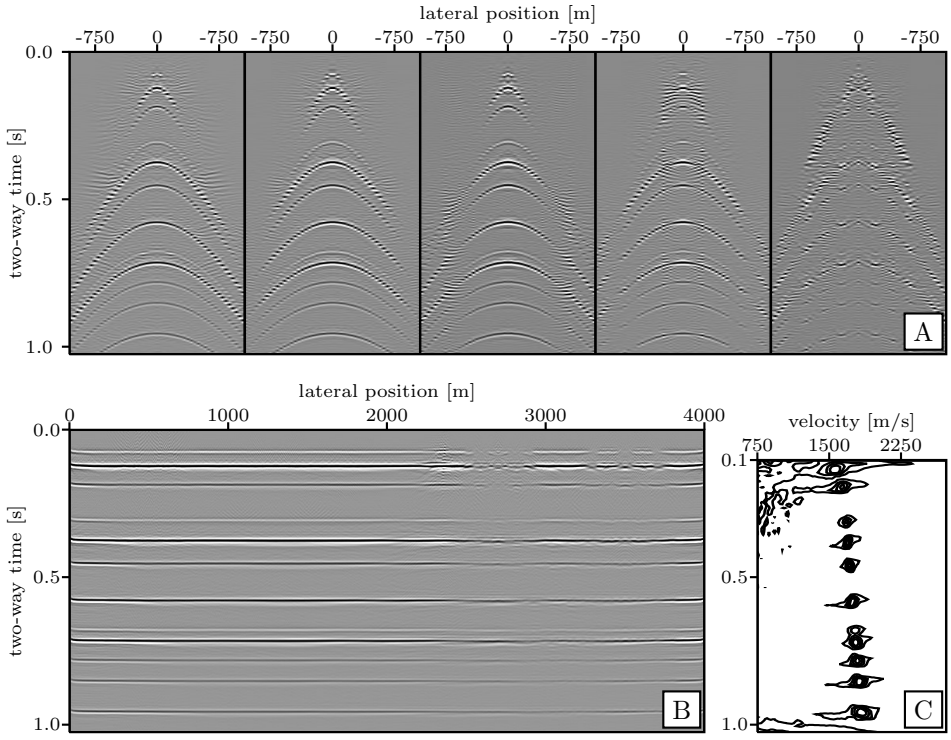

Figure 3.8: *(A) shows 5 CMP-gathers after inverse extrapolation of the shot-gathers of Figure 3.6. (B) shows the resulting stack after applying NMO corrections. (C) shows the average semblance after inverse extrapolation.*

high velocity layer is overlaying a low velocity layer.

3.5 Conclusions

From the experiments it is concluded that for simple cases, especially in areas with low velocity material overlying high velocity material, statics tend to improve data quality. However, in complex situations results deteriorate and reduce the quality of the final image. Moreover, results from velocity analysis can be far from correct, which will severely affect migration and time to depth conversion at later stages of the processing.

From the experiments it is also concluded that operators need to be estimated that represent the propagation characteristics of the near surface. These operators are used to remove the near surface effects from the data by redatuming. In the next chapters we will see how such operators are obtained from the data without explicit knowledge of the underlying velocity model.

4

Focusing of seismic data

4.1 Introduction

In this chapter the general theory for acoustic wavefield extrapolation and scattering is derived. This will provide the base for redatuming the wavefield through the near surface. Using a two-way wavefield propagation approach leads to complex wavefield extrapolators. It will be convenient to describe the wavefield in terms of one-way propagation, which simplifies the expressions for extrapolation and scattering significantly. Moreover, assumptions on underlying velocity models are less stringent for one-way wavefield extrapolation.

After derivation of the integral representations for one-way wavefield extrapolation and scattering, the continuous expressions are rewritten in a matrix notation which elegantly describes the processes to be investigated. This so called WRW model provides a clear insight in the redatuming process. Furthermore, the properties of the WRW formulation will provide us with tools that circumvent the need for an explicit formulation of a depth model to determine the propagators needed for the near-surface redatuming.

4.2 Two-way integral representations

To find integral expressions for forward and inverse extrapolation of wavefields, use will be made of the reciprocity theorem (Fokkema and van den Berg, 1993), as described in Appendix A. The reciprocity theorem is used to relate two field states. The first of the two states is the actual field state, $\{p(\mathbf{x}, \mathbf{x}^s, \omega), v_k(\mathbf{x}, \mathbf{x}^s, \omega)\}$ with material state $\{\rho, \kappa\}(\mathbf{x})$ and source state $\{q(\mathbf{x}, \omega), \mathbf{0}\}(\mathbf{x}, \omega)$, denoted by state A. The second

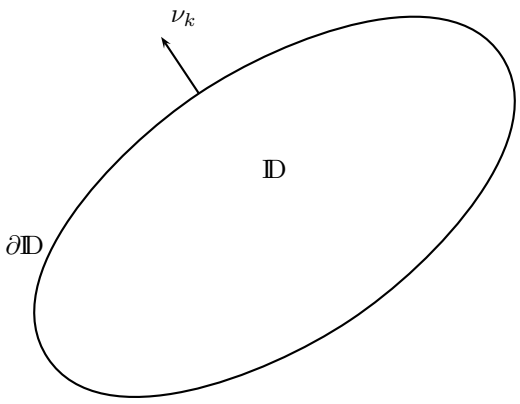

Figure 4.1: *Definition of the volume used for evaluating the wave-equation and reciprocity relations, in which $\partial\mathbb{D}$ denotes the boundary of volume $\mathbb{D}$ for which ν_k is the outward pointing normal vector with unitary length.*

state that will be used within the reciprocity theorem is the Green's state for volume injection, $\{G^q$ and $\Gamma_k^q\}$ with field state $\{p^q, v_k^q\}(\mathbf{x}, \omega)$, material state $\{\rho^q, \kappa^q\}(\mathbf{x})$ and source state $\{\delta(\mathbf{x} - \mathbf{x}^R), 0\}$, denoted as state B. From the reciprocity theorem the Kirchhoff integral can be derived which describes extrapolation of wavefields. Under certain conditions the Kirchhoff integral simplifies to the Rayleigh integrals, which form a convenient base for seismic imaging.

■ 4.2.1 The wave equation and reciprocity

As described in Appendix A, the volume injection source Green's states G^q and Γ_k^q are solutions for the pressure and particle velocity of the wave equation within a domain $\mathbb{D}$. The definition of $\mathbb{D}$ is depicted in Figure 4.1. The acoustic wave equation can be expressed in terms of two coupled equations,

$$\partial_k p(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\rho(\mathbf{x}) v_k(\mathbf{x}, \mathbf{x}^s, \omega) = f_k(\mathbf{x}, \mathbf{x}_s, \omega), \tag{4.2.1a}$$

$$\partial_k v_k(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\kappa(\mathbf{x}) p(\mathbf{x}, \mathbf{x}^s, \omega) = q(\mathbf{x}, \mathbf{x}^s, \omega). \tag{4.2.1b}$$

Then the Green's states G^q and Γ_k^q are solutions for these coupled equations for the scalar impulse response of the source state given by,

$$f_k(\mathbf{x}, \mathbf{x}^s, \omega) = 0, \tag{4.2.2}$$

$$q(\mathbf{x}, \mathbf{x}^s, \omega) = q\delta(\mathbf{x} - \mathbf{x}^s), \tag{4.2.3}$$

and the material state defined as $\{\rho, \kappa\}(\mathbf{x}) = \{\rho^q, \kappa^q\}(\mathbf{x})$. The Green's fields are then defined as,

$$\{p^q, v_k^q\}(\mathbf{x}, \mathbf{x}^s, \omega) = q\{G^q, -\Gamma_k^q\}(\mathbf{x}, \mathbf{x}^s, \omega). \tag{4.2.4}$$

	State A	State B
Field State	$\{p, v_k\}(\mathbf{x}, \omega)$	$\{G^q, \Gamma^q_k\}(\mathbf{x}, \omega)$
Material State	$\{\rho, \kappa\}(\mathbf{x})$	$\{\rho^q, \kappa^q\}(\mathbf{x})$
Source State	$\{q(\mathbf{x}^S), \mathbf{0}\}(\omega)$	$\{\delta(\mathbf{x}-\mathbf{x}^R), \mathbf{0}\}$

Table 4.1: *The two states for defining the reciprocity relations.*

The two states are depicted in Table 4.1. If the source functions of Equation 4.2.2 and 4.2.3 are substituted in the wave equations of Equations 4.2.1a and 4.2.1b it can be found that the following holds for the volume injection pressure Green's function G^q,

$$\rho^q(\mathbf{x})\partial_k(\frac{1}{\rho^q(\mathbf{x})}\partial_k G^q(\mathbf{x}, \mathbf{x}^R, \omega)) + \frac{\omega^2}{c^2(\mathbf{x})}G^q(\mathbf{x}, \mathbf{x}^R, \omega) = -j\omega\rho(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}^R), \quad (4.2.5)$$

with the velocity defined by, $c(\mathbf{x})^2 = \frac{1}{\rho^q(\mathbf{x})\kappa^q(\mathbf{x})}$. Substituting Equation 4.2.4 and the source functions into Equation 4.2.1a leads to,

$$\Gamma^q_k(\mathbf{x}, \omega) = \frac{1}{j\omega\rho^q(\mathbf{x})}\partial_k G^q(\mathbf{x}, \omega). \quad (4.2.6)$$

From here on we will leave out the reference to the frequency dependency in the argument of the field quantities, for reasons of brevity.

■ 4.2.2 The Kirchhoff integral

Substituting the two states as described in the previous section and depicted in Table 4.1 into the two-way reciprocity theorem, as derived in Appendix A, results into,

$$\begin{aligned} \int_{\mathbf{x}\in\partial\mathbb{D}} \left[G^q(\mathbf{x}, \mathbf{x}^R)v_k(\mathbf{x}) - p(\mathbf{x})\Gamma^q_k(\mathbf{x}, \mathbf{x}^R)\right] \nu_k \mathrm{d}S = \\ \int_{\mathbf{x}\in\mathbb{D}} j\omega \left[(\rho-\rho^q)\,\Gamma^q_k(\mathbf{x}, \mathbf{x}^R)v_k(\mathbf{x}) - (\kappa-\kappa^q)\,G^q(\mathbf{x}, \mathbf{x}^R)p(\mathbf{x})\right] \mathrm{d}V \\ + \int_{\mathbf{x}\in\mathbb{D}} q^S(\mathbf{x})G^q(\mathbf{x}, \mathbf{x}^R) + p(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}^R)\mathrm{d}V. \end{aligned} \quad (4.2.7)$$

With the aid of the equation of motion, Equation 4.2.4, for the two different states and using $\Delta\rho = (\rho - \rho^q)$ and $\Delta\kappa = (\kappa - \kappa^q)$ we arrive at the Kirchhoff-Helmholtz

integral,

$$p(\mathbf{x}^R,\mathbf{x}^S) = -\int_{\mathbf{x}\in\partial\mathbb{D}}\left[\frac{G^q(\mathbf{x},\mathbf{x}^R)\partial_k p(\mathbf{x})}{j\omega\rho} - \frac{p(\mathbf{x})\partial_k G^q(\mathbf{x},\mathbf{x}^R)}{j\omega\rho^q}\right]\nu_k \mathrm{d}S$$
$$+\frac{1}{j\omega}\int_{\mathbf{x}\in\mathbb{D}}\left[\frac{\Delta\rho}{\rho\rho^q}\partial_k G^q(\mathbf{x},\mathbf{x}^R)\partial_k p(\mathbf{x}) - \Delta\kappa G^q(\mathbf{x},\mathbf{x}^R)p(\mathbf{x})\right]\mathrm{d}V \qquad (4.2.8)$$
$$+\int_{\mathbf{x}\in\mathbb{D}} q^S(\mathbf{x})G^q(\mathbf{x},\mathbf{x}^R)\mathrm{d}V,$$

in which $\mathbb{D}$ denotes the domain under consideration with boundary $\partial\mathbb{D}$ and its outward pointing normal vector ν_k, as depicted by Figure 4.1. This equation states that the wavefield measured at $p(\mathbf{x}^R)$ consist of contributions of sources outside the domain, covered by the surface integral, contributions of scattering objects within the domain, covered by the first volume integral, and contributions of primary sources, covered by the second volume integral which resembles the direct wavefield.

More interesting in this thesis are the results, when the Green's medium is chosen to be the actual medium and sources of the actual wavefield are chosen to lay outside the domain under consideration. Under these choices, the volume integrals vanish, resulting into,

$$p(\mathbf{x}^R,\mathbf{x}^S) = -\int_{\mathbf{x}\in\partial\mathbb{D}}\left[\frac{G^q(\mathbf{x},\mathbf{x}^R)\partial_k p(\mathbf{x})}{j\omega\rho} - \frac{p(\mathbf{x})\partial_k G^q(\mathbf{x},\mathbf{x}^R)}{j\omega\rho}\right]\nu_k \mathrm{d}S, \qquad (4.2.9)$$

which describes forward propagation of the wavefield measured along the surface of the domain under consideration, as depicted in Figure 4.2A. This result will be used for forward extrapolation. Instead of considering the forward propagating causal Green's state, we can consider the anti-causal backward propagating Green's state $\{G^{q,*},\Gamma_k^{q,*}\}$ as well, since this is also a solution of the wave-equation of Equation 4.2.5. This backward propagating state results into,

$$p(\mathbf{x}^R,\mathbf{x}^S) = -\int_{\mathbf{x}\in\partial\mathbb{D}}\left[\frac{G^{q,*}(\mathbf{x},\mathbf{x}^R)\partial_k p(\mathbf{x})}{j\omega\rho} - \frac{p(\mathbf{x})\partial_k G^{q,*}(\mathbf{x},\mathbf{x}^R)}{j\omega\rho}\right]\nu_k \mathrm{d}S, \qquad (4.2.10)$$

which will be used for inverse extrapolation and is depicted by Figure 4.2B.

If the circular boundaries of Figure 4.2A and Figure 4.2B are extended to infinity then, under Sommerfeld's radiation condition, the contribution to the closed surface integral of the extended part vanishes for the forward propagation of Equation 4.2.9, resulting in analogy with Figure 4.2B into,

$$p(\mathbf{x}^R,\mathbf{x}^s) = -\int_{z=z_0}\left[\frac{G^q(\mathbf{x},\mathbf{x}^R)\partial_k p(\mathbf{x})}{j\omega\rho} - \frac{p(\mathbf{x})\partial_k G^q(\mathbf{x},\mathbf{x}^R)}{j\omega\rho}\right]\nu_k \mathrm{d}S. \qquad (4.2.11)$$

However, in case of backward propagation of Equation 4.2.10, Sommerfeld's radiation condition is not satisfied. As shown by Wapenaar and Berkhout (1989), however,

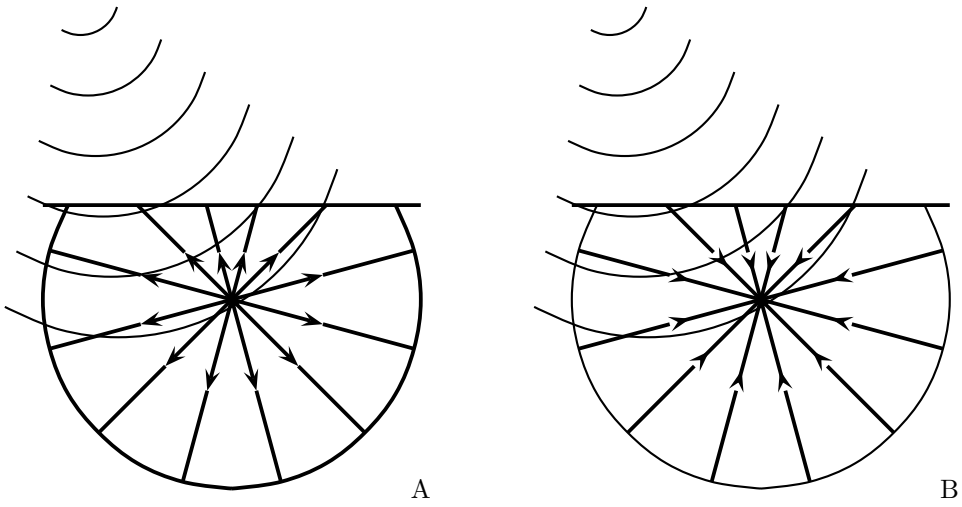

Figure 4.2: *The reciprocity theory leads to the notion that forward extrapolation is carried out through convolution with the causal Green's state (A) along the boundary of the domain under consideration, whereas for inverse extrapolation the anti-causal Green's state should be used (B). Sommerfeld's radiation condition shows that the contribution of the circular area will vanish if this part is extended to infinity for forward extrapolation. This is, however, not the case for inverse extrapolation but the evanescent part can be neglected.*

under the evanescent wave assumptions and a certain cylindrical choice of the domain under consideration, the integral can still be approximated by,

$$p(\mathbf{x}^R, \mathbf{x}^s) \approx -\int_{z=z_0} \left[\frac{G^{q,*}(\mathbf{x}, \mathbf{x}^R)\partial_k p(\mathbf{x})}{j\omega\rho} - \frac{p(\mathbf{x})\partial_k G^{q,*}(\mathbf{x}, \mathbf{x}^R)}{j\omega\rho} \right] \nu_k \mathrm{d}S. \qquad (4.2.12)$$

■ 4.2.3 The Rayleigh integrals

The equations for forward and inverse propagation require the knowledge of the pressure and the normal component of the particle velocity along the boundary. When the measurements are known along a flat boundary it is possible to eliminate either the pressure contribution or the normal component of the particle velocity by choosing the appropriate boundary conditions for the Green's states in Equation 4.2.11 and Equation 4.2.12:

- By using the Dirichlet boundary condition for the Green's state,

$$\lim_{\epsilon \downarrow 0} G^q(\mathbf{x} + \epsilon\nu, t) = 0, \qquad \mathbf{x} \in \partial\mathbb{D}, \qquad (4.2.13)$$

on $\partial\mathbb{D}$, $G^q(\mathbf{x} \in \partial\mathbb{D}) = 0$, and vanishes. Then the boundary acts as a perfect reflector with a negative reflection coefficient. This situation can be accomplished by constructing the Green's function through taking the free space Green's function and a mirrored free-space Green's function opposite in sign,

$$G^q(\mathbf{x}, \mathbf{x}^R) = G^{q,f}(\mathbf{x}, \mathbf{x}^R) - G^{q,f}(\mathbf{x}, \mathbf{x}^{R'}), \tag{4.2.14}$$

$$\nu_k \partial_k G^q(\mathbf{x}, \mathbf{x}^R) = \nu_k \partial_k G^{q,f}(\mathbf{x}, \mathbf{x}^R) - \nu_k \partial_k G^{q,f}(\mathbf{x}, \mathbf{x}^{R'}). \tag{4.2.15}$$

By choosing ν_k to be the outward pointing normal vector of the flat boundary under consideration this leads to,

$$\nu_k \partial_k G^q(\mathbf{x}, \mathbf{x}^R) = 2\nu_k \partial_k G^{q,f}(\mathbf{x}, \mathbf{x}^R). \tag{4.2.16}$$

If this expression is inserted into the reduced Kirchhoff integral of Equation 4.2.9 we arrive at,

$$p(\mathbf{x}^R) = 2 \int_{\mathbf{x} \in \partial\mathbb{D}} -\Gamma_k^{q,f}(\mathbf{x}, \mathbf{x}^R) p(\mathbf{x}) \nu_k \mathrm{d}A, \tag{4.2.17}$$

in which we used the definition of $\Gamma_k^{q,f}(\mathbf{x}, \mathbf{x}^R)$ as given by Equation 4.2.6. This equation is also known as the Rayleigh integral of the second kind.

- The Neumann boundary condition,

$$\lim_{\epsilon \downarrow 0} \nu_k \frac{1}{j\omega\rho} \partial_k G^q(\mathbf{x} + \epsilon\nu, t) = 0, \qquad \mathbf{x} \in \partial\mathbb{D}, \tag{4.2.18}$$

with ν the unit vector along the normal of $\partial\mathbb{D}$ pointing away from $\mathbb{D}$, dictates the normal component of $\Gamma_k(\mathbf{x} \in \partial\mathbb{D}, \mathbf{x}^R)$ to be zero. This can be accomplished by interpreting the boundary as a perfect reflector with a positive reflection coefficient. This can be achieved by mirroring again the Green's function but this time with positive sign,

$$G^q(\mathbf{x}, \mathbf{x}^R) = G^{q,f}(\mathbf{x}, \mathbf{x}^R) + G^{q,f}(\mathbf{x}, \mathbf{x}^{R'}). \tag{4.2.19}$$

Due to the choice of the normal vector, the Neumann condition is satisfied, leading to,

$$p(\mathbf{x}^R) = 2 \int_{\mathbf{x} \in \partial\mathbb{D}} \frac{G^{q,f}(\mathbf{x}, \mathbf{x}^R) \partial_k p(\mathbf{x})}{j\omega\rho} \nu_k \mathrm{d}A. \tag{4.2.20}$$

With $v_k(\mathbf{x}) = \frac{-1}{j\omega\rho} \partial_k p(\mathbf{x})$, this leads to

$$p(\mathbf{x}^R) = 2 \int_{\mathbf{x} \in \partial\mathbb{D}} G^{q,f}(\mathbf{x}, \mathbf{x}^R) v_k(\mathbf{x}) \nu_k \mathrm{d}A. \tag{4.2.21}$$

This equation is also known as the Rayleigh integral of the first kind.

	State A	State B
Field State	$\{G^{+,+}, G^{-,-}\}(\mathbf{x}, \mathbf{x}^R)$	$\{P^+, P^-\}(\mathbf{x}, \mathbf{x}^S)$
Material State	$-j\omega\mathbf{\Lambda}(\mathbf{x})$	$\mathbf{B}(\mathbf{x})$
Source State	$\delta(\mathbf{x} - \mathbf{x}^R)$	$\{S^+, S^-\}\delta(\mathbf{x} - \mathbf{x}^S)$

Table 4.2: *The two states used for defining the reciprocity relations for one-way wavefields.*

Two more Rayleigh integrals can be derived, expressing the particle velocity in terms of a surface integral over a known wavefield distribution, when we would have started with volume force injection sources for the Green's state (Appendix A).

Thus, the Kirchhoff integral can be simplified, leading to an extrapolation formulation depending on the velocity component or the pressure component of the field state along the boundary only. Note, however, that due to the choice of perfectly mirrored free-space Green's functions, these simplifications are only valid along a perfectly flat interface. If this is not the case, the full Kirchhoff integral as formulated in Appendix A should in principle be evaluated. Furthermore, if the Rayleigh integrals are used to forward and inverse extrapolate wavefields, the full two-way free surface response of the Green's state should be computed. Especially for complex models this can be a cumbersome and difficult task.

4.3 One-way integral representations

As imaging strategies, including redatuming, are based on back-propagation of primary one-way wavefields to their originating scattering contrast, it can be useful to use one-way reciprocity theorems. Moreover, by using two-way representations, the full Green's functions including, for instance, multiple energy have to be computed. Slight errors in the background model can introduce severe errors as shown by Wapenaar and Grimbergen (1996). As explained in Appendix A, one-way formulations can be derived by considering one particular direction of propagation only. As our measurements and processing algorithms are aimed at imaging in depth, our preferred axes is chosen to be the z-axis. Evaluations at the boundary of a volume under consideration will therefore only be made along ν_3, the z-component of the outward pointing unit normal vector.

■ 4.3.1 One-way reciprocity

Similar to the two-way integral representation, one-way wavefield representations can be derived through the one-way reciprocity theorem of the convolution type (Appendix A). Again the interaction of two states will be considered. For state A, the one-way Green's state is chosen to have propagation characteristics equal to the

actual medium but to be scatter free (no coupling). By virtue of this choice the Green's wavefield $\mathbf{G}$, as derived in Appendix A, reduces to a diagonal operator,

$$\mathbf{G}(\mathbf{x},\mathbf{x}^R) = \begin{pmatrix} G^{+,+}(\mathbf{x},\mathbf{x}^s) & G^{+,-}(\mathbf{x},\mathbf{x}^R) \\ G^{-,+}(\mathbf{x},\mathbf{x}^R) & G^{-,-}(\mathbf{x},\mathbf{x}^R) \end{pmatrix}, \tag{4.3.1a}$$

$$= \begin{pmatrix} \mathrm{h}(x_3 - x_3^R)W^+(\mathbf{x},\mathbf{x}^R) & 0 \\ 0 & -\mathrm{h}(x_3 - x_3^R)W^-(\mathbf{x},\mathbf{x}^R) \end{pmatrix}. \tag{4.3.1b}$$

For state B we choose the one-way wave-vector of the actual medium,

$$\mathbf{P}(\mathbf{x},\mathbf{x}^S) = (P^+(\mathbf{x},\mathbf{x}^S)P^-(\mathbf{x},\mathbf{x}^S))^T, \tag{4.3.2}$$

with source state $(S^+(\mathbf{x}^S), S^-(\mathbf{x}^S))^T$ and medium state $\mathbf{B}$. For the derivation of the different states, the reader is referred to appendix A. If the states, as summarized in Table 4.2, are inserted into the reciprocity relation of Equation A.3.36 we arrive at,

$$\begin{aligned} \int_{\mathbf{x}\in\partial\mathbb{D}} \mathbf{G}^T(\mathbf{x},\mathbf{x}^R)\bar{\mathbf{I}}\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mu_3 \mathrm{d}A = \\ \int_{\mathbf{x}\in\mathbb{D}} \mathbf{G}^T(\mathbf{x},\mathbf{x}^R)\bar{\mathbf{I}}\mathbf{S}(\mathbf{x},\mathbf{x}^S) + \mathbf{S}_G^T(\mathbf{x},\mathbf{x}^R)\bar{\mathbf{I}}\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V \\ + \int_{\mathbf{x}\in\mathbb{D}} \mathbf{G}^T(\mathbf{x},\mathbf{x}^R)\mathbf{\Phi}(\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V, \end{aligned} \tag{4.3.3}$$

in which $\bar{\mathbf{I}}$ is defined as

$$\bar{\mathbf{I}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{4.3.4}$$

The operator $\mathbf{\Phi}$ denotes the contrast operator between the two medium states. Since the propagating parts of the states were chosen equal and the Green's state was chosen to be coupling free, this contrast operator reduces to the scattering part of the actual medium as derived in appendix A,

$$\mathbf{\Phi}(\mathbf{x}) = \begin{pmatrix} \mathcal{T}^+(\mathbf{x}) & \mathcal{R}^-(\mathbf{x}) \\ -\mathcal{R}^+(\mathbf{x}) & -\mathcal{T}^-(\mathbf{x}) \end{pmatrix}, \tag{4.3.5}$$

in which $\mathcal{T}^+$ and $\mathcal{T}^-$ are operators which account for transmission of up- and down-going waves respectively and $\mathcal{R}^+$ and $\mathcal{R}^-$ are operators which account for reflection of up-going waves (into down-going waves) and down-going waves (into up-going waves) respectively.

■ 4.3.2 Forward extrapolation

Since $\mathbf{G}$ originates from a point source at $\mathbf{x}^R$ and by considering $\mathbf{P}$ to originate from a point source at $\mathbf{x}^S$, the first volume integral of the one-way reciprocity relation of

Equation 4.3.3 reduces to,

$$\int_{\mathbf{x}\in\mathbb{D}} \mathbf{G}^T(\mathbf{x},\mathbf{x}^R)\bar{\mathbf{I}}\mathbf{S}(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}^s) + \mathbf{S}_G^T(\mathbf{x},\mathbf{x}^R)\delta(\mathbf{x}-\mathbf{x}^R)\bar{\mathbf{I}}^T\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V = \mathbf{G}(\mathbf{x}^S,\mathbf{x}^R)\bar{\mathbf{I}}\mathbf{S}(\mathbf{x}^S) + \bar{\mathbf{I}}\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S). \tag{4.3.6}$$

The second volume integral of Equation 4.3.3 denotes the scattered wavefield. From this expression useful integral representations are derived. Multiplying Equation 4.3.6 with $\bar{\mathbf{I}}^{-1}$ and using the reciprocity relation for the Green's matrix, and interchanging $\mathbf{x}^s$ and $\mathbf{x}^r$ results in the following expression,

$$\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S) = \mathbf{G}(\mathbf{x}^R,\mathbf{x}^S)\mathbf{S}(\mathbf{x}^S) + \int_{\mathbf{x}\in\partial\mathbb{D}} \mathbf{G}(\mathbf{x}^R,\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}\mathbf{x} + \int_{\mathbf{x}\in\mathbb{D}} \mathbf{G}(\mathbf{x}^R,\mathbf{x})\mathbf{\Phi}(\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V. \tag{4.3.7}$$

If we choose the boundary to extend to infinity, the surface integral vanishes. If we furthermore assume the volume to be source free (equivalent to subtracting the direct wavefield), we end up with

$$\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S) = \int_{\mathbf{x}\in\mathbb{D}} \mathbf{G}(\mathbf{x}^R,\mathbf{x})\mathbf{\Phi}(\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V. \tag{4.3.8}$$

By splitting the up and down-going parts of this equation, we retrieve for the up-going wavefield,

$$P^-(\mathbf{x}^R,\mathbf{x}^S) = \int_{\mathbf{x}\in\mathbb{D}} -G^{-,-}\left[(\mathbf{x}^R,\mathbf{x})\mathcal{R}^+(\mathbf{x})P^+(\mathbf{x},\mathbf{x}^S) + \mathcal{T}^-(\mathbf{x})P^-(\mathbf{x},\mathbf{x}^S)\right]\mathrm{d}V, \tag{4.3.9}$$

in which the operator $\mathcal{R}^+(\mathbf{x})$ is defined as,

$$\mathcal{R}^+(\mathbf{x}) = \int_{\mathbf{x}'\in\mathbb{D}\|x_3'=x_3} R(\mathbf{x},\mathbf{x}')P(\mathbf{x}')\mathrm{d}A, \tag{4.3.10}$$

and operator $\mathcal{R}^-(\mathbf{x})$ is defined likewise.

Another useful representation is derived when domain $\mathbb{D}$ is taken free of sources and contrast. Then by virtue of these choices, the volume integrals vanish and the resulting integral reads,

$$\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S) = \int_{\mathbf{x}\in\partial\mathbb{D}} \mathbf{G}(\mathbf{x}^R,\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\nu_3\mathrm{d}A. \tag{4.3.11}$$

If we take the upper and lower half space source and scatter free for the Green's state, then the Green's state has purely outgoing waves at the boundaries. Meaning

$G^{+,+}$ at the lower boundary and $G^{-,-}$ at the upper boundary. Then we can write for the down-going wavefield,

$$P^{+}(\mathbf{x}^R,\mathbf{x}^S)=\int_{\mathbf{x}\in\partial\mathbb{D}}\left[G^{+,+}(\mathbf{x}^R,\mathbf{x})P^{+}(\mathbf{x},\mathbf{x}^S)+G^{-,-}(\mathbf{x}^R,\mathbf{x})P^{-}(\mathbf{x},\mathbf{x}^S)\right]\nu_3\mathrm{d}A. \tag{4.3.12}$$

Note that the only contribution to the integral arises from wavefields that are traveling in the same direction. Therefore, if the upper half space is contrast free, the integral over the upper boundary vanishes. In the same line of reasoning, the integral over the lower boundary will vanish when the wavefield is excited by a source in the upper halfspace. Figure 4.3A depicts forward extrapolation.

■ 4.3.3 Inverse extrapolation

The backward propagating version of Equation 4.3.7 is obtained by using the reciprocity relation of the correlation type, Equation A.3.36, and following the same line of argumentation as before. By using,

$$\mathbf{K}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}\bar{\mathbf{I}}=\bar{\mathbf{J}}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}, \tag{4.3.13}$$

we arrive at,

$$\begin{aligned}\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S)=\mathbf{K}\mathbf{G}^{*}(\mathbf{x}^R,\mathbf{x}^S)\mathbf{K}\mathbf{S}(\mathbf{x}^S)+\int_{\mathbf{x}\in\partial\mathbb{D}}\mathbf{K}\mathbf{G}^{*}(\mathbf{x}^R,\mathbf{x})\mathbf{K}\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}\mathbf{x}\\+\int_{\mathbf{x}\in\mathbb{D}}\mathbf{K}\mathbf{G}^{*}(\mathbf{x}^R,\mathbf{x})\mathbf{K}\boldsymbol{\Phi}(\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V.\end{aligned} \tag{4.3.14}$$

Again, with the proper choice of the halfspaces for the Green's state and actual state, the same choice of boundary conditions and the boundaries extended to infinity we can write for the scattered wavefield,

$$\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S)=\int_{\mathbf{x}\in\mathbb{D}}\mathbf{K}\mathbf{G}^{*}(\mathbf{x}^R,\mathbf{x})\mathbf{K}\boldsymbol{\Phi}(\mathbf{x})\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}V. \tag{4.3.15}$$

Furthermore, for the inverse extrapolation, when the domain under consideration is source and contrast free, we can write,

$$\mathbf{P}(\mathbf{x}^R,\mathbf{x}^S)=\int_{\mathbf{x}\in\partial\mathbb{D}}\mathbf{K}\mathbf{G}^{*}(\mathbf{x}^R,\mathbf{x})\mathbf{K}\mathbf{P}(\mathbf{x},\mathbf{x}^S)\mathrm{d}A. \tag{4.3.16}$$

If we now investigate the up-going wavefield,

$$P^{-}(\mathbf{x}^R,\mathbf{x}^S)=\int_{\mathbf{x}\in\partial\mathbb{D}}\left[P^{+}(\mathbf{x},\mathbf{x}^S)G^{-,-*}(\mathbf{x}^R,\mathbf{x})+P^{-}(\mathbf{x},\mathbf{x}^S)G^{+,+*}(\mathbf{x}^R,\mathbf{x})\right]\mathrm{d}A, \tag{4.3.17}$$

we conclude that for inverse propagation only opposite traveling wavefields contribute, as depicted in Figure 4.3B.

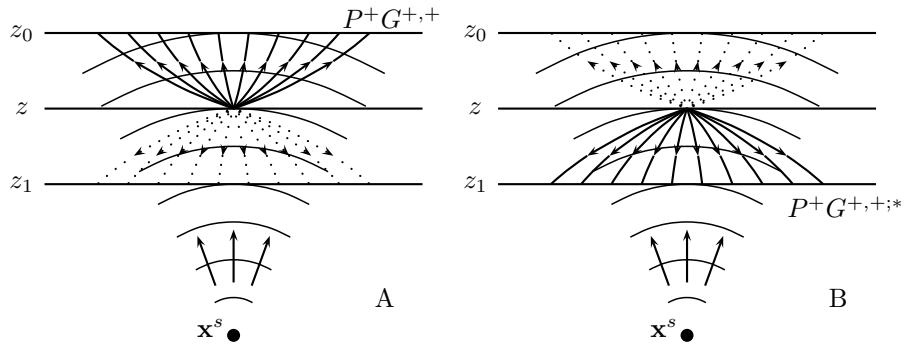

Figure 4.3: *Forward and inverse one-way wavefield extrapolation. When the wavefield is extrapolated from z to z0, only wavefields traveling in the same direction give contribution (A), whereas for inverse extrapolation from z to z1, only opposite traveling wavefields contribute (B).*

4.4 Physical field measurements, the WRW-matrix representation

In the previous section all the tools were derived to describe forward and inverse wavefield propagation as well as forward and inverse wavefield scattering. In this section the WRW model is derived, which is merely a discrete description of the foregoing results. Since measurements are generally taken in a discrete sense, the WRW model describes wavefield propagation and scattering in a convenient way.

From this model scattering series are easily derived, in which multiple reflections can be included and in which layer stripping methods as well as a redatuming schemes can be derived. Moreover, it provides an opening to model-independent operator determination, which will be used later on to describe the near surface in a kinematic sense. These operators can than be used to redatum the data to a level underneath the near surface, as if sources and receivers were initially placed at that level and hence do not suffer from near surface effects.

■ 4.4.1 Derivation of the WRW-model

To derive the WRW model a cylindrical volume $\mathbb{D}$ is considered bounded in the x_3-direction by two planes z_0 and z_1 with its radius extended to infinity. Sources and receivers are placed in the plane z_0. The scattered wavefield is described by

Equation 4.3.9 as,

$$\begin{aligned} P^-(\mathbf{x}^R,\mathbf{x}^S) = &\int_{\mathbf{x}\in\mathbb{D}} G^{-,-}(\mathbf{x}^R,\mathbf{x})\mathcal{R}^+(\mathbf{x})P^+(\mathbf{x},\mathbf{x}^s)\mathrm{d}V \\ &+ \int_{\mathbf{x}\in\mathbb{D}} G^{-,-}(\mathbf{x}^R,\mathbf{x})\mathcal{T}^-(\mathbf{x})P^-(\mathbf{x},\mathbf{x}^S)\mathrm{d}V, \end{aligned} \tag{4.4.1}$$

in which the reflection operator $\mathcal{R}(\mathbf{x})$ is defined by,

$$\mathcal{R}^+(\mathbf{x})P^+(\mathbf{x},\mathbf{x}^S) = \int_{\mathbf{x}'\in\partial\mathbb{D}} \left[R^+(\mathbf{x},\mathbf{x}')P^+(\mathbf{x}',\mathbf{x}^S)\right]_{x_3'=x_3} \mathrm{d}A. \tag{4.4.2}$$

The transmission operator $\mathcal{T}^-(\mathbf{x})$ is defined likewise and it is assumed that the medium outside z_0 and z_1 is free of sources and scatterers. Then for the primary scattered wavefield at z_1 we take $P^-(\mathbf{x},\mathbf{x}^S) = 0$ and for the incident wavefield we take,

$$P^+(\mathbf{x}') = \int_{\mathbf{x}\in\partial\mathbb{D},x_3=z_0} S^+(\mathbf{x})G^{+,+}(\mathbf{x}',\mathbf{x})\mathrm{d}A. \tag{4.4.3}$$

If we take for the source distribution,

$$S^+(\mathbf{x}) = S(\omega)\delta(\mathbf{x}-\mathbf{x}^S), \tag{4.4.4}$$

Equation 4.4.1 reduces to,

$$P^-(\mathbf{x}^R,\mathbf{x}^S) = \int_{x_3>z_0} W^-(\mathbf{x}^R,\mathbf{x}) \int_{\mathbf{x}'\in\mathbb{D},x_3'=x_3} R^+(\mathbf{x},\mathbf{x}')W^+(\mathbf{x}',\mathbf{x}^S)S(\omega)\mathrm{d}A\mathrm{d}\mathbf{x}, \tag{4.4.5}$$

in which we used the definition of the scattered free Green's operator of Equation 4.3.1b. This can be interpreted as the first order term of a Bremmer series, which forms the basis of the WRW formulation as introduced by Berkhout (1982).

■ 4.4.2 Surface free datum response

Since in seismics measurements are discrete, the WRW model of Equation 4.4.5 is discretized, resulting in a matrix formulation,

$$\mathbf{P}^-(z_r,z_s) = \sum_{m=1}^{M} \mathbf{W}^-(z_r,z_m)\mathbf{R}^+(z_m,z_m)\mathbf{W}^+(z_m,z_s)\mathbf{S}^+(z_s). \tag{4.4.6}$$

Each column of $\mathbf{P}^-$ represent the measurements at the discrete detector positions for one seismic experiment for one frequency. The matrix $\mathbf{S}^+$ represents the source distributions for all seismic experiments, for one temporal frequency. Each column of $\mathbf{W}^+$ describes downward propagation from one source location towards all locations along $\mathbf{R}^+$ and one row of $\mathbf{W}^-$ describes upward propagation from all locations along $\mathbf{R}^+$ toward one receiver location. Furthermore, the matrix $\mathbf{R}^+$ represents the

discretized version of the reflection operator for each depth level m, between z_0 and z_1.

If the sources are distributed as point-sources, as defined by Equation 4.4.4, the source matrix would read,

$$\mathbf{S}^+ = \mathbf{I}S(\omega). \tag{4.4.7}$$

Physical source directivity can be accomplished by including off-diagonal terms. It should also be noted that all operators depend on the angular frequency, ω, of which the notation is only left out for reasons of brevity.

For the expression of the scattered up-going wavefield with a reflection free surface, we can write,

$$\mathbf{P}^-(z_0) = \mathbf{X}(z_0, z_0)\mathbf{P}^+(z_0), \tag{4.4.8}$$

$$\mathbf{X}(z_0, z_0) = \sum_{m=1}^{M} \mathbf{W}^-(z_0, z_m)\mathbf{R}^+(z_m, z_m)\mathbf{W}^+(z_m, z_0), \tag{4.4.9}$$

in which internal multiples are neglected. In Wapenaar and Berkhout (1989) it is stated that to be exact, $\mathbf{R}$ is of more complex nature. However, it is also recommended not to use the exact, more complex definition of $\mathbf{R}$ for matters of stability in the presence of noise. Note that the downgoing wavefield, $\mathbf{P}^+$ is written here as a generalized form of the downgoing source wavefield $\mathbf{S}^+$.

The emphasis in this thesis is laid on redatuming through the overburden. Thus we will distinct between the overburden and the target response. With,

$$\mathbf{W}^+(z, z_0) = \mathbf{W}^+(z, z_d)\mathbf{W}^+(z_d, z_0), \tag{4.4.10}$$

$$\mathbf{W}^-(z_0, z) = \mathbf{W}^-(z_0, z_d)\mathbf{W}^-(z_d, z), \tag{4.4.11}$$

in which z_d is the datum level to which we would to like redatum our data, Equation 4.4.9 can be written as

$$\mathbf{X}(z_0, z_0) = \sum_{m=0}^{d} \mathbf{W}^-(z_0, z_m)\mathbf{R}^+(z_m, z_m)\mathbf{W}^+(z_m, z_0) + \mathbf{W}^-(z_0, z_d)\mathbf{X}(z_d, z_d)\mathbf{W}^+(z_d, z_0). \tag{4.4.12}$$

The first term of Equation 4.4.12 represents the scattered wavefield of the overburden, resulting from a scatter distribution within the overburden, which will be denoted from here onward as,

$$\mathbf{X}_{ob}(z_0, z_0) = \sum_{m=0}^{d} \mathbf{W}^-(z_0, z_m)\mathbf{R}^+(z_m, z_m)\mathbf{W}^+(z_m, z_0). \tag{4.4.13}$$

The removal of this term is not direct subject to this thesis and for state of the art techniques the reader is referred to Ernst (1999), where Green's functions are estimated explicitly for near surface guided waves, which are used to invert for the scatter distribution. If this part is neglected or removed, then the second part of Equation 4.4.12 resembles the target response as measured at the surface (see also Figure 4.5).

■ 4.4.3 Free surface datum response

To this point the surface was assumed to be reflection free. In practice, surface z_0 resembles the free surface, which is a perfect reflector. Thus, the total down-going wavefield should be written as,

$$\mathbf{P}^+(z_0, z_0) = \mathbf{P}_s^+(z_0, z_0) + \mathbf{P}_r^+(z_0, z_0), \tag{4.4.14}$$

in which the $\mathbf{P}_s^+(z_0, z_0)$ is the direct down-going wavefield and $\mathbf{P}_r^+(z_0, z_0)$ the reflected up-going wavefield. The down-going wavefield can then be written as,

$$\mathbf{P}^+(z_0, z_0) = \mathbf{R}_{fr}^-(z_0)\mathbf{P}^-(z_0, z_0) + \mathbf{P}_s^+(z_0, z_0), \tag{4.4.15}$$

in which $\mathbf{R}_{fr}^-(z_0)$ resembles the reflectivity of the free surface which is equal to -1 in case of an acoustic free surface. Thus we can write for the total up-going wavefield,

$$\mathbf{P}^-(z_0, z_0) = \mathbf{X}(z_0, z_0)\left[\mathbf{R}_{fr}^-(z_0)\mathbf{P}^-(z_0, z_0) + \mathbf{P}_s^+(z_0, z_0)\right], \tag{4.4.16}$$

which is explicitly written as,

$$\mathbf{P}^-(z_0, z_0) = \mathbf{X}_{fr}(z_0, z_0)\mathbf{P}_s^+(z_0, z_0). \tag{4.4.17}$$

The free surface response matrix is then defined as,

$$\mathbf{X}_{fr}(z_0, z_0) = \left[\mathbf{I} - \mathbf{X}(z_0, z_0)\mathbf{R}_{fr}^-(z_0)\right]^{-1}\mathbf{X}(z_0, z_0). \tag{4.4.18}$$

This can be written as a series expansion,

$$\mathbf{X}_{fr}(z_0, z_0) = \left[\mathbf{I} + \sum_{m=1}^{\infty}\left(\mathbf{X}(z_0, z_0)\mathbf{R}_{fr}^-(z_0)\right)^m\right]\mathbf{X}(z_0, z_0), \tag{4.4.19}$$

which clearly shows the infinite number of multiple reflections which are generated by the free surface as schematically displayed in Figure 4.4.

■ 4.4.4 Two-way target response

Until here, the forward model was derived in terms of one-way propagation of one-way wavefields. In practice, use is made of two-way seismic data. In Section 3.3 the

decoupling for two-way wavefields into one-way wavefields was discussed. Here the relationship between the two-way seismic source or detector and the one way up-and down-going wavefields as influenced by the free surface is discussed.

For a pressure source at point locations along the free surface we only emit a down-going pressure wavefield,

$$\mathbf{S}^{+}(z_0) = \mathbf{I}S(\omega). \tag{4.4.20}$$

For a buried volume injection source, at point locations along a depth z_s, the down-going wavefield contains a ghost, resulting from reflection of the up-going wavefield at the free surface, which can be derived as,

$$\mathbf{S}^{+}(z_0) = \frac{1}{2}\left(\mathbf{W}^{-,*}(z_s, z_0) - \mathbf{W}^{-}(z_s, z_0)\right) S(\omega). \tag{4.4.21}$$

Similarly, the measured two-way wavefield at the detector is constructed from the up-going one-way wavefield. For a velocity detector, a geophone, at the free surface the measured wavefield is simply,

$$V_z(z_0) = \frac{-1}{j\omega\rho}\frac{\partial P^{-}(z_0)}{\partial z} \tag{4.4.22}$$

For buried receivers, hydrophones in marine seismics, the measured wavefield can be expressed by,

$$\mathbf{P}(z_0) = \left(\mathbf{W}^{+,*}(z_r, z_0) - \mathbf{W}^{+}(z_r, z_0)\right)\mathbf{P}^{-}(z_r), \tag{4.4.23}$$

n which z_r resembles the depth of the actual receiver. Equation 4.4.23 again contains a ghost.

The source and receiver responses can be included within respectively operators $\mathbf{D}^{+}$ and $\mathbf{D}^{-}$, such that the total data equation reads,

$$\mathbf{P} = \mathbf{D}^{-}(z_0)\mathbf{X}_{fr}(z_0, z_0)\mathbf{D}^{+}(z_0)S(\omega), \tag{4.4.24}$$

as depicted in Figure 4.4B, in which $\mathbf{X}_{fr}(z_0, z_0)$ resembles the multiple series of Equation 4.4.19.

To derive the wavefield as if it was measured underneath the near surface, or overburden, the following processing steps should be accounted for:

- Decomposition of sources and receivers. This will effectively remove the operators $\mathbf{D}^{+}$ and $\mathbf{D}^{-}$. In Schoolmeesters (2001) a detailed study is described to remove the ghost effects present in the operators. Volker (2002), describes the effects of directivity and how, by means of the WRW model, acquisition designs can be assessed to control the directivity patterns in order to optimize target illumination.

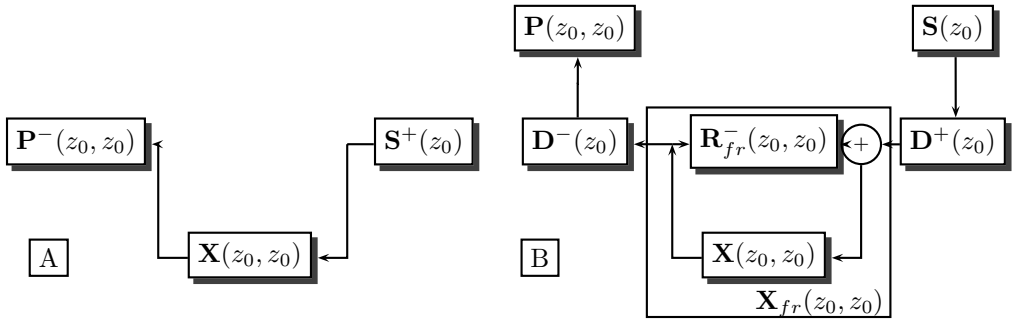

Figure 4.4: *a. The forward model of the primary wavefield. b. The model of the measured wavefield, including multiple contributions and directivity patterns.*

- Multiple attenuation. After correcting for source and receiver characteristics, multiple energy should be removed or attenuated. The reader is referred to van Dedem and Verschuur (2005) for extensive discussions on multiple attenuation, based on the formulation of Equation 4.4.19.

- Removal of near surface scattering. After the previous two processing steps, the data is described by,

$$\mathbf{P}^-(z_0) = \mathbf{X}(z_0, z_0)S(\omega) + \mathbf{X}_{ob}(z_0, z_0)S(\omega). \tag{4.4.25}$$

  The overburden response can be removed by techniques described by Ernst (1999) and Campman (2005).

- Near surface redatuming. In the last stage, the propagation through the overburden is removed. According to Equation 4.4.12, the target response after removal of the overburden response would read,

$$\mathbf{X}(z_0, z_0) = \mathbf{W}^-(z_0, z_d)\mathbf{X}(z_d, z_d)\mathbf{W}^+(z_d, z_0). \tag{4.4.26}$$

  The operators $\mathbf{W}^-(z_0, z_d)$ and $\mathbf{W}^+(z_d, z_0)$ describe the one-way, scattering free propagation characteristics of the overburden. The aim in this thesis is to define inverse propagators that effectively remove the forward propagation characteristics through,

$$\begin{aligned} &\mathbf{F}^-(z_d, z_0)\mathbf{X}(z_0, z_0)\mathbf{F}^+(z_0, z_d) = \\ &\mathbf{F}^-(z_d, z_0)\mathbf{W}^-(z_0, z_d)\mathbf{X}(z_d, z_d)\mathbf{W}^+(z_d, z_0)\mathbf{F}^+(z_0, z_d), \end{aligned} \tag{4.4.27}$$

  in which ideally the operators $\mathbf{F}^-(z_d, z_0)$ and $\mathbf{F}^+(z_0, z_d)$, that need to be defined, should be the inverse of the forward propagators $\mathbf{W}^-(z_0, z_d)$ and $\mathbf{W}^+(z_d, z_0)$. Through this procedure synthetically data will be generated that resembles signal excited and recorded at datum-level z_d.

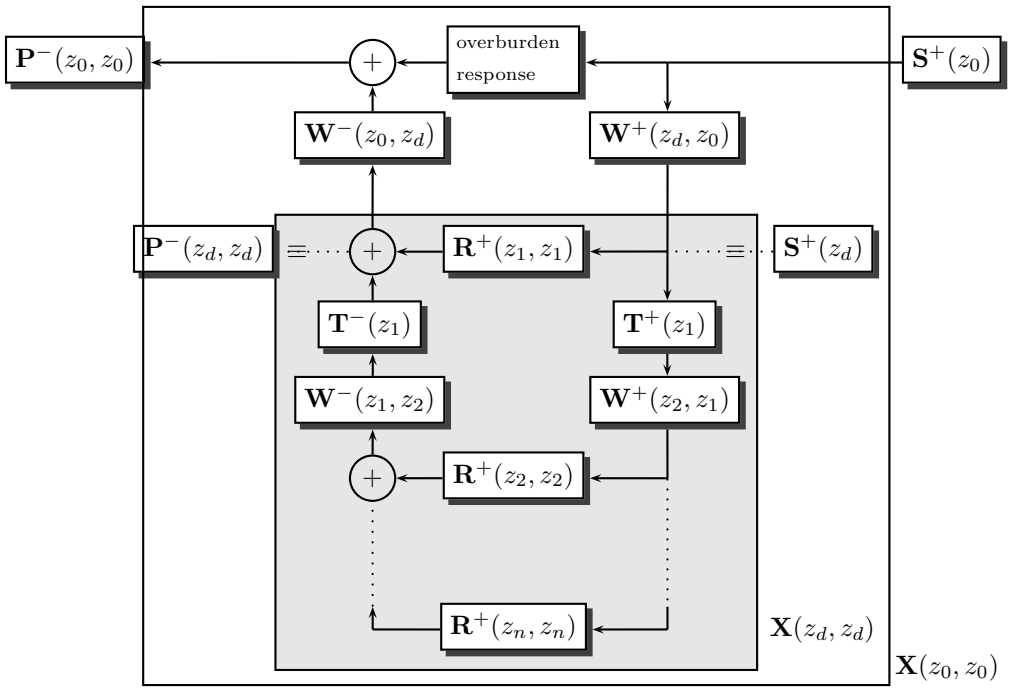

Figure 4.5: *The primary wavefield is expressed in terms of reflecting events. The earth response can be decoupled in a target response and an overburden response.*

In theory multiples should be removed right after source and receiver decomposition, since multiple energy is produced by all sources, including near surface scattering. However, in practice multiple energy will be removed after correcting for near surface scattering, since this would be more robust in terms of processing under the assumption that near surface scatter energy reflected multiple times at the free surface will be weak. The method described in Campman (2005) estimates and images near surface scatterers after which scatter energy is estimated and removed for each subsequent arrival. As such it will become more practical to remove near surface scatter energy before multiple removal within a processing scheme.

The last processing step is subject to the remainder of this thesis. The propagators $\mathbf{W}$ describe the kinematics through the overburden and were derived under the assumption of a scatter-free halfspace. However, in general the near surface overburden is not scatter free at all. It actually are the contrasts and scatterers

that cause the specific nature of the near surface overburden. On top of these unaccounted transmission losses quite often coupling effects at the sources and receivers result into additional unaccounted amplitude effects. Furthermore, their amplitude behavior depend on the losses in the overburden, which were not accounted for in the previous steps. For instance coupling effects at the sources and receivers can be present. Furthermore, the amplitude behavior of the propagators, due to transmission losses, is not easy to account for.

In the next sections we will further focus on the kinematic part of the inverse operators. In Chapter 7 we will further discuss how to account for amplitude effects when the inverse operators found are used for redatuming purposes.

4.5 Near surface redatuming

In the previous section is was shown that according to Equation 4.4.12 the total wavefield can be divided into an overburden response and a target response. The first term of the response operator of Equation 4.4.12 defines the overburden response, while the second term defines the target response.
It is our aim to inverse extrapolate the data to the target zone on which we will focus in this section.

■ 4.5.1 Derivation of the inverse propagation operators

The goal is to define the forward operators $\mathbf{W}^-(z_0, z_d)$ and $\mathbf{W}^+(z_d, z_0)$, which can be used for inverse extrapolation to remove the forward propagation through,

$$\mathbf{P}^-(z_d, z_d) = \mathbf{F}^-(z_d, z_0)\mathbf{P}^-(z_0, z_0)\mathbf{F}^+(z_0, z_d) = \mathbf{X}(z_d, z_d)S(\omega), \tag{4.5.28}$$

when the inverse propagators are defined as,

$$\mathbf{F}^-(z_d, z_0) = \left[\mathbf{W}^-(z_0, z_d)\right]^{-1}, \tag{4.5.29}$$

$$\mathbf{F}^+(z_0, z_d) = \left[\mathbf{W}^+(z_d, z_0)\right]^{-1}. \tag{4.5.30}$$

Due to its filtering nature, physically understood through the inclusion of evanescent waves, a direct inversion of the forward operators can not be established; the forward operators are rank deficient. To define the inverse propagators various possible methods are available, such as,

- Inversion by Singular Value Decomposition of the forward propagators. Singular Value Decomposition provides a matrix factorization through which it is possible to determine a pseudo inverse, which describes the invertible part of the forward operator. Following Strang (1988) any operator $\mathbf{W}$ of size $m \times n$ can be factored into,

$$\mathbf{W} = \mathbf{U}_1\mathbf{\Sigma}\mathbf{U}_2^H \tag{4.5.31}$$

in which matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ are unitary (complex orthogonal) and $\mathbf{\Sigma}$ is a diagonal matrix of size $r \times r$. The size of $\mathbf{\Sigma}$ is give by r, which is smaller than m and n and denotes the rank of the operator. The elements on the main diagonal of $\mathbf{\Sigma}$, $(\sigma_1^2, \cdots, \sigma_r^2)$, resemble the so-called singular values of $\mathbf{W}$. Once the Singular Value Decomposition is established, the pseudo-inverse of $\mathbf{W}$ is found to be,

$$\mathbf{F} = \mathbf{U}2\mathbf{\Sigma}^{-1}\mathbf{U}_1^H. \tag{4.5.32}$$

For further details and algorithms to define the Singular Value Decomposition, the reader is referred to Strang (1988) and Golub and van Loan (1996).

- Weighted least-squares inversion of the forward propagators. The inverse propagator is described by,

$$\mathbf{F} = \left[\mathbf{W}^H\mathbf{C}_1\mathbf{W} + \lambda^2\mathbf{C}_2\right]^{-1}\mathbf{W}^H\mathbf{C}_1, \tag{4.5.33}$$

which holds for the bot the combination $\{\mathbf{F}^-(z_d, z_0), \mathbf{W}^-(z_0, z_d)\}$ as well as the combination $\{\mathbf{F}^+(z_0, z_d), \mathbf{W}^+(z_d, z_0)\}$. The operators $\mathbf{C}_1$ and $\mathbf{C}_2$ depend on the properties of signal and noise. A more detailed discussion on least-squares inversion can be found in Appendix B. In Chapter 7, least-squares redatuming will be used to account for amplitude effects, in case of severe energy loss due to near surface scattering and transmission effects.

- The matched filter, which is exact in homogeneous media if the contribution of evanescent waves is neglected,

$$\mathbf{F}^+(z_0, z_d) = \left[\mathbf{W}^-(z_d, z_0)\right]^*, \tag{4.5.34}$$

$$\mathbf{F}^-(z_d, z_0) = \left[\mathbf{W}^+(z_0, z_d)\right]^*. \tag{4.5.35}$$

In the case of the medium being homogeneous with sources and receivers at the same positions the following holds,

$$\mathbf{W}^+(z_d, z_0) = \left[\mathbf{W}^-(z_0, z_d)\right]^T. \tag{4.5.36}$$

Note here, that the matched filter is actually a first step in least-squares inversion for homogeneous media when sources and receivers are located at the same positions.

Especially for the derivation of the kinematic part of the operators, the matched filter approach will prove its attractiveness due to its simplicity, which is demonstrated in the remainder of this chapter. The properties as derived for the matched filter approach will advantageously be used to determine the correct propagators in terms of phase. In Chapter 7 we will use the least-squares approach for operators with correct phase, to determine the proper amplitudes for true amplitude redatuming when necessary.

■ 4.5.2 Focusing operators

The process of removing the forward propagators by applying inverse operators can be considered as a two-step procedure, which subsequently focuses the source emitted energy and the receiver detected energy in a synthetic sense. As described in Berkhout (1997a) and Thorbecke (199j) this two-step procedure can be considered as focusing in emission and focusing in detection.

Focusing in emission

Focusing in emission describes the removal of the propagation part, $\mathbf{W}^{+}(z_d, z_0)$, emanating from the source side, which by definition equals

$$\mathbf{P}^{-}(z_0, z_0)\mathbf{F}^{+}(z_0, z_d) = \mathbf{W}^{-}(z_0, z_d)\mathbf{X}(z_d, z_d)S(\omega). \tag{4.5.37}$$

Focusing in emission, describes the process of backward propagating the sources through the near-surface, toward the defined boundary between near-surface and deeper subsurface and resembles the discrete analogy of Equation 4.2.12 with a Dirichlet boundary condition (Rayleigh equation of the second kind), for two-way wavefields and Equation 3.3.12 for one-way wavefields. The principle is graphically illustrated in Figure 4.6.

Focusing in detection

Focusing in detection describes the removal of the propagation part, $\mathbf{W}^{-}(z_0, z_d)$, detected at the receiver side, which equals,

$$\mathbf{F}^{-}(z_d, z_0)\mathbf{P}^{-}(z_0, z_0) = \mathbf{X}(z_d, z_d)\mathbf{W}^{+}(z_d, z_0)S(\omega). \tag{4.5.38}$$

Focusing in detection, describes the process of backward propagating the receivers through the near-surface, toward the defined boundary between near-surface and deeper subsurface and resembles the discrete analogy of Equation 4.2.12 with a Dirichlet boundary condition (Rayleigh equation of the second kind), for two-way wavefields and Equation 3.3.12 for one-way wavefields. The principle is graphically displayed by Figure 4.6

If we are only interested in the phase characteristics and ignore the amplitudes, then by virtue of the definitions of the matched filter inverse propagator, focusing in detection and emission respectively can be written as,

$$\mathbf{W}^{+,*}(z_d, z_0)\mathbf{P}^{-}(z_0, z_0) = \mathbf{X}(z_d, z_d)\mathbf{W}^{+}(z_d, z_0)S(\omega), \tag{4.5.39}$$

$$\mathbf{P}^{-}(z_0, z_0)\mathbf{W}^{-,*}(z_0, z_d) = \mathbf{W}^{-}(z_0, z_d)\mathbf{X}(z_d, z_d)S(\omega). \tag{4.5.40}$$

In our analysis we will focus on one particular event within the recorded data. This event is defined to mark the transition of the assigned near surface and the deeper subsurface. This particular event is resembled by its reflection response $\mathbf{R}^{+}(z_d)$ and

is deduced from the subsurface response $\mathbf{X}(z_0, z_0)$. When we are only interested in phase characteristics and ignore angle dependent reflection, we can write for the earth response,

$$\begin{aligned} \mathbf{X}(z_0, z_0) &= \mathbf{W}^-(z_0, z_d)\mathbf{R}^+(z_d, z_d)\mathbf{W}^+(z_d, z_0), & (4.5.41)\\ &= r_t\mathbf{W}^-(z_0, z_d)\mathbf{W}^+(z_d, z_0). & (4.5.42)\end{aligned}$$

If the sources and receivers are assumed to be placed at the same surface locations, such that Equation 4.5.36 holds, Equations 4.5.39 and 4.5.40 can be written as,

$$\left[\mathbf{W}^-(z_0, z_d)\right]^H \mathbf{P}^-(z_0, z_0) = r(z_d)\left[\mathbf{W}^-(z_0, z_d)\right]^T S(\omega), \quad (4.5.43)$$

$$\mathbf{P}^-(z_0, z_0)\left[\mathbf{W}^+(z_d, z_0)\right]^H = r(z_d)\left[\mathbf{W}^+(z_d, z_0)\right]^T S(\omega). \quad (4.5.44)$$

From Equation 4.5.43 and Equation 4.5.44, it can be concluded that when the correct (in terms of phase) inverse propagator is applied to the measured data, apart from amplitude terms, the result equals to the complex conjugate of the applied operator for a specific event, which is referred to as the *principle of equal traveltime*(Berkhout, 1997a; Thorbecke, 199j).

Amplitude effects

Here we consider one focusing operator $\mathbf{f}_i(z_d, z_0)$, a row-vector of the inverse propagator $\mathbf{F}(z_d, z_0)$ at the detector side. By selecting only one row from the inverse propagator, the inverse propagation will focus the detected energy to only one particular subsurface point i of interest. Here, a specific point just above the target level z_d is considered. Then for correct focusing the operator has to obey,

$$\begin{aligned} \mathbf{f}_i(z_d, z_0)\mathbf{W}^-(z_0, z_d) &= \mathbf{i}_i, & (4.5.45)\\ i_{ij} &= \begin{cases} 0, & \text{for } i \neq j \\ 1, & \text{for } i = j \end{cases}. & (4.5.46)\end{aligned}$$

If we assume the scatter distributions to behave non-dispersively, i.e. the reflection effect is not depending on frequency, the inverse operators differ only in amplitude and are correct under the evanescent wave assumption.

To show the effect of amplitude behavior, we consider here a stack of N homogeneous laterally invariant layers. In this case the matched filter approach for a set of operators describing one particular reflector reads,

$$\begin{aligned} \mathbf{F}^-(z_N, z_0)\mathbf{W}^-(z_0, z_N) = &\prod_{n=1}^{N}\left[\mathbf{W}^+(z_{n-1}, z_n)\right]^* \mathbf{T}^+(z_{n-1}) \\ &\times \prod_{n=1}^{N}\mathbf{W}^-(z_{n-1}, z_n)\mathbf{T}^-(z_{n-1}). \end{aligned} \quad (4.5.47)$$

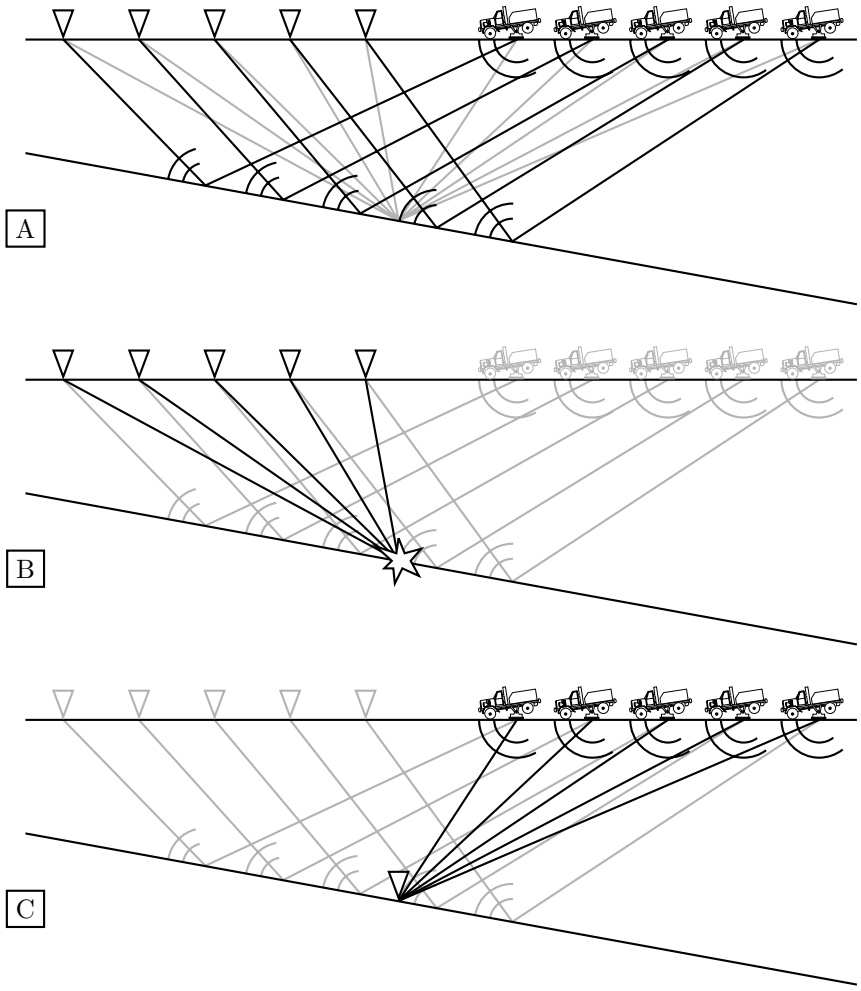

Figure 4.6: *Physical principle of focusing in emission and focusing in detection. A) shows the specular rays for one specific common offset selection, with in gray the raypaths associated with one focusing operator. B) shows the resulting rays after focusing in emission of all measured data. C) shows the resulting rays after focusing in detection of all measured data.*

For homogeneous layers, each row of the propagation operator is a shifted version of the other rows. Such a structure is known as a Toeplitz structure and resembles a convolution. For this reason all operators become diagonal operators when transformed to the wavenumber domain. Each diagonal element represents a specific wavenumber k_x, for which we can write,

$$\bar{F}(k_x)\bar{W}^-(k_x) = \prod_{n=1}^{N} \left[\bar{W}_n^+(k_x)\right]^* \bar{T}_n^+(k_x) \prod_{n=1}^{N} \bar{W}_n^-(k_x)\bar{T}_n^-(k_x). \tag{4.5.48}$$

From Berkhout (1982), we can write for the transmission and reflection in the wavenumber domain,

$$1 + \bar{R}^+(k_x) = \bar{T}^+(k_x), \tag{4.5.49}$$
$$1 + \bar{R}^-(k_x) = \bar{T}^-(k_x), \tag{4.5.50}$$
$$1 - \bar{R}^+(k_x) = \bar{T}^-(k_x). \tag{4.5.51}$$

Thus in the wavenumber domain the focusing response can be written as

$$\bar{F}^-(k_x)\bar{W}^-(k_x) = \prod_{n=1}^{N} (1 - [\bar{R}_n^+(k_x)]^2). \tag{4.5.52}$$

Similarly, for focusing in detection we can write,

$$\bar{W}^+(k_x)\bar{F}^+(k_x) = \prod_{n=1}^{N} (1 - [\bar{R}_n^+(k_x)]^2). \tag{4.5.53}$$

Then, after both focusing in emission as well as focusing in detection, the data can be written in the wavenumber domain as,

$$\bar{F}^-(k_x)\bar{W}^-(k_x)\bar{R}_d^+\bar{W}^+(k_x)\bar{F}^+(k_x) = \bar{R}_d^+ \prod_{n=1}^{N} \left(1 - \left[\bar{R}_n^+(k_x)\right]^2\right)^2, \tag{4.5.54}$$

which can result in significant reduction of energy if the matched filter approach is used for near surface redatuming.

Using the matched filter approach for focusing, resolves the phase behavior of the measured wavefield for a particular subsurface point of interest. However, especially if we try to describe complex near surface behavior by means of a single focusing step, the risk exists that amplitudes will be degraded significantly. Therefore we will use the matched filter approach to retrieve the phase behavior of the near surface, whereas a least-squares approach will be used in Chapter 7 to actually redatum the measured data to a chosen datum underneath the complex near surface.

4.6 Focusing operator updating

In the previous sections the measured, primary reflection data is described in terms of downward propagation from the source to the scattering domain, scattering and propagation from the scatterer to the receiver. Furthermore, the inverse propagators were defined in terms of focusing in detection and focusing in emission operations. However, in general the medium properties are not known on forehand and therefore no direct knowledge is available on the correctness of forward propagators and hence the inverse propagators. With respect to the phase behavior of the propagators two properties can be determined in order to define a data-driven methodology to derive the correct inverse propagators, as was recognized by Bolte (1997) as well. These properties are the so called *principal of equal traveltime* and *Fermat modeling*. A third property can be derived but is not directly a data determined property.

■ 4.6.1 Principle of equal traveltime

An operator $\Gamma_i(\vec{z_0})$, defining a synthesized source $\vec{\Gamma}_i(z_0)S(\omega)$, can be applied to the recorded data, which will synthesize the recorded wavefield into a Common Focus Point (CFP) gather,

$$\mathbf{P}^-(z_0)\vec{\Gamma}_i(z_0) = \mathbf{W}^-(z_0,z_d)\mathbf{R}^+(z_d,z_d)\mathbf{W}^+(z_d,z_0)\vec{\Gamma}_i(z_0)S(\omega). \tag{4.6.55}$$

By demanding the operator to focus on a point $\mathbf{x}_i$ at depth z_d,

$$\mathbf{W}^+(z_d,z_0)\vec{\Gamma}_i(z_0,z_d) = [0,\cdots,0,1,0,\cdots,0]^T = \vec{I}_i, \tag{4.6.56}$$

and using $[\mathbf{W}^+(z_d,z_0)]^{-1} \approx [\mathbf{W}^-(z_0,z_d)]^*$, the focusing operator $\vec{\Gamma}_i(z_0,z_d)$ can be defined through,

$$\vec{\Gamma}_i(z_0,z_d) = \left[\mathbf{W}^-(z_0,z_d)\right]^* \vec{I}_i. \tag{4.6.57}$$

When we are only interested in structural imaging, and not in e.g. amplitude properties, the reflection operator $\mathbf{R}^+$ is assumed to be a diagonal matrix. Synthesizing the recorded wavefield with the defined synthesis operator yields the CFP gather,

$$\begin{aligned} \mathbf{P}^-(z_0,z_0)\vec{\Gamma}_i(z_0,z_d) &= \mathbf{W}^-(z_0,z_d)\vec{I}_i R_i^+ S(\omega), & (4.6.58)\\ &\propto \left[\vec{\Gamma}_i(z_0,z_d)\right]^*. & (4.6.59) \end{aligned}$$

Apart from an amplitude factor, the CFP gather should coincide with its time reversed focusing operator. This is also known as the 'principal of equal traveltime'(Berkhout, 1997b) which is shown in Figure 4.6.1 The operator illuminating a subsurface point should coincide with the wavefield induced by a point source at the same subsurface location. By convolving the phase response of the operator and CFP gather along the time-coordinate, we should obtain a flat event around $t = 0s$ in the so called Differential Time Shift (DTS) panel,

$$\Delta\mathbf{P}^-(z_0,z_d) = Re\left\{\mathbf{W}^-(z_0,z_d)\right\}\vec{I}_i R_i^+ S(\omega), \tag{4.6.60}$$

in which the phase of the CFP gather is corrected for, whereas the amplitude remains.

The DTS panel provides a control whether a correct focusing operator has been applied or not, without having knowledge about the velocity model. If an erroneous operator is chosen, the event the focusing operator was estimated for will no longer align around $t = 0s$. within the DTS panel.

Suppose a forward propagator $< \mathbf{W}^-(z_0, z_d) >$ was modeled within an incorrect velocity model,

$$< \mathbf{W}^-(z_0, z_d) >= \mathbf{W}^-(z_0, z_d)\Delta\mathbf{W}, \tag{4.6.61}$$

in which $\Delta\mathbf{W}$ is equivalent with the made error. Then one of the focusing operators would read,

$$\begin{aligned} \vec{\Gamma}_i(z_0, z_d) &= \left[\mathbf{W}^-(z_0, z_d)\right]^* \Delta\mathbf{W}^* \vec{I}_i, & (4.6.62)\\ &\approx \left[\mathbf{W}^+(z_d, z_0)\right]^{-1} \Delta\mathbf{W}^* \vec{I}_i. & (4.6.63) \end{aligned}$$

Synthesizing the recorded wavefield, consistent with the correct model, with this incorrect operator would then yield,

$$\mathbf{P}^-(z_0, z_0)\vec{\Gamma}_i(z_0, z_d) = \mathbf{W}^-(z_0, z_d)\mathbf{R}^+(z_d, z_d)\Delta\mathbf{W}^* \vec{I}_i S(\omega). \tag{4.6.64}$$

If this result is compared with the complex conjugate of the focusing operator used,

$$\left[\vec{\Gamma}_i(z_0, z_d)\right]^* = \mathbf{W}^-(z_0, z_d)\Delta\mathbf{W}\vec{I}_i, \tag{4.6.65}$$

it is observed that a phase error $\Delta\mathbf{W}$ in the operator would lead to an opposite phase error $\Delta\mathbf{W}^*$ in the synthesized result. This knowledge can be used to improve on the erroneous focusing operator: by choosing to update the erroneous operator by adding half of the times picked in the DTS panel, convergence to an operator which describes the propagation characteristics better will be obtained without having knowledge of the underlying velocity model.

To demonstrate the principle of equal traveltime, fixed-spread shot-records were modeled over a simple model consisting of 3 horizontal reflectors at a depth of 400m, 600m and 800m respectively. The shot and receiver spacing was chosen to be 25m. The contrast consists of density contrasts only. Figure 4.8A shows 3 CFP gathers, after one focusing step, using the correct operators which are plotted within the figure. By correlating the operators in time with their corresponding CFP gathers, the DTS gathers are obtained, depicted in Figure 4.8B. Clearly, for the focused event the DTS gathers align around $t = 0$.

When the operators are modeled with an erroneous estimate of the earth's propagation characteristics, the DTS gathers are no longer aligned around $t = 0$. Figure

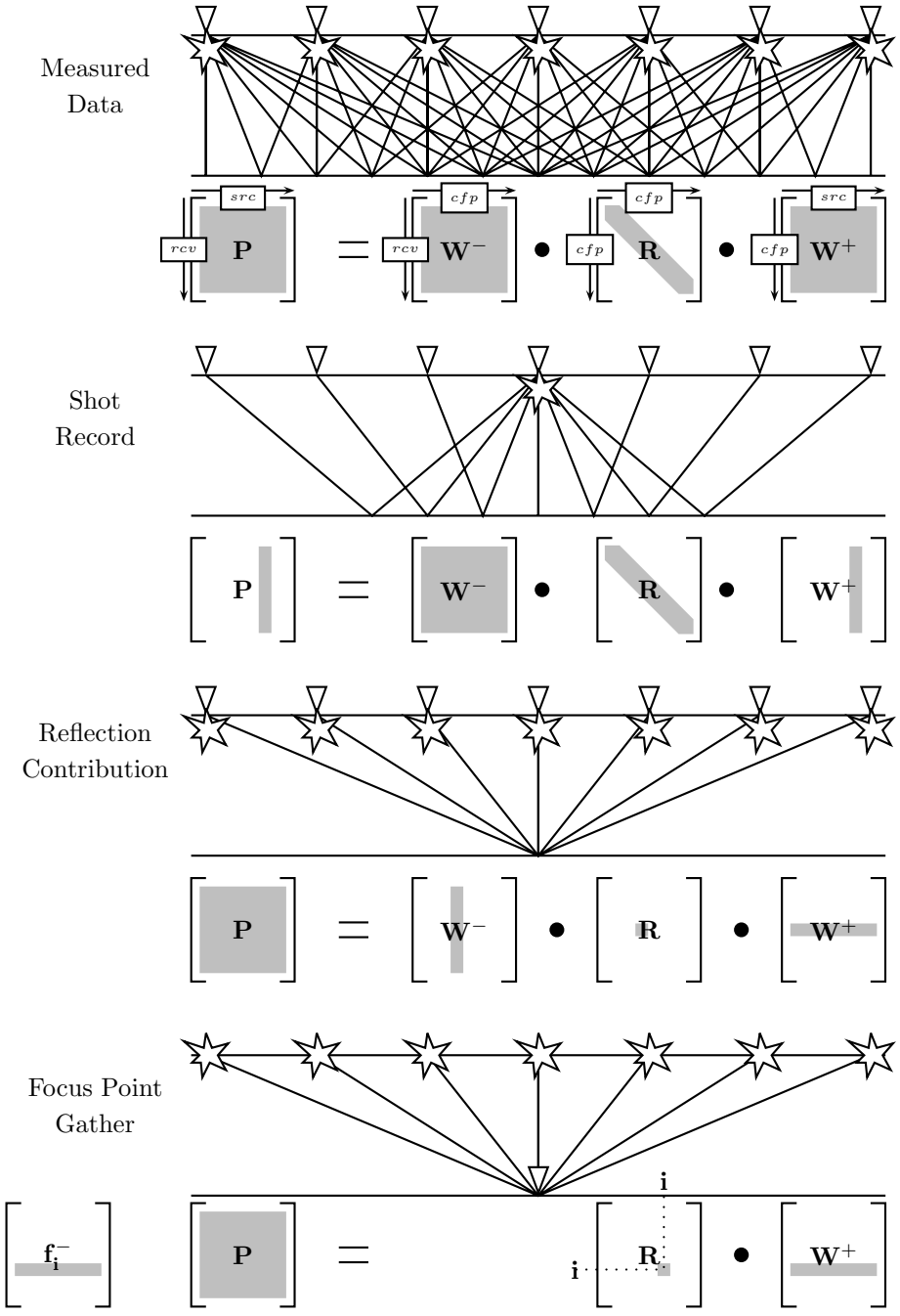

Figure 4.7: *Schematic representation of WRW model and associated focusing.*

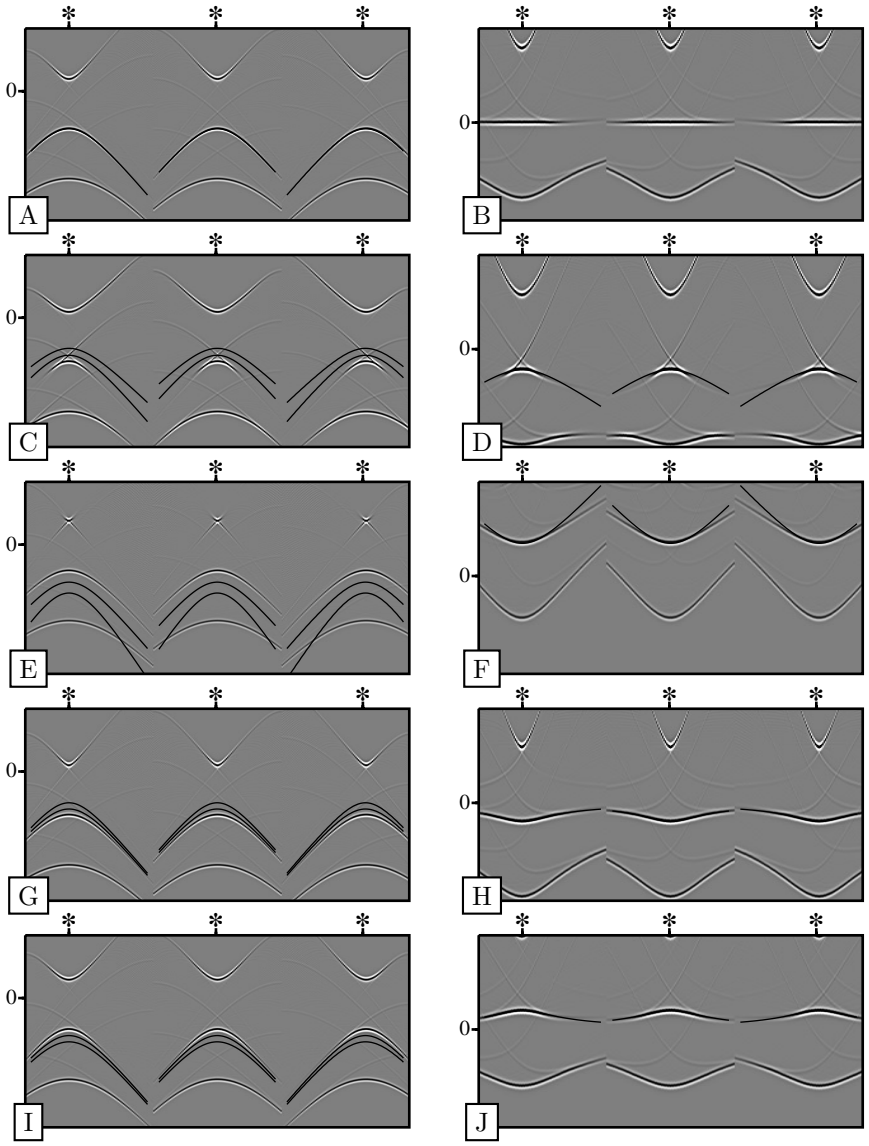

Figure 4.8: *Principle of equal traveltime in the space time domain. The left column shows CFP gathers for three CFP's, overlain with their focusing operators. The right column shows the DTS gathers, overlain with 2 times the difference between the operators and the error free operators. (A,B) Correct focusing operators. (C,D) too high velocity used in operator modeling. (E,F) too low velocity used in operator modeling. (G,H) too deep depth used in operator modeling. (I,J) too shallow depth used in operator modeling.*

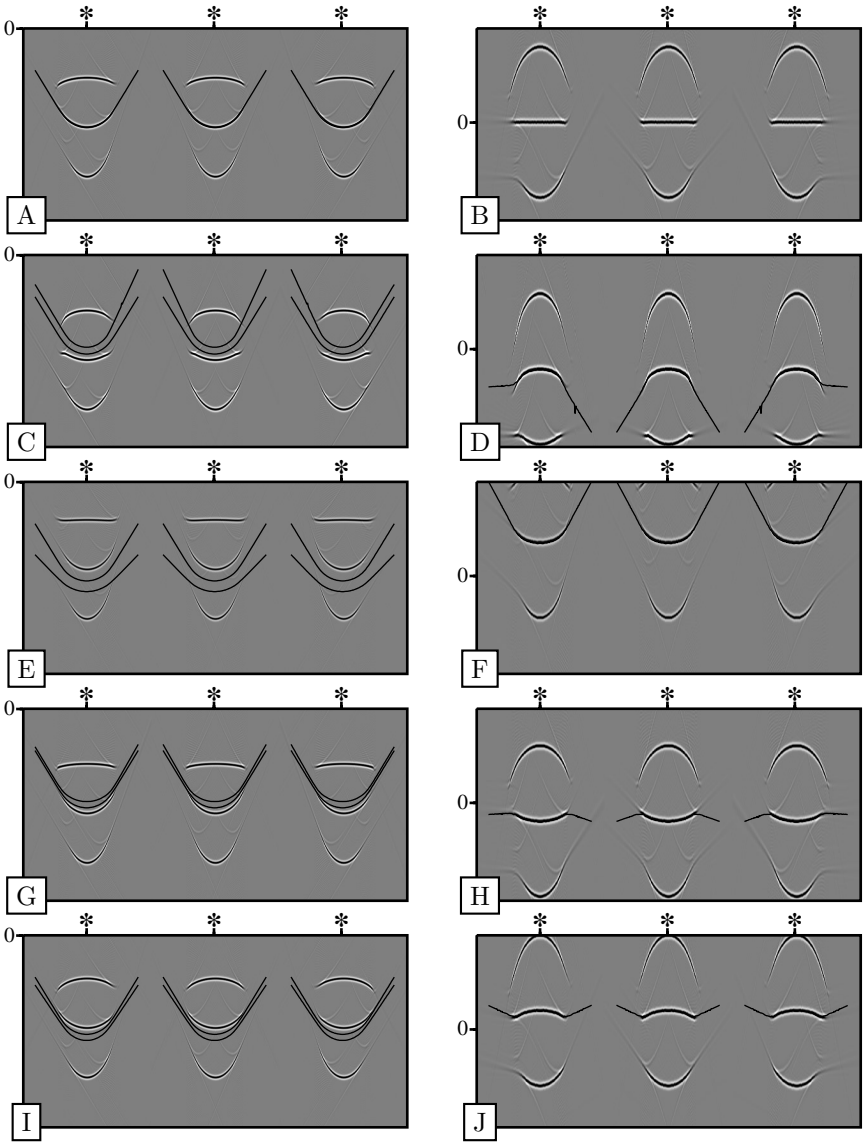

Figure 4.9: *Principle of equal traveltime in the linear Radon domain. The left column shows CFP gathers for three CFP's, overlain with their focusing operators. The right column shows the DTS gathers, overlain with 2 times the difference between the operators and the error free operators. (A,B) Correct focusing operators. (C,D) too high velocity used in operator modeling. (E,F) too low velocity used in operator modeling. (G,H) too deep depth used in operator modeling. (I,J) too shallow depth used in operator modeling.*

4.8C and Figure 4.8D show respectively the CFP-gathers and DTS-gathers when operators are used with a too high velocity. The operator used to construct the CFP gather and the correct focusing operator are displayed within the CFP gathers. It is remarkable to see that the correct operator seems to lie in between the erroneous operator and its CFP response. Thus, if the difference between the correct operator and the used operator is multiplied by two, we should obtain approximately the differential traveltimes that would be observed in the DTS gather for the specific event for which we estimated the focusing operator. This time curve, the difference between the correct operator and used operator times two, is displayed within the DTS gathers.

By picking the time difference and apply half the result to the used, erroneous operator, the new operator will be closer to the correct operator, without updating the model. From the DTS gather we can can conclude that the proposed update does not fully match the true update, therefore the updating has to be done in an iterated fashion. Figure 4.8E and 4.8F show comparable results with respect to the choice of a too low initial velocity, from which the same conclusion can be derived. If the initial depth is taken wrongly , which is depicted in figures 4.8G, 4.8H (too shallow) and figures 4.8I, 4.8I (too deep), we see that for a 1D medium, only one update step is needed if the error in depth is not too large or the aperture is relatively small. This was mathematically shown by Hegge (2001) under the assumption that the events show a parabolic move out. For arbitrary 2D and 3D media, the updating will behave iteratively as well.

Figure 4.9 shows the same results after transforming the data to the linear Radon domain, from which we draw the conclusion that in the Radon domain, updating in the situation of laterally invariant velocity medium can be carried out with one step (Thorbecke, 199j). In practice, however, the updating is carried out within the spatial domain. As the proposed algorithm depends on defining single time picks for each trace within the DTS gathers, transforming back and forth between the Radon and spatial domain introduces multiple pick definitions for each trace when the seismic responses become more complex, due to more complex earth models: The Radon transform of a traveltime functional is not straightforward and analytically defined.

■ 4.6.2 Fermat modeling

The second property derived from the WRW model that will be used advantageously is called to *Fermat modeling*, which is based on the principle of minimum or maximum traveltime. From a set of focusing operators describing the full lateral extent of a specific event, the measured surface response in terms of phase can be simulated

as well. Since

$$\mathbf{F}^{-}(z_d, z_0) = \left[\mathbf{W}^{+}(z_d, z_0)\right]^{*}, \tag{4.6.66}$$

$$\mathbf{F}^{+}(z_0, z_d) = \left[\mathbf{W}^{-}(z_0, z_d)\right]^{*}, \tag{4.6.67}$$

a measured wavefield $< \mathbf{P}(z_0, z_0) >$ can be synthesized by,

$$< \mathbf{P}(z_0, z_0) > = \left[\mathbf{F}^{+}(z_0, z_d)\right]^{*} \left[\mathbf{F}^{-}(z_d, z_0)\right]^{*} S(\omega), \tag{4.6.68}$$

$$= \mathbf{W}^{-}(z_0, z_d)\mathbf{W}^{+}(z_d, z_0)S(\omega). \tag{4.6.69}$$

When only locally reacting media are considered, thus no emphasis will be laid on the amplitudes, then the phase behavior of the data reconstruction of Equation 4.6.69 should coincide with the phase of the particular event described by the operators as recorded within the data,

$$\sphericalangle \left[\mathbf{F}^{+}(z_d, z_0)\mathbf{F}^{-}(z_0, z_d)S^{*}(\omega)\right]^{*} = \sphericalangle \left[\mathbf{W}^{-}(z_0, z_d)\mathbf{R}^{+}(z_d, z_d)\mathbf{W}^{+}(z_d, z_0)S(\omega)\right], \tag{4.6.70}$$

$$= \sphericalangle \left[\mathbf{P}^{-}(z_0, z_0)\right], \tag{4.6.71}$$

in which only one non-dispersive reflection event is considered only, assuring that the reflection function $\mathbf{R}^{+}$ is frequency independent. For convenience we will assume sources and receivers to be placed exactly on the same surface locations,

$$\mathbf{W}(z_d, z_0) = \mathbf{W}^{+}(z_d, z_0) = \left[\mathbf{W}^{-}(z_0, z_d)\right]^{T}. \tag{4.6.72}$$

Under the high frequency approximation, the phase of the focusing operator relates to the traveltime through the medium between a certain focal point and surface location. We can then approximate the operators by an amplitude term and a phase term

$$W(\mathbf{x}_1, \mathbf{x}_2) = A(\mathbf{x}_1, \mathbf{x}_2) \exp(-j\omega\tau(\mathbf{x}_1, \mathbf{x}_2)). \tag{4.6.73}$$

The synthesized data of Equation 4.6.69 can then be written in a continuous formulation[1] as,

$$P^{-}(\mathbf{x}_{src}, \mathbf{x}_{rcv}) = \int_{\mathbf{x}\in\partial\mathbb{D}_d} A(\mathbf{x}_{src}, \mathbf{x}) \exp(-j\omega\tau(\mathbf{x}_{src}, \mathbf{x})) \; A(\mathbf{x}, \mathbf{x}_{rcv}) \exp(-j\omega\tau(\mathbf{x}, \mathbf{x}_{rcv}))S(\omega), \tag{4.6.74}$$

$$= \int_{\mathbf{x}\in\mathbb{D}_d} A(\mathbf{x}_{src}, \mathbf{x})A(\mathbf{x}, \mathbf{x}_{rcv}) \; \exp(-j\omega(\tau(\mathbf{x}_{src}, \mathbf{x}) + \tau(\mathbf{x}, \mathbf{x}_{rcv})))S(\omega), \tag{4.6.75}$$

[1] Here we revert from the discrete formulation of the WRW-model back to a continuous formulation as the following determination of two-way traveltimes is based on stationary phase solution of an integral expression.

in which $\partial\mathbb{D}_d$ is the datum plane containing the CFP locations.

By using stationary phase analysis (Erdelyi, 1956; Bleistein and Handelsman, 1986), the most significant contribution to the integral of Equation 4.6.75 can be determined, defined by the stationary point $\mathbf{x}_s$, with respect to the CFP locations,

$$\frac{\partial\tau'(\mathbf{x}_{src},\mathbf{x}_{rcv},\mathbf{x}_s=\mathbf{x})}{\partial\mathbf{x}}=0, \qquad \mathbf{x}\in\partial\mathbb{D}_d, \tag{4.6.76}$$

in which,

$$\tau'(\mathbf{x}_{src},\mathbf{x}_{rcv},\mathbf{x})=\tau(\mathbf{x}_{src},\mathbf{x})+\tau(\mathbf{x},\mathbf{x}_{rcv}). \tag{4.6.77}$$

Then the integral may be approximated by

$$\begin{aligned}P(\mathbf{x}_{src},\mathbf{x}_{rcv}) =&\sqrt{\frac{2\pi}{\omega\frac{\partial^2\tau}{\partial x^2}}}A_1(\mathbf{x}_{src},\mathbf{x}_s)A_2(\mathbf{x}_{rcv},\mathbf{x}_s)S(\omega)\\ &\exp(-j\omega(\tau(\mathbf{x}_{src},\mathbf{x}_s)+\tau(\mathbf{x}_{rcv},\mathbf{x}_s)-j\mu\frac{\pi}{4}),\end{aligned} \tag{4.6.78}$$

with

$$\mu=sign\left[\omega\frac{\partial^2\tau}{\partial x^2}\right]. \tag{4.6.79}$$

According to the stationary phase of Equation 4.6.76,

$$\nabla(\tau(\mathbf{x}_{xsrc},\mathbf{x})+\tau(\mathbf{x}_{rcv},\mathbf{x}))|_{\mathbf{x}=\mathbf{x}_s}=0, \tag{4.6.80}$$

$$\nabla\tau(\mathbf{x}_{xsrc},\mathbf{x})|_{\mathbf{x}=\mathbf{x}_s}=-\nabla\tau(\mathbf{x}_{rcv},\mathbf{x})|_{\mathbf{x}=\mathbf{x}_s}, \tag{4.6.81}$$

which selects the specular ray as schematically displayed by Figure 4.10. Then the approximate solution to the integral of Equation 4.6.75 becomes,

$$P(\mathbf{x}_{src},\mathbf{x}_{rcv})=A(\mathbf{x}_{src},\mathbf{x}_s)A(\mathbf{x}_{rcv},\mathbf{x}_s)\exp\left(-j\omega(\tau(\mathbf{x}_{src},\mathbf{x}_s)+\tau(\mathbf{x}_{rcv},\mathbf{x}_s))\right)S(\omega). \tag{4.6.82}$$

Note, that Equation 4.6.80 resembles Fermat's principle, stating that the traveltime between source and receiver should be the traveltime along the shortest path in time. Furthermore, Equation 4.6.81 resembles Snell's law, stating that the angle of incidence should equal the angle of reflection. Thus, the two-way traveltimes can be constructed from the traveltime table of the forward propagator, by combining the two columns representing the considered source and receiver location and picking the minimum traveltime (Fermat's principle) or by summing the traveltimes found at the locations within the two selected columns for which the derivatives are equal in amplitude but opposite in sign (Snell's law). The particular focuspoint, for which the stationary phase analysis renders the two-way traveltime, is also the reflection point of the specular ray between source and receiver. For the case we have an set of erroneous focusing operators,

$$<\mathbf{W}^-(z_0,z_d)>=\mathbf{W}^-(z_0,z_d)\Delta\mathbf{W}, \tag{4.6.83}$$

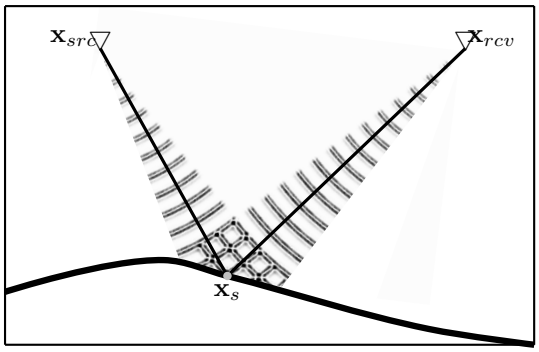

Figure 4.10: *The two-way traveltime can be determined by finding the stationary point* $\mathbf{x}_s$ *with respect to the traveltime of the combined operators originating from a specific source and receiver pair* $(\mathbf{x}_{src}, \mathbf{x}_{rcv})$.

the synthesized data reads,

$$\mathbf{P}^-(z_0, z_0) = \mathbf{W}(z_0, z_d)\Delta\mathbf{W}\Delta\mathbf{W}^T \left[\mathbf{W}(z_0, z_d)\right]^T S(\omega), \tag{4.6.84}$$

under the assumption $\mathbf{W}^+(z_d, z_0) = \left[\mathbf{W}^-(z_0, z_d)\right]^T = \left[\mathbf{W}(z_0, z_d)\right]^T$. Clearly an error in traveltime (the phase characteristics of the complex valued $\mathbf{W}$) would render a doubled error within reconstructed data, compared to the real data. However, as long as the error is real and unitary with orthogonal rows and columns,

$$\Delta\mathbf{W}\Delta\mathbf{W}^T = \mathbf{I}, \tag{4.6.85}$$

the error will not be detected if the synthesized data is compared to the real data.

Besides the two criteria on the system errors derived through the principle of equal traveltime and Fermat modeling, the operators should also be orthogonal,

$$\mathbf{I} = \mathbf{F}^-\mathbf{W}^-, \tag{4.6.86}$$

Any error in the forward operator,

$$< \mathbf{W}^- >= \mathbf{W}\Delta\mathbf{W}, \tag{4.6.87}$$

results in

$$\Delta\mathbf{W}^H\mathbf{W}^H\mathbf{W}\Delta\mathbf{W} = \mathbf{I}. \tag{4.6.88}$$

With the true operator, $\mathbf{W}$, being orthogonal it is easily concluded that an error in the focusing operators should also satisfy,

$$\Delta\mathbf{W}^H\Delta\mathbf{W} = \mathbf{I}. \tag{4.6.89}$$

Combining all the criteria on the system errors leads to the conclusion that if an error $\Delta\mathbf{W}$ in the focusing operators satisfies,

$$\Delta\mathbf{W}^T = \Delta\mathbf{W}^H = \Delta\mathbf{W}^{-1}, \tag{4.6.90}$$

it will not directly be detected.

The Fermat modeling method is illustrated in Figure 4.11. Figure 4.11A displays a full lateral set of focusing operators, containing time-undulations due to near surface anomalies. The operators were sorted to their common surface point (CSP) location and used within a minimum phase analysis to construct the two-way traveltimes for all offsets related to the reflector from which the focusing operators originated. In Figure 4.11B the two-way traveltimes are displayed. Along the line $n_{cfp} = n_{csp}$ the zero-offset times are recognized, which clearly show the impact of the near surface.

In practice, the propagators are stored as discrete matrices and will contain a noise term. Then the derivatives are determined through finite differencing the columns. As derivatives are extremely sensitive to noise, picking the minimum traveltime is preferred. However, in the next chapter initial operators will be constructed through closed form expressions, in which it is more convenient to use the derivatives.

4.7 Synthetic Example

To illustrate the outlined methodologies, data was modeled using finite difference modeling on the model shown in Figure 4.12A. A laterally varying low velocity overburden has a gradually changing boundary. The shot record in Figure 4.12 Bclearly visualizes the statics problem. In total 301 shots with 301 receivers each were modeled. By using conventional statics methods as the PLUS-MINUS method and residual statics based on maximum stack power energy (see Chapter 2), a corrected stack was derived as displayed in Figure 4.12C. The overburden was replaced by a layer with an acoustic velocity of 1900 m/s, estimated by the static procedures.

Clearly the conventional methods show some short comings for this example. The replacement velocity was estimated a little bit too low: 1900 m/s versus 2000 m/s. The replacement velocity in combination with the estimated and static free two-way traveltime results in an estimated reflector-depth of about 420 m, which is 20 meters too deep. Around $x = 200m$ the stack shows a dim spot and the stack seems to show structure (the reflector shows "smiling"). Within the next section we will see how common focuspoint processing is performing on this example.

■ 4.7.1 Operator updating

A simplified model (Figure 4.13A) is used in this example to determine initial operators in a first attempt to describe the propagation effects to focus points defined on

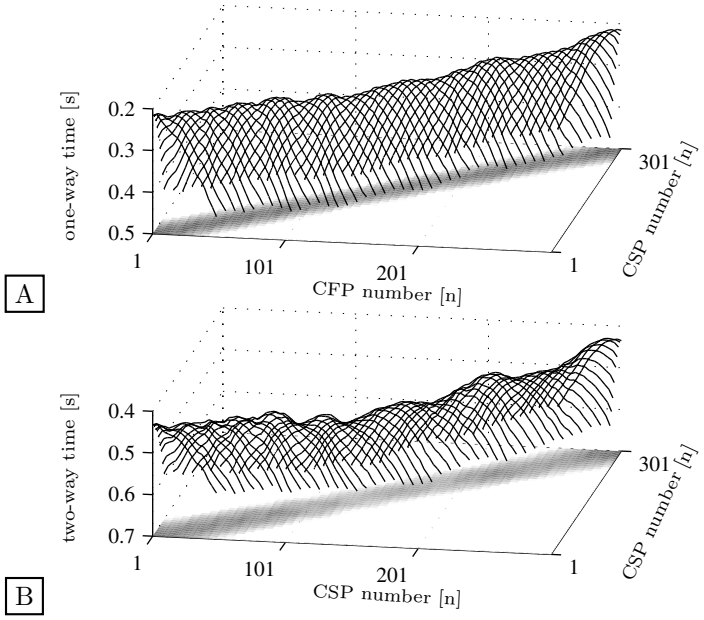

Figure 4.11: *By reordering the focusing operators of A) to CSP operators, which is possible when a full lateral set of operators is available, the two-way traveltimes for all shot-receiver experiments can be reconstructed through a stationary phase analysis as displayed in B).*

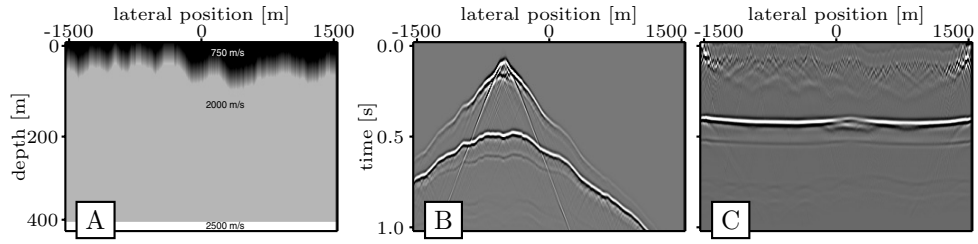

Figure 4.12: *Complex near surface example. B) displays a shot-record modeled within the model of A), which contains an undulating and non reflecting near surface transition and a reflector at 400 m depth. After application of refraction and residual statics, the time-image of the reflector at 400 m depth shows apparent structure and dim spots as displayed in C).*

the reflector at 400 m depth (Figure 4.13B). Here we used a laterally homogeneous model. In practice one could determine initial operators from the stack and stacking velocities, even in combination with conventional static solutions. An initial Common Focus Point gather is constructed (Figure 4.13C) with the aid of the modeled initial operators. When the CFP gather and the operator are correlated, clearly they do not align around $t = 0$ in the initial DTS panel (Figure 4.13D). After updating

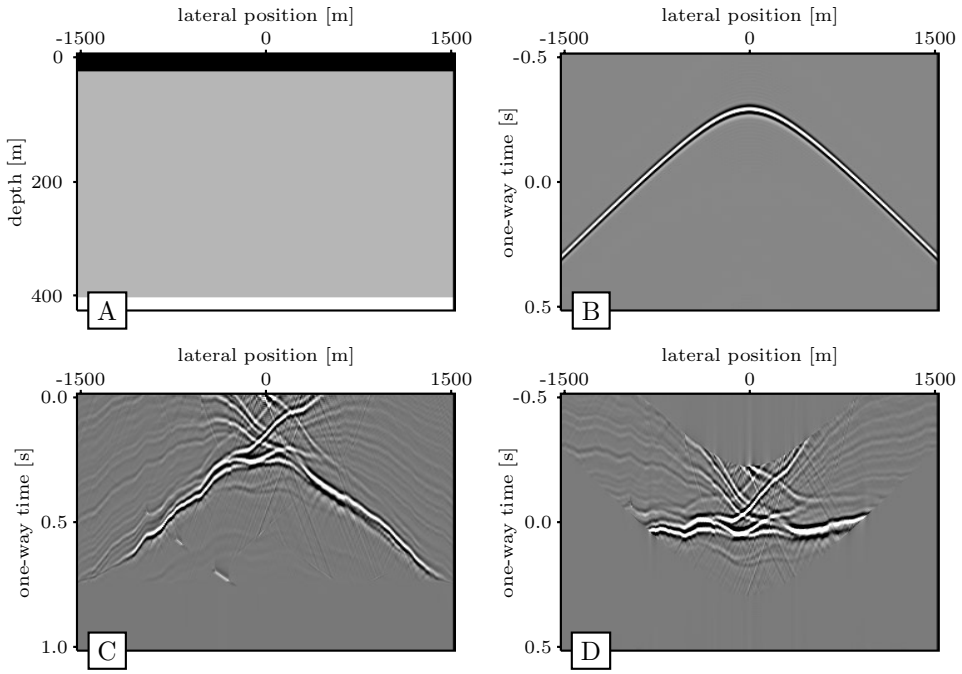

Figure 4.13: *An initial operator (B) is modeled in a simple initial model (A). The corresponding erroneous CFP gather (C) clearly does not coincide with the operator as can be seen in the DTS gather (D).*

the operator using the principle of equal traveltime, through picking times in the DTS panel, a final operator is derived after a few iterations, displayed in Figure 4.14A. When correlating the final operator with its constructed CFP gather, the DTS panel does align around $t = 0$. The principal of equal traveltime is satisfied by this condition, as can be seen in Figure 4.14B. This procedure was repeated for all focal points along the boundary. The final operators found can be used for further analysis and processing.

■ 4.7.2 Analysis and processing of final operators

The final operators as found in the previous section were determined by using the principle of equal traveltime only. In this section it is demonstrated how these operators can be used to construct time images and to redatum the recorded data. As only the principle of equal traveltime is used an analysis relating to the other two principles will be made as well. First the final operators will be used to construct

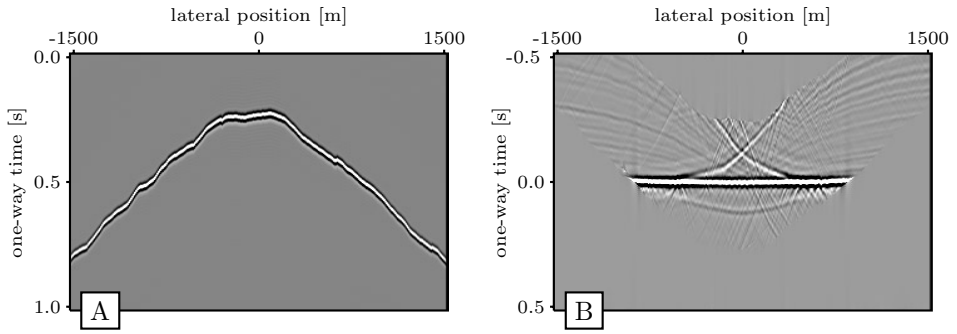

Figure 4.14: *After updating an operator (A) is found with gives good alignment in the DTS gather around t=0 (B), when correlated with its corresponding CFP gather.*

a CFP time-image. A CFP time-image is constructed by stacking the DTS panels, rendering the imaged energy, and placing the results at the apex-times of the used operators. The image of Figure 4.15A shows a time-image constructed with the true operators, while the image of Figure 4.15B shows a time-image constructed with the estimated operators. The result looks quite promising, at the right hand side the estimated image looks a little bit different. This is partly due to the error in the operators as well as drift along the lateral coordinate, which will be explained further on. However, when compared to the stack obtained using conventional statics as displayed by Figure 4.12C, a large improvement is observed. No significant false structure is introduced by the estimated operators, compared to the image obtained by using the true operators. Furthermore, dim spots are not observed. It should be noted here that the near surface is not replaced by a constant velocity layer as is done within the conventional procedures, therefore the image directly shows the structure of the near surface anomalies when compared to the image of Figure 4.12C.

In the next step, the estimated operators were used to redatum the data, by applying the operators twice, according to Equation 3.4.22. To examine the results after redatuming an anticline was modeled underneath the reflector at $400m$ depth. The image of Figure 4.16A shows a redatumed shot gather after redatuming with the estimated operators. The image of Figure 4.16B shows the same result, after redatuming with the true operators. In comparison we see that both apices of the two events have drifted. Furthermore, the first arrival shows a drift as well. The drift is introduced due to the fact that during updating the control over the assumed lateral locations of the CFP operators is lost. Within the analysis made by Hegge (2001), an analytical result is derived for the lateral drift of a focus point during updating along a dipping reflector. For other, more complex models, no analytical solutions can be derived. In Volker (2002) an assessment is made of the spatial res-

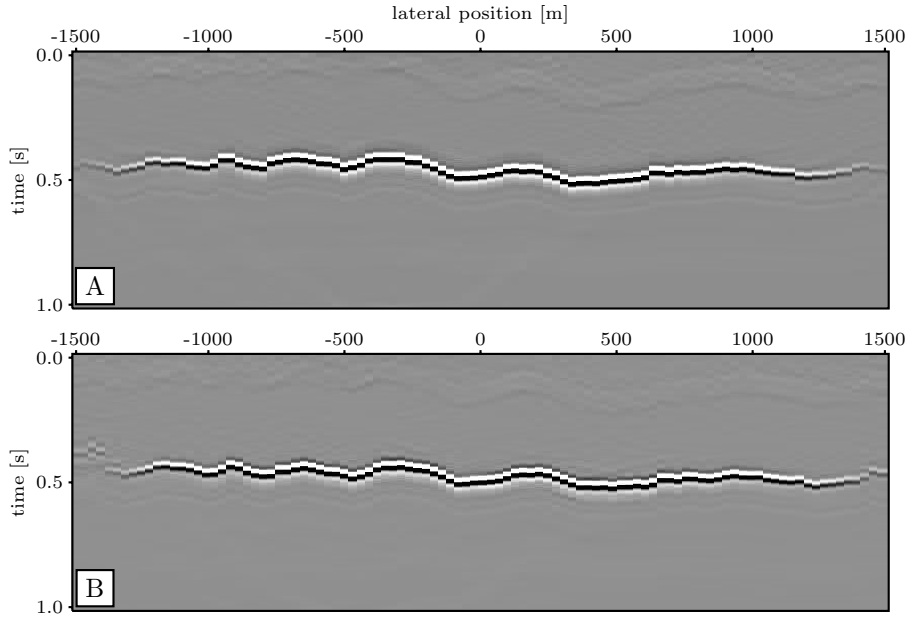

Figure 4.15: *The final operators are used to construct a CFP time image. A) The image when constructed with the correct operators. B) The image constructed with the estimated operators. Note the small differences.*

olution that can be obtained through migration as function of the spatial sampling of sources and receivers. If the redatumed result using updated operators is taken as input for migration, without a proper control of the lateral drift of the focusing operators, this will degrade the resolution. Therefore, this aspect in operator updating needs to be addressed. One has to put the operators subject to a tomographic inversion to estimate the exact locations. This would result in redatumed data which is unequally sampled in space for which one could correct by using regularization tools such as described in Hindriks and Duijndam (2000).

■ 4.7.3 Focusing analysis

In Figure 4.16 not only drift of CFP locations is observed but some spurious energy as well. Although the principle of equal traveltime provides an elegant way in describing a data-driven updating of focusing operators, it is not intrinsically assured whether the correct operator will actually be found as shown in Section 3.5. Already from the principle of equal traveltime, it can be shown that a null space exists during operator construction for the space of operators that will focus the data. When one

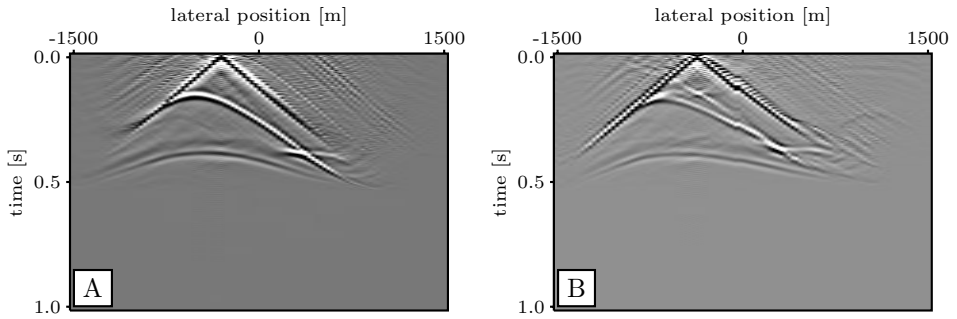

Figure 4.16: *The focusing operators can be used to perform a redatuming to the focused reflector. Figure A) shows a correctly redatumed shot record. Figure B) shows a redatumed shot gather when estimated operators are used. The differences are caused mainly by drifting of CFP locations. This can be solved by using a regularization procedure.*

focusing operator is constructed from two rows of the inverse propagator,

$$\vec{\Gamma}(z_0, z_d) = \left[\mathbf{W}^-(z_0, z_d)\right]^* \mathbf{I}_i + \left[\mathbf{W}^-(z_0, z_d)\right]^* \mathbf{I}_j, \tag{4.7.91}$$

$$\vec{\Gamma}(z_0, z_d) = \left[\mathbf{W}^-(z_0, z_d)\right]^* \left[\mathbf{I}_i + \mathbf{I}_j\right], \tag{4.7.92}$$

and is applied on the data, the result can be written in two ways,

$$\mathbf{P}(z_0, z_0)\vec{\Gamma}(z_0, z_d) = \mathbf{W}^-(z_0, z_d)\mathbf{R}(z_d, z_d)\mathbf{W}^+(z_d, z_0)\left[\mathbf{W}^-(z_0, z_d)\right]^* \left[\mathbf{I}_i + \mathbf{I}_j\right], \tag{4.7.93}$$

$$= \mathbf{W}^-(z_0, z_d)\mathbf{R}(z_d, z_d)\mathbf{I}_i + \mathbf{W}^-(z_0, z_d)\mathbf{R}(z_d, z_d)\mathbf{I}_j, \tag{4.7.94}$$

which both obey the principle of equal traveltime. As the focusing result of Equation 4.7.94 coincides in time with the reversed time of the applied operator of Equation 4.7.91, any linear combination of rows of the inverse propagator obeys to the principle of equal traveltime, which leads to an infinite large null-space(Tarantola, 1987). However, such a combined operator does not satisfy orthogonality with the remaining set of operators, which can be verified through the focusing principle. But, since no data dependency is available within the focusing principle, a route to a proper operator in terms of focusing into a single focus point can not directly be provided. In Bolte an additional non-uniqueness property is described, showing that next to linear combinations of full operators, also new operators can be constructed from subsets of full operators, which still obey the principle of equal traveltime.

To check to which extent the operators were correctly determined, the operators were subject to a pre-stack Kirchoff depth migration using the exact model. This procedure is intrinsically equivalent to the focusing principle, in the sense that if an

operator is not a focusing operator it will render energy at other locations within the model than the aimed focuspoint. Figure 4.17A shows the image after migrating correctly modeled operators. Figure 4.17B shows the image after migration of estimated operators. The correct operators clearly collapse to the correct location. Overall, the estimated operators are imaged quit well. However, the locations to which the operators image are a little bit smeared, show spurious events and drifted laterally, which was observed by Bolte et al. (1999) as well. Furthermore, on the right hand side of the model, the CFP locations show a large structure. Close examination of the operators learned that the corresponding events in the DTS gathers still had a constant misfit from $t = 0$ for which we should have include some additional updating iterations. From this little example it can be concluded that the locations of the focal points drifted away from the aimed locations and the migrated result shows that the operators are not perfectly focusing.

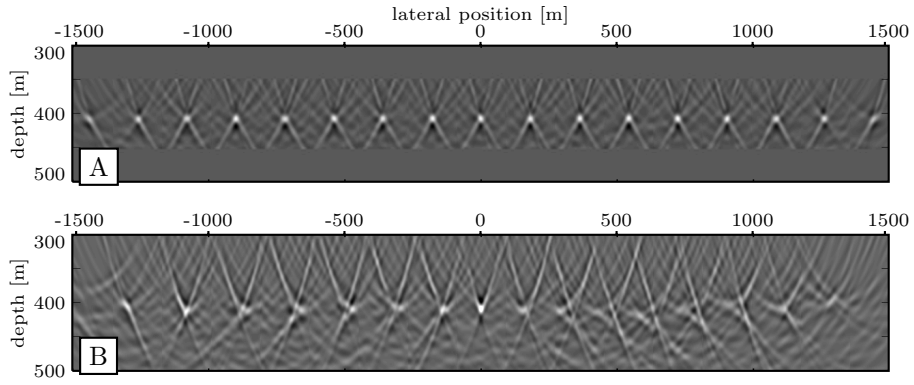

Figure 4.17: *To check the operators found after single-point updating, the subset of the operators was depth-migrated using the correct model. Figure A) shows the migrated result of the correct operators clearly collapsing into the focus points. Figure B) shows the migrated estimated operators. Through the updating process smearing and drifting of CFP locations was introduced.*

For a more detailed analysis of the estimated operators, a full set of operators, Figure 4.18A, was modeled in the true near surface model with an interval of 2.5m in between the target focuspoints. Thus a very dense set of operators was obtained. The updated operators, Figure 4.18B were correlated, one by one, with the full set of densely sampled operators. By picking for each updated operator, the operator from the dense correct set, which gave the minimum rms-error, both the fitness of each operator is obtained, indicating how well the operator focuses to a single point, as well as the drift from the assumed CFP location. In Figure 4.18C, the blue

line connecting the dots denotes the rms error for each operator. The closed points within the graph of Figure 4.18C denote the aimed lateral locations (locations that were aimed at), the open points denote the actual locations after updating. Thus, we see that the operators are not perfectly focusing and their locations are drifting. Note that the rms-error for the focusing operator around $x = 0$ has the lowest value, which is confirmed by the clear image at the same location in Figure 4.15A and also in Figure 4.17B. Also, most of the drift is found at the right-hand side of the model, which confirms the observations in Figure 4.16 and Figure 4.17.

4.8 Conclusions

The CFP method provides an elegant way of determining one-way propagation operators from the data. A full set of operators provides a full inverse propagator that can be used for redatuming the recorded data through the complex near surface. In practice, the updating procedure for a single point may not result in a unique solution.

In the next chapter we show that this problem can be solved by updating a full range of focus points simultaneously.

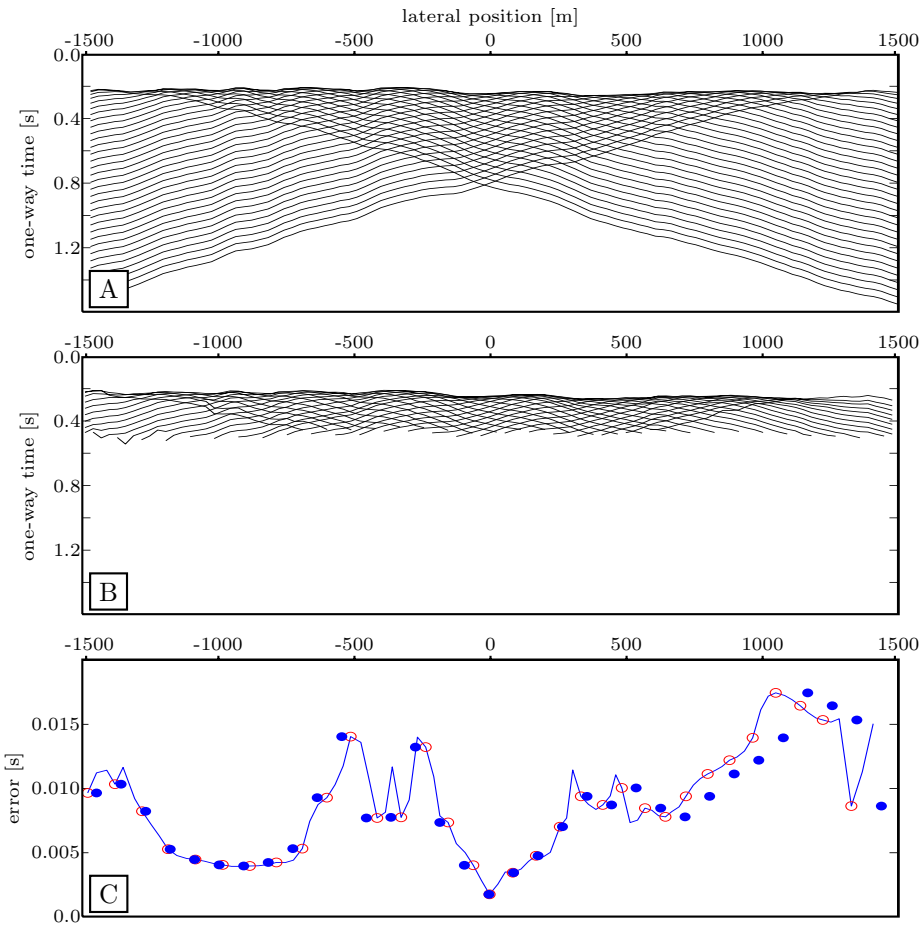

Figure 4.18: *Within the single point updating procedure, the estimated lateral locations of the focuspoints show a shift. Figure A) shows updated operator-times, figure B) shows modeled operator-times and figure C) the correlation between updated operator-times and operator-times modeled at a 1m interval distance. The line displays the rms time error, the open dots denote the estimated locations, while the closed dots denote the actual CFP locations.*

5

Intialization of focusing operators

5.1 Introduction

In the previous chapters the general methodology has been outlined to correct for propagation through the subsurface by means of focusing operators. In practice, however, it is quite often a cumbersome and non-trivial task to directly implement the outlined procedures in case of a complex near surface. As shown in Chapter 2, the effect of a complex near-surface, traditionally called statics, reveals itself in distortion of the continuity of the seismic signal and associated coherent stacking of energy during imaging. Within this chapter the principle behind coherent energy stacking within the CFP methodology is demonstrated as well as the effects a complex near surface will have on this procedure. Furthermore, a method is outlined that will capture these near surface effects as good as possible within the construction of an initial set of focusing operators. As a result a convergent updating is assured to further improve on the focusing operators through the principle of operator updating.

5.2 Fresnel zone reconstruction in case of a complex near surface

The construction of a Common Focus Point-gather consists of an integral evaluation, or stacking procedure; focusing in detection describes an integral evaluation over the receiver locations, while focusing in emission describes an integral evaluation over the source locations. Such integrals evaluate the energy resulting from the interference pattern of two convolved wave-fields, the source-gather and the focusing operator, which resembles an anti-causal Green's wavefield excited by a secondary source (see Chapter 4, Section 3). The main contribution in these integrals arise from the first Fresnel zone, where the data and the operator are tangent.

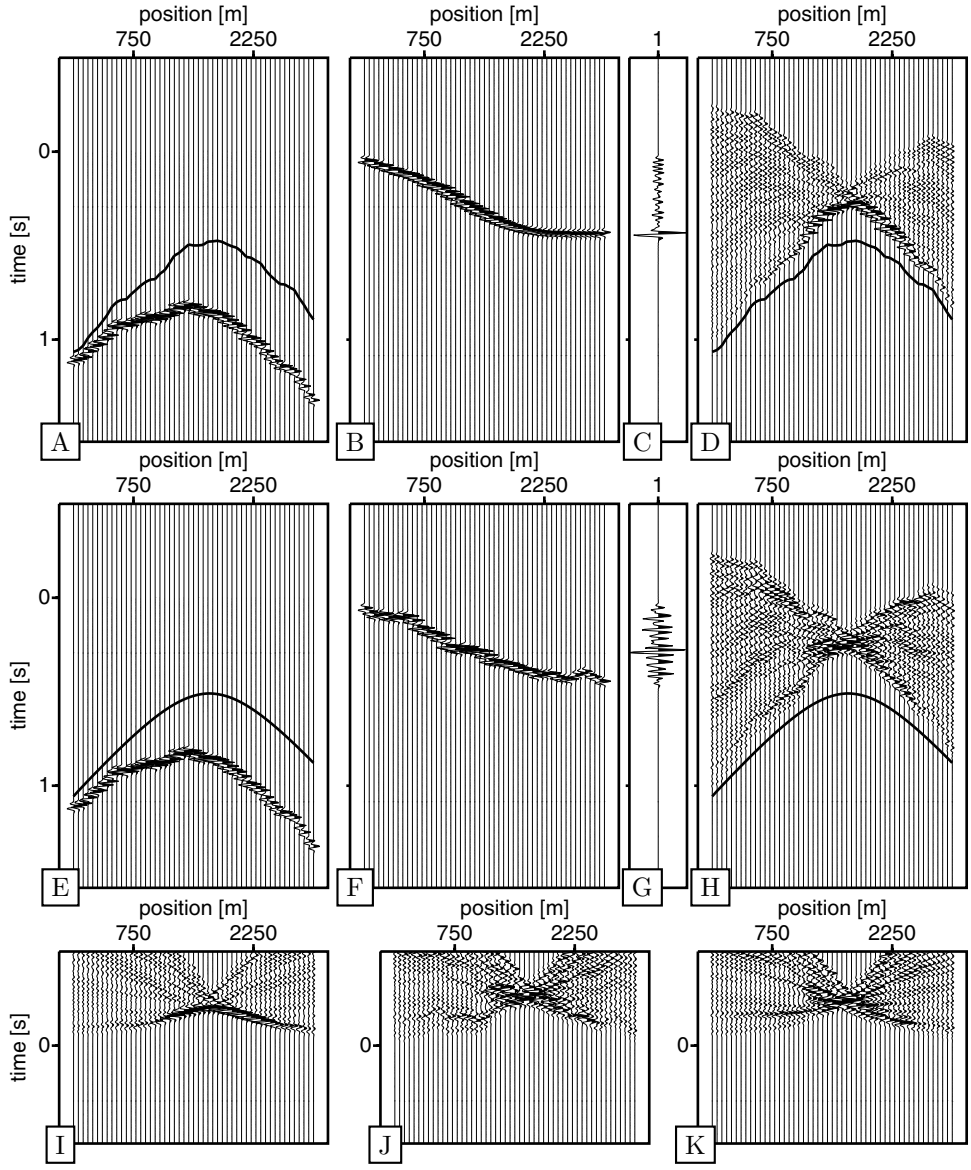

Figure 5.1: *Construction of a CFP-gather. Each shot-gather (A,E) is corrected in time by the CFP operator (B,F, operators are denoted by the curve in A and E) and stacked (C,G) resulting in one trace of the CFP-gather (D,H). The displayed shot-gathers are contaminated with near surface effects. If near surface effects are not implemented in the operator, stacking of the corrected shot-gather introduces artifacts (G), hampering the analysis of the DTS-gather (J). By using initial static corrections in the operator (A), the corrected shot-gathers (B) produce a cleaner stack (C), resulting in better defined DTS-gathers (I). Even if the erroneous CFP gather of (H) is corrected with the operator containing initial statics of (A), the DTS-gather of (K) has not significantly improved.*

Figure 5.1 displays in detail the principles behind the construction of a CFP-gather, in a graphical sense. In Figure 5.1A, a shot-gather is shown with the recorded reflection of one event within a model with a serious near surface problem, clearly expressed by the undulating nature of the recorded event. The focusing operator, for which the CFP-gather will be constructed, is denoted within the shot-gather by its traveltime curve. The initial focusing operator is modeled from an incorrect reflection depth but it does capture the propagation effects of the complex near surface. The next step in computing the CFP-gather, is shifting the traces of the shot-gather in time corresponding to the traveltimes of the focusing operator. This procedure resembles the correlation of the shot-gather with the operator. Figure 5.1B shows the correlated result. It is clearly observed that the statics within the initial operator remove the fast undulations arising from the near surface within the shot-record, leaving a smooth synthesized shot-record. The stacked result will be dominated by the contribution of the single stationary point within the corrected shot-record, resulting, apart from some noise, in a clear single peak as observed in Figure 5.1C. By repeating the procedure for all shot-records, the CFP-gather of Figure 5.1D is obtained, containing as expected a clear single event. The used reversed time operator overlies the CFP-gather. Due to the use of an erroneous depth for the initial operator, the CFP-gather and operator do not coincide, but the effects of the near surface clearly do coincide as the CFP-gather now contains the same near surface effects at the source side.

If, however, the initial operator is modeled using an average (erroneous) velocity without any attempt of incorporating fast variations arising from the near surface (Figure 5.1E), the corrected shot-gather still contains the undulations caused by the near surface as can be observed in Figure 5.1F. Due to the remaining undulations, which form multiple local stationary points, the resulting stack does not contain a clear single event as shown in Figure 5.1G. The final CFP-gather is therefore contaminated with artifacts and the single reflecting event is hardly recognized as displayed in Figure 5.1H.

To assure correct updating, the DTS-gathers need to be examined to define traveltime updates for the CFP operators. Using the operator containing initial static corrections clearly leads to an easy interpretable DTS-gather as depicted in Figure 5.1I. However, the usage of a smooth operator does not lead to a clearly interpretable event, as can be seen in Figure 5.1J. Figure 5.1K displays the DTS-gather constructed by correlating the operator containing the near surface effects with the CFP-gather that was constructed by using the smooth operator that didn't contain any near surface effects. Although near surface effects seem to be described reasonably well in the CFP-gathers for the far offset (the event is fairly well detectable), within the near offset section hardly any improvement is recognized when compared to the result of Figure 5.1J. Apparently, trying to detect the near surface effect within the DTS-gathers is a non-trivial task and since the near offsets eventually will have the largest contribution in focusing and imaging, such a workflow will even-

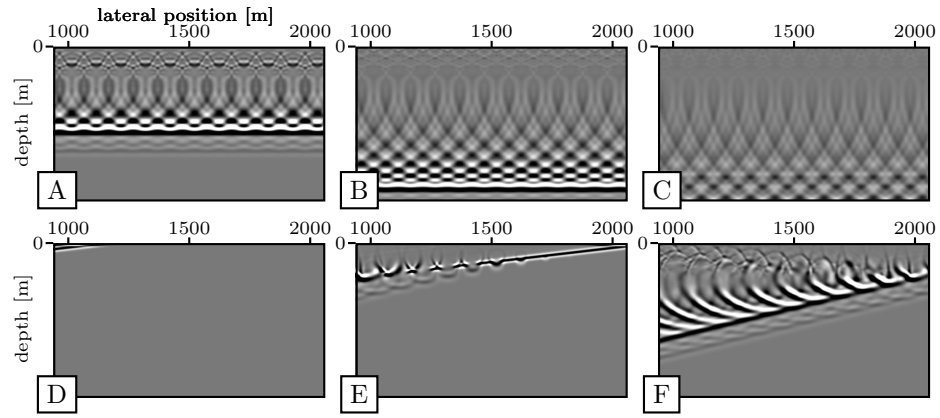

Figure 5.2: *The evolution (from left to right panels) of a vertically traveling plane-wave (A,B,C) and a dipping plan-wave (D,E,F), traveling through a low velocity sinusoidal shaped near surface (wavelength=100m). The first arrival wavefront is slow healing from the near surface effects.*

tually lead to a strongly distorted and contaminated image.

In the remainder of this chapter first an evaluation of the nature of near surface anomalies an their impact on wave-propagation is given, followed by two means of identifying zero-offset times and initial stacking velocities. These zero-offset times have a close relation to conventional surface consistent static solutions. Next, the found zero-offset times and stacking velocities are perturbed within a non-linear optimization scheme such that optimized initial operators are obtained through Fermat modeling, as explained in Chapter 5, and capture the near surface effects within the initial operators as good as possible.

5.3 Near surface impact on wave propagation

Near surface anomalies mainly affect the seismic wavefield close to the sources and receivers. However, its impact is visible on reflections from the complete subsurface. As within the WRW-formulation two-way seismic wave-fields are formulated in terms of convolving one-way propagators, it makes sense to investigate the near surface effects on the one-way propagators under consideration.

The wavefield $\mathbf{W}^+$, emitted by a source at the surface will propagate down, gets reflected at the deeper interfaces and propagates back through the propagator $\mathbf{W}^-$ to the receivers. Thus, the measured wavefield travels through the near surface two times. However, the near surface has a different effect on the separate one-way propagators. The downward propagator from the source to the reflector recovers

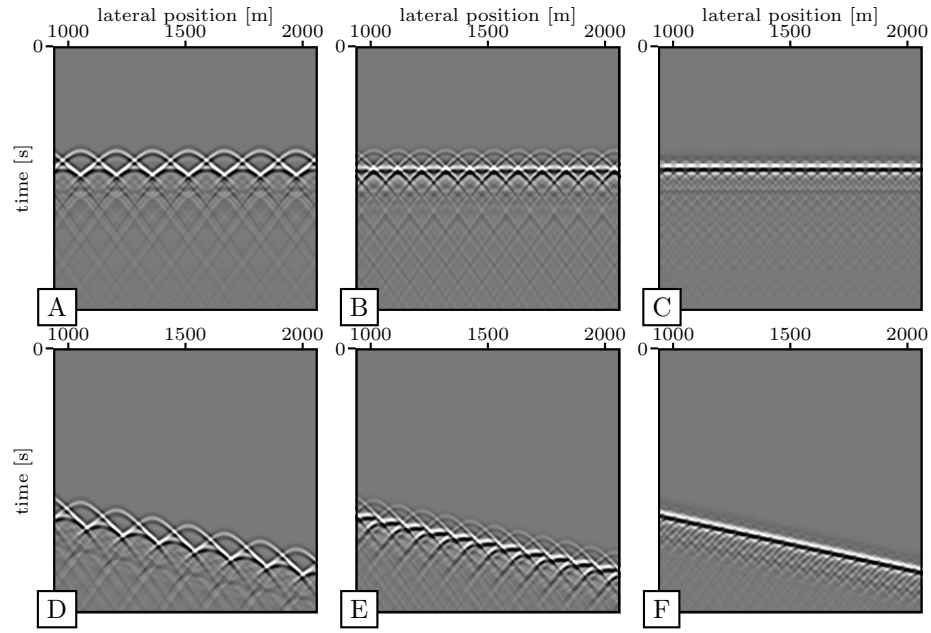

Figure 5.3: *Plane waves recorded at a level 500 m below a sinusoidal shaped low velocity near surface through which the plane waves have traveled. The wavelengths of the near surface are 200 m (A,D), 100 m (B,E) and 50 m (C,F).*

in the far field (near the reflector) from the near surface impact, whereas the upward propagator recorded at the receivers is affected by the near surface close to the receivers and has, therefore, not yet recovered from the near surface impact. In Aki and Richards (1980) a theoretical derivation is given to describe the far field resulting from inhomogeneous and periodic near field scattering. Here, we will suffice describing the healing phenomena by visual inspection.

Figure 5.2 displays for two plane-waves the propagation effects through a sinusoidal near surface (close to depth zero) interface between a medium with a relative slow wave-velocity and a medium with a relative high wave-velocity. Clearly a difference in traveltime is developing along the wavefield. Further downward in the far field still different modes are recognized, however, away from the near surface a plane-wave envelope is clearly recognized as a first arrival. Figure 5.3 shows the recording of the same two plane waves as a function of time and different wave-lengths of the near surface inhomogeneities, measured at a fixed distance away from the near surface. For a smaller wavelength of the sinusoidal near-surface inhomogeneity, the plane wave nicely heals from the near surface effects.

From these experiments it can be concluded that the different plane wave modes have a tendency to heal from the near surface inhomogeneities depending on the distance from and the wavelength of the near surface inhomogeneities: The larger the wavelength, the further away sufficient healing will take place with the notion that for wavelengths much larger compared to the wavelength of the seismic signal, the near surface will no longer cause a problem within the traveltime updating procedure.

Figure 5.4 illustrates how on the one hand the near surface impacts the up-going wavefield, whereas on the other hand the down-going wavefield will heal from the near surface, through a combination of snap shots of the wavefield originating from a point source respectively placed away from the near surface (Figure 5.4A) and placed within the near surface (Figure 5.4B). The measured wavefield within the near surface clearly shows the imprint of the near surface in terms of traveltime. The wavefield recorded within the far field with respect to the near surface inhomogeneities has healed from the near surface effects in terms of traveltime undulations although still different modes are recognized by the secondary arrivals. However, these modes will not hamper the tracking of the first arrival. The result of Figure 5.4C can be regarded as the physical measurement of a focusing operator, whereas the result of Figure 5.4D can be regarded as the physical measurement of a so called common surface point (CSP) operator: an operator describing the propagation characteristics from a single surface location toward all the considered common focus point locations. Note that for a full consistent set of focusing operators, $\mathbf{F} = \mathbf{W}^{+,*}$, the focusing operators are defined by the rows of the propagation matrix $\mathbf{W}^{+}$, see also Section 3.6.1. The CSP operators are defined by the columns of the same propagation matrix. Therefore we are able to construct smooth CSP operators from a consistent set of non-smooth focusing operators sufficing from a complex near surface and vice versa.

The aim of the remainder of this chapter is to define a proper definition of the smoothness of the CSP operators. Note, however, that this description does not need to be complete in the sense of describing the near surface characteristics exactly. It merely serves as a method to capture the trend of near surface irregularities such that we are able to constructively update the CFP operators.

■ 5.3.1 Domain of smallest operator complexity

The concept of wavefield healing and wavefield destruction can be a confusing phenomena, easily interpreted as violating reciprocity. This conclusion was drawn on a similar observation in salt modeling by Muerdter, Kelly and Ratcliff (The Leading Edge, july 2001). However, as indicated by the reply of Wapenaar and Fokkema, it should be noted that reciprocity relates a point receiver and a point source in 3D (line receivers in 2D), which is still obeyed. Also in the experiment in Figure 5.4 reciprocity holds, as the middle trace in Figure 5.4C and Figure 5.4D are identical,

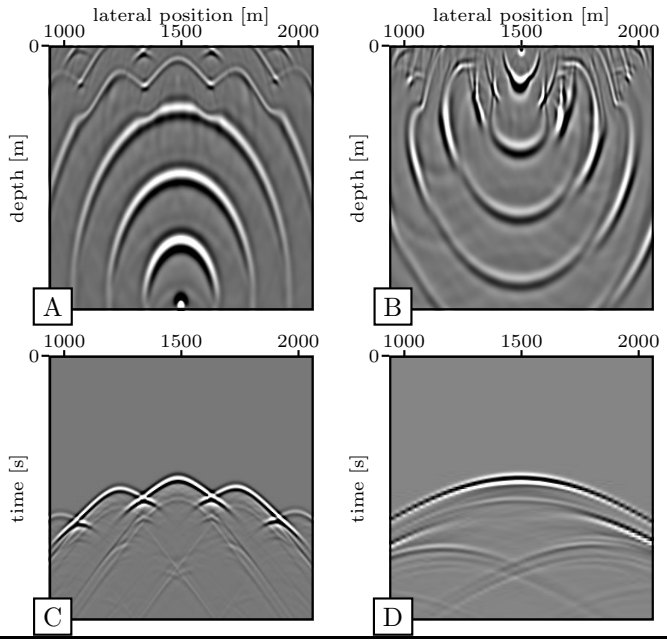

Figure 5.4: *A) displays the evolution in time of a pressure wavefield excited by a point-source at ($x = 1500, z = 600$). C) displays the signal recorded in time by a line of receivers place at $z = 0$. B) displays a pressure wavefield excited by a point-source at ($x = 1500, z = 0$). D) displays the signal recorded in time by a line of receivers placed at $z = 600$.*

whereas source and receiver are interchanged. Actually, reciprocity will be used advantageously, to determine initial operators describing the near surface phase effects approximately, through a sparse set of parameters.

For that matter consider the following experiment. Within a homogeneous background model a low velocity anomaly is placed respectively near the surface and near the target reflector. A third experiment contains a high velocity anomaly near the surface. For the three sketched situations traveltime operators are modeled from all surface locations to a laterally dense set of predefined focus point locations at the target reflector. Next, the computed traveltimes are corrected by the traveltimes through the homogeneous background without the presence of the inhomogeneities. In Figure 5.5 the corrected residual traveltimes are displayed as function of surface location and reflector location. Clearly, an inhomogeneity close to the surface has only an impact on the common surface location for the whole suite of operators, while an inhomogeneity close to the reflector has impact only on the common reflector location of the whole suite of operators. Since we assume that both sources and

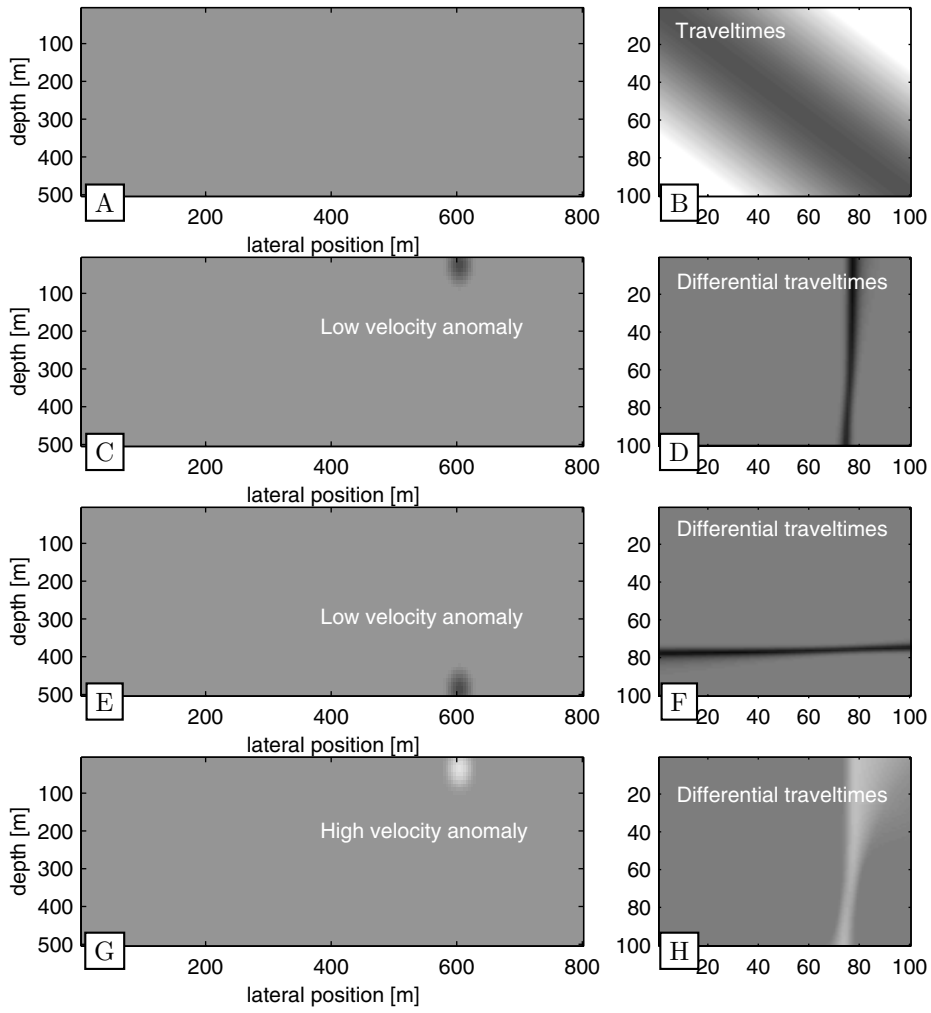

Figure 5.5: *Imprint of subsurface anomalies on operator traveltimes. A) displays a background velocity model in which rays are traced from all lateral locations at a depth of 500m to all lateral locations at a depth of 0m. Traveltimes are displayed in B). C),E) and G) display anomalies placed within the background model. The traveltime differences with respect to the anomaly free model are displayed in D),F) and H). The anomalies become visible in a row and a column of the focusing operator, depending on their location in depth.*

receivers are placed at the same lateral locations,

$$\mathbf{W}^{+}(z_0, z_d) = \mathbf{W}^{-,T}(z_d, z_0), \tag{5.3.1}$$

the near surface effects affecting the upward propagator can approximately be covered by smoothly modeled downward propagators when a complete dense set of downward propagators is considered.

5.4 Construction of initial focusing operators

As outlined in the previous section, the goal of the remainder is to determine initial focusing operators with the notion that,

- Variations at the surface side should be included, to prevent that the focusing of either receivers or sources produce artifacts that obscure the focusing analysis;
- Lateral variations in velocity and reflector position close to the focus-point locations do not affect the focusing when they are not accounted for and can be corrected for during the final focusing analysis. Hence, the initial assumption of smooth move-out for the downward propagators will hold.

A smoothness within the CSP operators will be enforced. This smoothness allows for a parameterization that can be solved with an inversion process using the fact that we have a redundancy in the data if we opt for a sparse set of parameters. The final set of parameters will be determined through a genetic algorithm. Within this algorithm the parameters will be used to construct CSP operators. These CSP operators will then be combined according to the principle of data matching. Through this principle, the two-way traveltimes are constructed from the one-way CSP operators. The objective of the inversion algorithm is to maximize the energy when the data is stacked over the modeled two-way traveltimes.

To make sure that the procedure will lead to a set of focusing operators describing the same event in a lateral extent a proper search window for the parameters needs to be defined, to prevent mixing of recorded events.

The parameter windows are chosen around an initial set of parameters. The initial set of parameters should be defined such that we can choose the search windows as small as possible.

■ 5.4.1 Forward model and objective

To enforce smooth CSP operators throughout the updating procedure, a parameterization is introduced to describe the near surface CFP operators. The healing effect of the operators can be captured by several models.

In principle at this stage we are not interested in the correct near surface model, but merely in a model that will describe the smoothness for CSP operators while capturing the near surface effects manifested in the transposed CSP operators, the CFP operators.

To describe the smoothness of the CSP operators, many methods could be used, of which spline interpolation and Chebyshev interpolators would be among the most general descriptions. To remain somewhat close to a physical meaning, instead of using a general interpolation scheme such as a spline, the interpolation will be restricted to a generalized formulation of Normal Move-Out (NMO).

The traveltime between a surface location $\mathbf{x}_{src}$ and a CFP location $\mathbf{x}_{cfp}$ is a function of the arc length R of the ray connecting these two positions, and the spatially varying slowness field, $p(\mathbf{x})$, which is the reciprocal of the wave-propagation velocity. the offset, which will be used later on, is defined as,

$$R_2 = \sqrt{(x_{1;cfp} - x_{1;src})^2 + (x_{2;cfp} - x_{2;src})^2}. \tag{5.4.2}$$

For reasons of brevity, a 2D situation is considered only,

$$\mathbf{x} = (x = x_1, z = x_3). \tag{5.4.3}$$

Traveltime, as described by the Eikonal equation, can be thought of as a summation of segments multiplied by slowness,

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = \int_{\mathbf{x}_{src}}^{\mathbf{x}_{cfp}} p(\mathbf{x}) \mathrm{d}s, \tag{5.4.4}$$

in which s is the arc in space defining the shortest traveltime path, the ray. In Chapter 6 more details can be found about seismic rays and the computation of seismic rays. A discretized version of the traveltime equation, Equation 5.4.4, would read,

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = \sum_{i=1}^{N} p_i \Delta s_i, \tag{5.4.5}$$

in which it is assumed that N distinctive layers with differing slowness exist between the acquisition level and the datum level.

By defining an average slowness, $\tilde{p}(\mathbf{x}_{src}, \mathbf{x}_{cfp})$ along the ray under consideration[1],

$$\tilde{p}(\mathbf{x}_{src}, \mathbf{x}_{cfp}) = \frac{\sum_{i=1}^{N} p_i \Delta s_i}{\sum_{i=1}^{N} \Delta s_i}, \tag{5.4.6}$$

the traveltime of Equation 5.4.5 can then be written as a sum over segments multiplied by the average slowness,

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = \tilde{p}(\mathbf{x}_{src}, \mathbf{x}_{cfp}) \sum_{i=1}^{N} \Delta s_i, \tag{5.4.7}$$

[1]The use of $\tilde{\cdot}$ refers in this case to *average* and not to a field quantity transformed to the Radon domain

and through Pythagoras,

$$\sum_i \Delta s_i > \|\mathbf{x}_{cfp} - \mathbf{x}_{src}\|, \tag{5.4.8a}$$

$$\sum_i \Delta s_i = c(\mathbf{x}_{src}, \mathbf{x}_{cfp})\|\mathbf{x}_{cfp} - \mathbf{x}_{src}\|, \tag{5.4.8b}$$

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = c(\mathbf{x}_{src}, \mathbf{x}_{cfp})\tilde{p}(\mathbf{x}_{src}, \mathbf{x}_{cfp})\|\mathbf{x}_{cfp} - \mathbf{x}_{src}\|, \tag{5.4.8c}$$

in which c corrects for the curvature of the path between $\mathbf{x}_{cfp}$ and $\mathbf{x}_{src}$. The dimensionless correction factor c will depend on the lateral distance between $\mathbf{x}_{cfp}$ and $\mathbf{x}_{src}$, $x_{src} - x_{cfp}$. Thus, traveltime Equation 5.4.8c can then be written as,

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = \sqrt{c^2(x_{src}, x_{cfp})\|z_{src} - z_{cfp}\|^2\tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp}) + c^2(x_{src}, x_{cfp})R_2^2\tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp})}. \tag{5.4.9}$$

In order to implement a smoothness constraint to the operator describing the traveltime between a single surface location and all cfp-locations, $c(x_{src}, x_{cfp})\tilde{p}$ should be parameterized by a smooth function. The parameterization will be constrained such that for zero-offset, $R_2 = 0$, the correction factor equals to one, $c(x_{src}, x_{src}) = 1$. As such $\tilde{p}$ is no longer the average velocity along the ray, but merely the apparent velocity between the surface and datum depth at the zero-offset. We are free to do so, as long as the combination of c and $\tilde{p}$ holds in Equation 5.4.8c. Then the first term within the square-root of Equation 5.4.9 should equal the squared zero-offset time,

$$T_0(\mathbf{x}_{cfp}) = T_0(\mathbf{x}_{src}) \tag{5.4.10a}$$

$$= c(x_{src}, x_{src})\tilde{p}(\mathbf{x}_{src}, \mathbf{x}_{cfp}|x_{src} = x_{cfp})\|(z_{cfp} - z_{src})\|. \tag{5.4.10b}$$

For the parameterization of c we will choose a polynomial expansion as function of offset,

$$\begin{aligned} c^2(x_{cfp}, x_{src})\tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp}) =& \tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp}|x_{src} = x_{cfp}) \\ &+ \frac{(x_{cfp} - x_{src})}{\mathrm{MAX}(x_{cfp} - x_{src})}\Delta p_1^2(x_{src}) \\ &+ \frac{(x_{cfp} - x_{src})^2}{\mathrm{MAX}^2(x_{cfp} - x_{src})}\Delta p_2^2(x_{src}) + \mathcal{O}(N_3), \end{aligned} \tag{5.4.11}$$

in which $\mathrm{MAX}(x_{cfp} - x_{src})$ renders the maximum offset in the system. This factor is introduced such that the higher order velocity terms $(\Delta p_1^2, \Delta p_2^2, \ldots, \Delta p_N^2)$ will not be too small and hence will stabilize the inversion procedure in which we invert for the parameters $(\tilde{p}, \Delta p_1, \Delta p_2, \ldots, \Delta p_N)$. The number of orders to invert for influences

the stability of the system. For a first order description of the correction factor c we can define,

$$p_1^2(x_{src}) = \tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp}|x_{src} = x_{cfp}) + \Delta p_1^2(x_{src}) \tag{5.4.12a}$$

$$p_2^2(x_{src}) = \tilde{p}^2(\mathbf{x}_{src}, \mathbf{x}_{cfp}|x_{src} = x_{cfp}) - \Delta p_1^2(x_{src}) \tag{5.4.12b}$$

$$\alpha(x_{src}, x_{cfp}) = \frac{\mathrm{MAX}(x_{cfp} - x_{src}) - (x_{cfp} - x_{src})}{\mathrm{MAX}(x_{cfp} - x_{src})} \tag{5.4.12c}$$

which will lead, when combining Equation 5.4.10b, Equation 5.4.9 and Equation 5.4.12c, to,

$$T(\mathbf{x}_{cfp}, \mathbf{x}_{src}) = \sqrt{T_0^2(x_{src}) + \left[(1 - \alpha(x_{src}, x_{cfp}))p_1^2(x_{src}) + \alpha(x_{src}, x_{cfp})p_2^2(x_{src})\right](x_{src} - x_{cfp})^2}. \tag{5.4.13}$$

The optimization procedure should then search for each surface location the set of three parameters, $\{T_0, p_1, p_2\}(x_{src})$. Note that the two slowness parameters are linearly weighted to determine an effective slowness for each offset. In Figure 5.6 the parameterization is explained in a graphical form.

In Verschuur and Mahrfoul (2005), an additional parameter γ was introduced to define a nonlinear weighting as well as an additional offset perturbation. The γ value was set to a fixed number within a trial and error exercise whereas the additional offset perturbation was taken as an additional parameter to invert for, resulting in a parameter-set of four unknowns for each surface location. By introducing an additional order within the smoothness description of Equation 5.4.11 a comparable effect would be obtained.

The outlined forward model will be used in combination with the principle of data construction to determine the set of 3xN_{src} parameters that best describe the measured two-way traveltimes. Compared to the number of parameters to estimate, the number of data-points, N_{src}xN_{rcv}, assures an overdetermined system. Through the stationary phase analysis which was used to derive the Fermat modeling method, Equation 4.6.80 in Chapter 4, for a known set of parameters the cfp-location $\mathbf{x}_{cfp}$ hit by the specular ray of the reflected wavefield between source location $\mathbf{x}_{src}$ and receiver location $\mathbf{x}_{rcv}$ is automatically determined by solving,

$$0 = \nabla \Bigg[\sqrt{T_0^2(x_{src}) + \left[\alpha(x_{src}, x)p_1^2(x_{src}) + (1 - \alpha(x_{src}, x)p_2^2(x_{src})\right](x_{src} - x)^2} + \sqrt{T_0^2(x_{rcv}) + \left[\alpha(x_{rcv}, x)p_1^2(x_{rcv}) + (1 - \alpha(x_{rcv}, x))p_2^2(x_{rcv})\right](x_{rcv} - x)^2} \Bigg] \Bigg|_{x = x_{cfp}}, \tag{5.4.14}$$

which resembles Fermat's principle. It describes all possible two-way traveltimes from one source, via all possible CFP locations toward the receiver. The stationary

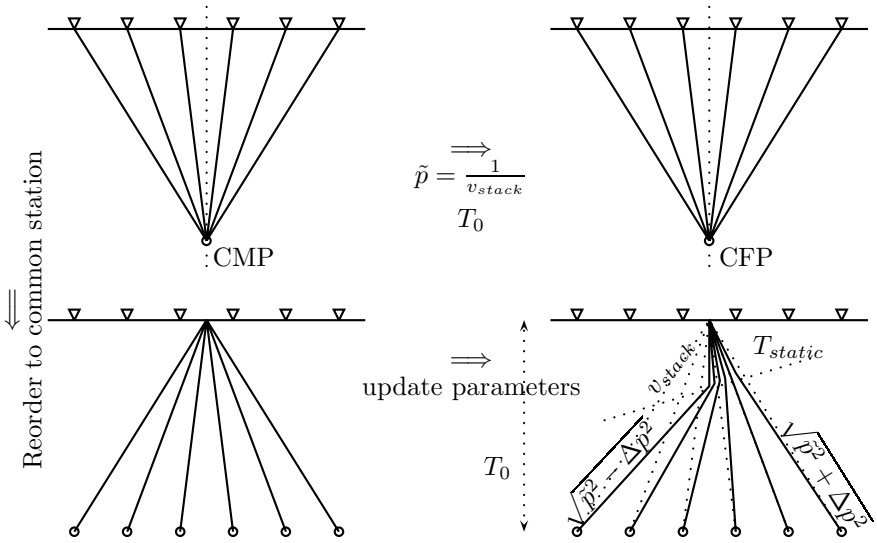

Figure 5.6: *When constructing traditional initial CFP-gathers, use is made of a rough estimate of the stacking velocities and zero offset times (top Figures). By reordering the initial operators, better first estimates are obtained for the operators in presence of complex near surface problems (bottom Figure). Within the updating procedure, the operators are corrected and fine tuned for the localized near surface effects.*

point yields the actual two-way traveltime. For each forward model realization, this expression will be evaluated for each source and receiver combination. The procedure is schematically sketched in Figure 5.7.

To verify whether the proposed parameterization is adequate to determine the initial operators through data-matching, the operators found through standard updating on a synthetic dataset are subject to an inversion procedure. To this aim, a linearized inversion procedure is used to determine the described parameters from the correctly modeled operators derived by computing the over the model of Figure 5.8C, using an Eikonal solver.

Opposite to the data-matching criterion that will be used to find the CFP operators, in this example the operators are already given. Their traveltimes are a function of the exact CFP locations as well. These CFP locations are no longer a given as explained in Chapter 4, because of drift during the updating. Therefor, to test our parameterization on a given set of operators we need to extend the standard set of parameter with the lateral CFP-positions; The final inversion result of this example will thus not only contain the set $(T_0, p1, p2)\{\mathbf{x}_{src}\}$ for all surface locations but also

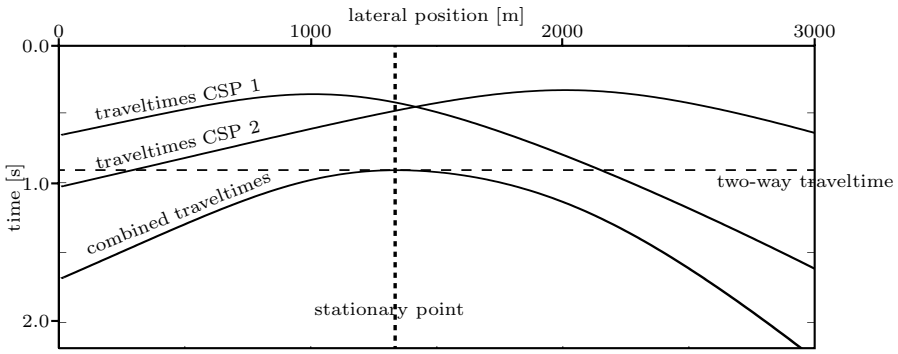

Figure 5.7: *By combining the traveltimes from surface location CSP 1 to all subsurface locations along a single reflector with the traveltimes from surface location CSP 2 to all subsurface locations along the same reflector, the two-way traveltime from CSP 1 to CSP 2 for the reflection energy can be found at the stationary point of the combined traveltime curve. The lateral position of the stationary point denotes the reflection position of the specular ray.*

the set $(x_{1;1} \dots x_{1;N_{cfp}})$ of lateral coordinates of the CFP-locations. The total number of parameters to invert for will therefore be $3\text{x}N_{src} + N_{cfp}$. The data-set with $N_{src}\text{x}N_{src}$ data-points is still redundant to allow for an overdetermined system.

The linearized inversion scheme used is described in Appendix B. Figure 5.8A displays 3 operators in the common surface point domain, overlain by the operators described by the best fitting parameters found through inversion. Note that the residuals are small, such that the two sets of curves almost completely overlap. Figure 5.8B shows 3 operators within the Common Focus Point domain. The effects caused by the near surface layer are clearly covered by the described parameters. Since the parameters also include the approximate locations in space of the CFP's, it is worthwhile to compare them to the location of the reflector on which the operators are focused. Although the locations are not exactly derived, as could be expected from an approximate description of the near surface, they do resemble remarkably well the true locations of the reflector as can be seen in Figure 5.8C. Especially when taken into account the initial location, resembled by the crosses, which were used to initialize the inversion procedure.

■ 5.4.2 Inversion for CSP operators by Simulated Evolution

Genetic algorithms or simulated evolution methods are a subset of so called Monte Carlo methods (Gallagher et al., 1991; Robert and Casella, 2005). Monte Carlo methods are based on random search techniques through parameter space. Within these random search techniques it is possible to optimize the search strategy and

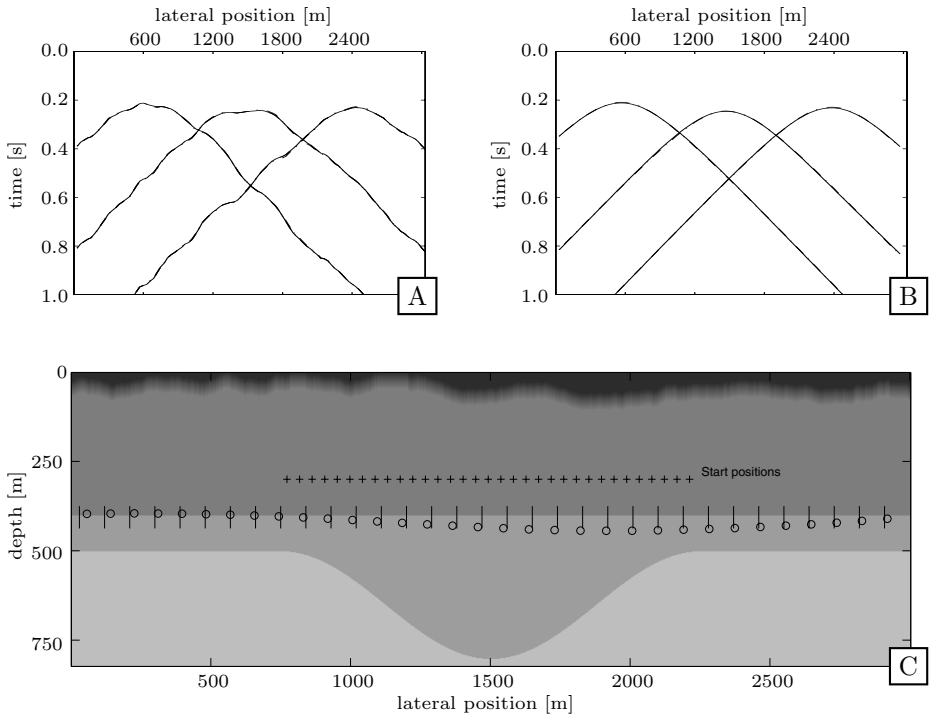

Figure 5.8: *Validation of the Common Shot Point operator parameterization. Figure A shows the original CFP operators overlain by the estimated parameterized operators. Figure B shows the same results after reordering to CSP operators. Figure C displays the velocity model for which the original operators were modeled. Within the model, the initial CFP locations are displayed by crosses. The open circles display the final estimated locations; the true locations are indicated by the bars. Clearly the parameterization resolves the operator times very well and the locations are estimated reasonably accurate as well.*

thus obtaining more efficient algorithms. Simulated evolution is one of these more sophisticated search methods and is based on evolutionary processes as observed in nature and was used by (Wilson et al., 1994) to determine residual statics. Another more sophisticated method is the Heat Bath Algorithm, which is based on cooling down processes and recrystallization of certain elements contained within the heated fluid, also referred to as simulated annealing (Cerny, 1985).

In this thesis the simulated evolution is adopted to invert for the parameterized traveltimes as described by Equation 4.4.13. The objective is to minimize the difference between the measured traveltimes of a certain event $T_d(\mathbf{x}_{rcv}, \mathbf{x}_{src})$ and the forward modeled traveltimes,

$$T(\mathbf{x}_{rcv}, \mathbf{x}_{cfp}) + T(\mathbf{x}_{src}, \mathbf{x}_{cfp}). \tag{5.4.15}$$

The focuspoint location, $\mathbf{x}_{cfp}$ is determined within the inversion scheme by Equation 4.4.14.

The method is based on the evolution of populations, where repeatedly two members of a current generation are selected to produce two children of the next generation. Within each iteration (in terms of evolution of populations, a generation of the population), population members are selected according to a certain fitness, to form a pair of parents. Such a pair of parents will produce a pair of children during the crossover phase. In terms of evolution, genes of both parents are mixed. Finally, at random some children or parents have a chance to mutate in the mutation phase. This last phase keeps some randomness in the population pool and prevents the algorithm to converge prematurely. In Appendix B an extensive description is given on the simulated evolution method.

The fitness of a population member is determined by the stack power obtained by stacking the data along the two-way traveltimes computed by Equation 5.4.15. The genes of a particular member equals the parameter vector. To initialize a population for such a genetic algorithm it is recommended to use the static methods as described in Chapter 2, Section 2.7. These alternative methods provide initial stack velocities in combination with an initial T_0 that is constructed from the picked zero-offset times of a certain event in combination with the computed static shifts. A full population is created by selecting random values within a band around the initial parameter values.

The genetic algorithm is demonstrated for obtaining adequate initial focusing operators for data related to the subsurface model of Figure 4.12A. The operators are parameterized within the Common Surface Point domain, to describe the measured data in terms of phase through the use of the principle of data construction. The result of the genetic algorithm is displayed in Figure 5.10. The top row, Figure 5.10A, displays five of the 301 shot-records that are used within the genetic data matching procedure. The next five rows display the shot records corrected with the found

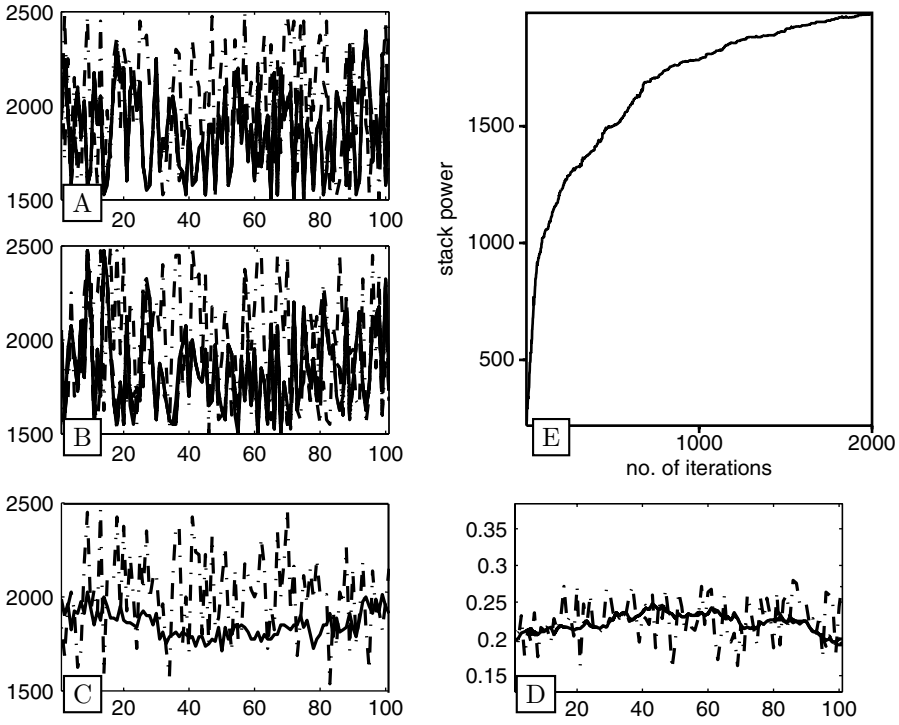

Figure 5.9: *Performance of the genetic algorithm to estimate initial operators through the principle of data construction. A) B) and C) show the estimated values for* $p1$*,* $p2$ *and* $(p1+p2)/2$ *respectively. D) shows the estimated* T_0*. E) shows the stack power as function of the iteration-number.*

two-way traveltimes after respectively 100, 500, 1000, 1500 and 2000 iterations of the genetic algorithm. The obtained operators are displayed in Figure 5.11. As can be concluded from Figure 5.11 the parameters tend to roughly describe the measured data in terms of phase, capturing the longer wavelengths of the near surface anomalies. Some jitter is still observed, however the operators found can still be subject to a final update procedure to capture the full kinematic description of the operators, as the proposed procedure was only meant to define initial operators.

The performance of the algorithm is displayed in Figure 5.9A and Figure 5.9B show the estimated $p1$ and $p2$ respectively. A lot of jitter is observed, such that when $p1$ is large $p2$ is small and vice versa. The average however, behaves quite smooth. This is probably because of the fact that the normal to the near surface interface is changing rapidly laterally, causing a rapid changing skew in the CSP op-

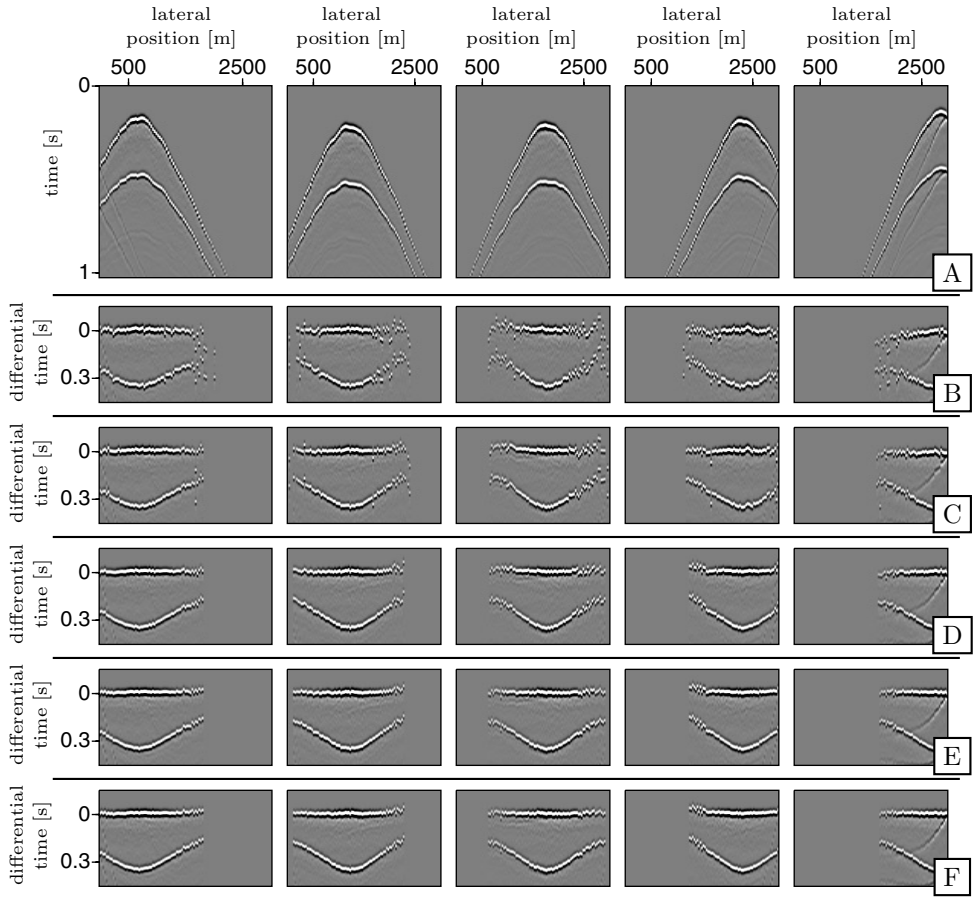

Figure 5.10: *Results of a genetic algorithm for updating the parameters describing CSP operators in order to optimize the alignment of energy for a particular event within the prestack shot-gathers. A) 5 of the original shot-gathers. B)-F) shot-gathers corrected with the modeled two-way traveltime after B) 100 iterations, C) 500 iterations, D) 1000 iterations, E) 1500 iterations, F) 2000 iterations.*

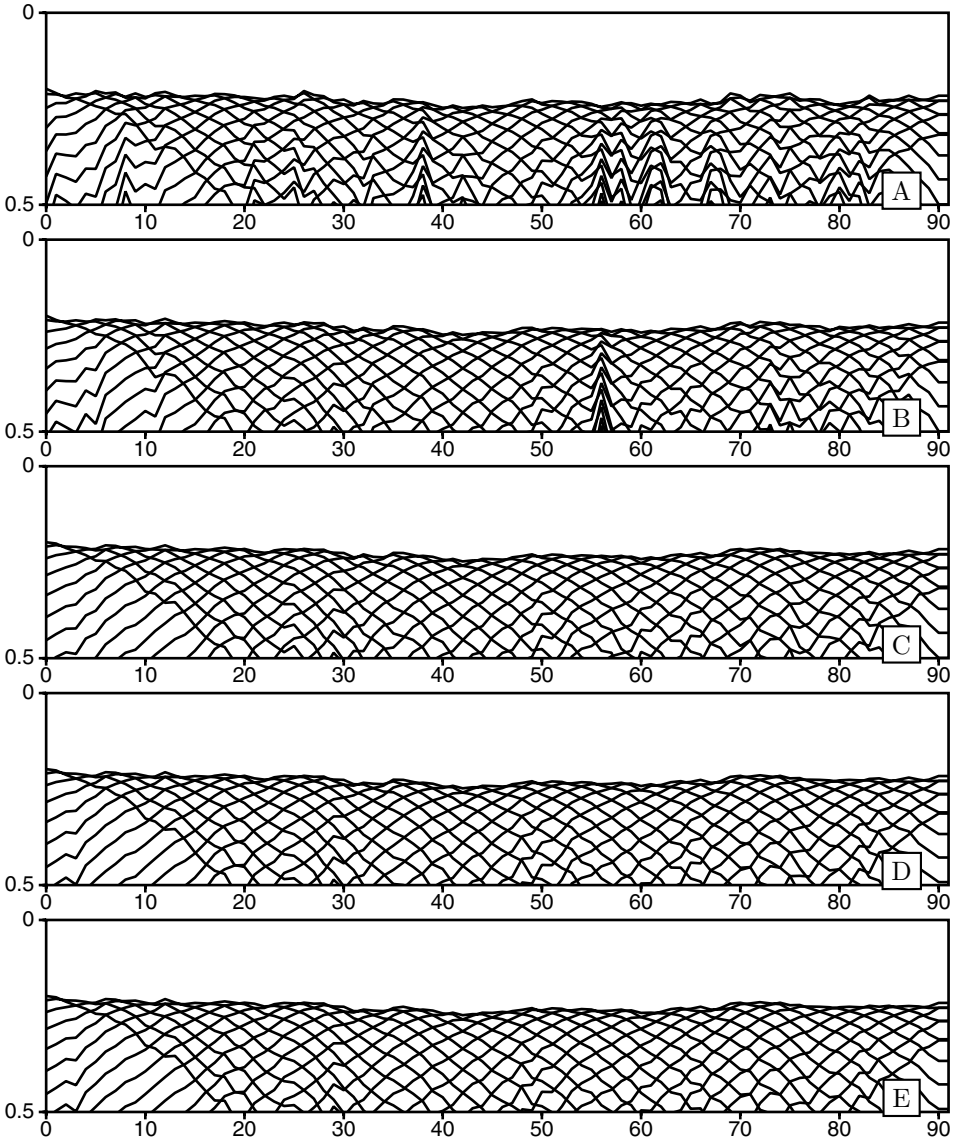

Figure 5.11: *The evolution of CFP operators during updating of the parameters that describe parameterized CSP operators, using a genetic algorithm after A)100, B)500, C)1000, D)1500 and E)2000 iterations.*

erators. Note that the estimated T_0 profile of Figure 5.9D quite well follows the near surface behavior. Note that the stackpower, which is the objective function to be maximized, is continuously increasing during the iterations, as visible in Figure 5.9E.

Once the final operators are determined through procedures described in this chapter and the previous chapter, the operators can be subjected to an inversion procedure to determine a near surface model as described in Chapter 6, or be used to derive amplitudes and to redatum the measured data, described in Chapter 7.

6

Traveltime operator inversion, a tomographic solution

6.1 Introduction

Within the previous chapters it has been outlined and demonstrated how to obtain one-way operators from the measured seismic data, contaminated by near surface heterogeneities. Since the described procedures make use of the CFP methodology, the results fully describe the near surface behavior in terms of propagation characteristics, directly obtained from the data itself. A big advantage is that the results are not constrained by any simplified underlying static model. Therefore, when the found operators are used to redatum the measured data, as will be described in Chapter 7, any time and offset dependency of the near surface anomalies is taken properly into account. However, as the near-surface anomalies are only described in terms of propagation characteristics, any information on the shape, depth, spatial sampling of the event on which the operator estimation procedure has been focused, is not available. As this event also serves as the new datum level, it will be worthwhile to gain knowledge on its geometry in space. For this purpose one could use tomographic inversion, as the found operators are basically describing the traveltimes between a point source at the datum, of which the location is still unknown, and a series of receivers at surface. The tomographic problem therefore closely resembles an earthquake localization problem. In Cox (2004) such type of tomographic inversion is extensively discussed and described for the general CFP methodology. Here, we only briefly describe the tomographic method and describe a specific parameterization and forward modeling step which was used for the near surface problems described in this thesis.

Finally, it should be noted that the tomographic procedure will not be able to

fully capture the near surface anomalies, but only its lower frequency parts. As long as the datum-reflector, used to determine the CFP-operators, is behaving relatively smooth, we should be able to obtain information on its geometry. Opposite to conventional static models, the method outlined in this thesis results in maximum resolution given by the data as the determination of the final focusing operators is fully data-driven, whereas the conventional methods are mainly model driven. The estimation of the near-surface model as described in this thesis merely serves to aid in capturing the final CFP locations rather than determining the final focusing and redatuming result. Model-driven solutions will hardly ever be able to capture all the propagation effects of the near surface adequately.

6.2 Traveltime inversion of CFP operators

In the previous two chapters, the focusing operators were defined through updating, for which initial operators with a surface consistent imprint were employed. Though the updating was based on parameterized models, these models were only used to initially constrain the updating process, and the final parameters are not considered to be reliable values for depth conversion. To appreciate the redatuming results, we need to know the lateral positions and depth of the focus points. Therefore, the found operators will be used as input for a tomographic inversion problem. The model parameters that will be defined through inversion are the CFP-locations and parameters that describe a velocity field of the near surface to explain the propagation characteristics of the defined operators. The actual parameterization will be addressed in Section 5.4. Here we will suffice describing the general inversion process.

The tomography problem tries to minimize the following objective function,

$$F = \mathbf{e}^T\mathbf{e}, \tag{6.2.1}$$

with,

$$\mathbf{e}^T = \left[\mathbf{C}_n^{\frac{1}{2}}\left(\mathbf{d}-\mathbf{t}(\mathbf{m})\right), \mathbf{C}_m^{\frac{1}{2}}\left(\mathbf{m}_p-\mathbf{m}\right)\right], \tag{6.2.2}$$

in which $\mathbf{d}$ resembles the data, being the estimated one-way traveltimes, $\mathbf{t}(\mathbf{m})$ the modeled operator times from parameters $\mathbf{m}$ with $\mathbf{m}_p$ being the a priori values of the parameters with a priori variances $\mathbf{C}_m^{\frac{1}{2}}$. The parameters $\mathbf{m}$ describe both the velocity depth model as well as the focal point locations at the datum (see Section 5.4). In general the parametric inversion of traveltimes is a non-linear problem and could be solved iteratively through local linearization. As outlined in Appendix B, starting with an initial set of parameters, $\mathbf{m}_0$, we can update the parameters such that after k iterations we would arrive at,

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha\mathbf{p}_k, \tag{6.2.3}$$

$$\alpha\mathbf{p}_k = \left(\mathbf{J}^T(\mathbf{m}_k)\mathbf{J}(\mathbf{m}_k)\right)^{-1}\mathbf{J}(\mathbf{m}_k)^T\mathbf{e}_k, \tag{6.2.4}$$

$$\mathbf{e}_k^T = \left[\mathbf{C}_n^{\frac{1}{2}}\left(\mathbf{d}-\mathbf{t}(\mathbf{m}_k)\right)\right], \tag{6.2.5}$$

in which we neglected the Hessian and assumed a null state of information on the a priori values of the parameters. The term $(\mathbf{d} - \mathbf{t}(\mathbf{m}_k))$ resembles an error in traveltime between the measured data and forward modeled data. In that respect the iterative solution to determine the parameters that describe the near surface model can be found in two ways.

The first strategy takes the operator times as determined through the steps described in Chapters 4 and 5 as the true operators, describing the near surface propagation characteristics as good as possible, and assumes the DTS-gathers have already flattened perfectly around $t = 0$s. Next the operators are compared directly with synthetically modeled operators given the current set of model operators. The iterative sequence of Equation 6.2.3 through Equation 6.2.5 will then update the parameters until the forward modeled operator times satisfy the operator times which were found through the principle of equal traveltime.

The second strategy takes the modeling misfit as the picked times within the DTS gathers. Following the principle of equal traveltime as explained in Chapter 4,

$$\frac{1}{2}\mathbf{t}_{DTS} \approx \mathbf{d} - \mathbf{t}(\mathbf{m}_k), \tag{6.2.6}$$

in which $\mathbf{t}_{DTS}$ are the traveltimes picked in the DTS gathers and $\mathbf{d}$ is considered to be the set of traveltimes of the correct and true operators. The estimated operators, $\mathbf{t}(\mathbf{m}_k)$, are forward modeled within a model given by the parameter-set $\mathbf{m}_k$. After which the DTS gathers are constructed. Then the picked misfits are divided by 2 and used to update the model parameters. This strategy takes into account that the operator times, modeled with the current set of parameters, are offset to the correct operators by approximately half the times picked within the DTS-gathers; the problems of model estimation and operator time estimation are solved simultaneously.

Both strategies are schematically displayed in Figure 6.1. The first strategy hardly puts any model constraints to the process of operator time estimation and can therefore be considered as a step in a model independend migration algorithm. The constraining parameterization is postponed. Within the second strategy, the operator updating process has become a model constrained process and it is therefore not assured that the operator times are found within a reasonable error margin. The first strategy was is adopted in this thesis, whereas the second strategy was adopted by e.g. Kabir and Verschuur (2000). The forward modeling step will be explained within the next section.

Note that the linearized formulation of Equation 6.2.3 through Equation 6.2.5 does not restrict ourselves to a single data set. In principle we could combine both strategies and add for instance additional Tikhonov regularizations and even alternative data, such as first-break measurements.

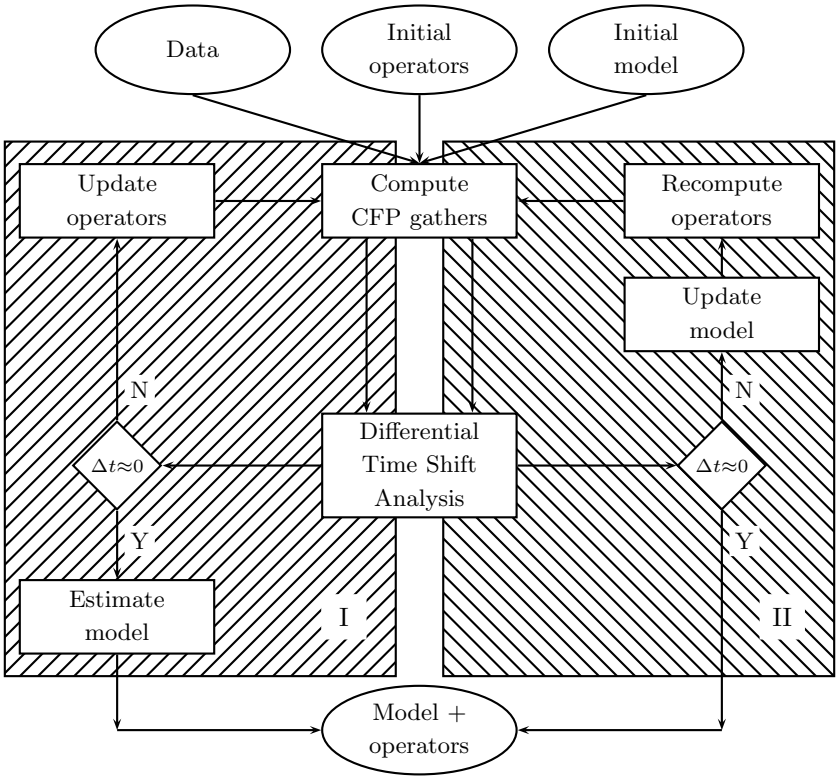

Figure 6.1: *Estimating the near surface model can be accomplished in two ways. Either the operators will be updated until the principle of equal traveltime is satisfied, after which the model will be estimated such that the it explains the final operators, as demonstrated through the left-hand route. Alternatively, estimating the model will be coupled to the operator updating directly, by updating the model-parameters such that the newly computed operators will be closer to satisfying the principle of equal traveltime, as demonstrated by the right-hand route.*

The Jacobian, $\mathbf{J}(\mathbf{m}_k)$ in Equation 6.2.4, is the operator containing the operator traveltime derivatives with respect to the current parameter set $\mathbf{m}_k$. In most instances these parameters have a localized effect. For instance the CFP-locations impact only the traveltimes between the surface and the specific CFP-locations. Hence, traveltime derivatives with respect to the coordinates of a specific CFP-location will render zero for all traveltimes which are not related to that particular CFP location. Velocity parameters also will have localized effects. Velocity nodes that are not in the neighborhood of a particular CFP-location or surface location will not affect the traveltimes measured between those two locations. Thus, in general the Jacobian operator $\mathbf{J}$ that is needed for non-linear optimization can be approached as being a sparse operator. For the 2D situation with N surface locations, M unknown CFP locations and L unknown velocity model parameters, the Jacobian matrix can be written as,

$$\mathbf{J} = \begin{bmatrix} \begin{bmatrix}\mathbf{J}_{11}\end{bmatrix} & \cdots & \begin{bmatrix}\mathbf{J}_{1N}\end{bmatrix} & \begin{bmatrix}\mathbf{X}_1\end{bmatrix} & \begin{bmatrix}\mathbf{Z}_1\end{bmatrix} \\ \vdots & & \vdots & & \\ \begin{bmatrix}\mathbf{J}_{M1}\end{bmatrix} & \cdots & \begin{bmatrix}\mathbf{J}_{MN}\end{bmatrix} & \begin{bmatrix}\mathbf{X}_M\end{bmatrix} & \begin{bmatrix}\mathbf{Z}_M\end{bmatrix} \end{bmatrix}, \tag{6.2.7}$$

in which $\mathbf{J}_{mn}$, containing the traveltime derivatives with respect to the model velocities p_l, is defined by,

$$\begin{aligned}
\mathbf{J}_{m1} &= \begin{bmatrix} \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_1^1} & \cdots & \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_l^1} & \cdots & \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_L^1} \\ & & \mathbf{0} & & \end{bmatrix}, \\
\mathbf{J}_{mn} &= \begin{bmatrix} & & \mathbf{0} & & \\ \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_1^n} & \cdots & \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_l^n} & \cdots & \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_L^n} \\ & & \mathbf{0} & & \end{bmatrix}, \\
\mathbf{J}_{mN} &= \begin{bmatrix} & & \mathbf{0} & & \\ \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_1^N} & \cdots & \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_l^N} & \cdots & \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_m^{cfp}})}{\partial p_L^N} \end{bmatrix},
\end{aligned} \tag{6.2.8}$$

and $\mathbf{X}_m$, containing the derivatives with respect to the lateral component of the CFP locations, is defined by,

$$\mathbf{X}_1 = \begin{bmatrix} \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_1^{cfp}})}{\partial x_1^{cfp}} & \mathbf{0} \\ \vdots & \\ \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_1^{cfp}})}{\partial x_1^{cfp}} & \mathbf{0} \\ \vdots & \\ \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_1^{cfp}})}{\partial x_1^{cfp}} & \end{bmatrix}, \quad \mathbf{X}_m = \begin{bmatrix} \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_m^{cfp}})}{\partial x_m^{cfp}} & \\ \vdots & \\ \mathbf{0} \quad \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_m^{cfp}})}{\partial x_m^{cfp}} & \mathbf{0} \\ \vdots & \\ \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_m^{cfp}})}{\partial x_m^{cfp}} & \end{bmatrix}, \quad \mathbf{X}_M = \begin{bmatrix} & \frac{\partial t(\mathbf{x_1^{rcv}},\mathbf{x_M^{cfp}})}{\partial x_M^{cfp}} \\ & \vdots \\ \mathbf{0} & \frac{\partial t(\mathbf{x_n^{rcv}},\mathbf{x_M^{cfp}})}{\partial x_M^{cfp}} \\ & \vdots \\ & \frac{\partial t(\mathbf{x_N^{rcv}},\mathbf{x_M^{cfp}})}{\partial x_M^{cfp}} \end{bmatrix}, \tag{6.2.9}$$

and $\mathbf{Z}_m$ is defined alike for the depth of the CFP locations. Each submatrix $\mathbf{J}_{mn}$, of size $N \times L$, contains only one row with non-zero entries. If velocity parameters only

have a local effect, in other words, only influence a limited number of rays, a large number of entries within the non-zero row will become zero as well. The submatrices $\mathbf{X}_m$ and $\mathbf{Z}_m$, of size $N \times M$, contain only one column with non-zero entries. As a result, the Jacobian matrix is extremely sparse. Especially algorithms like Conjugate Gradient are well suited to implicitly invert sparse systems in a least-squares manner, as given by Equation 6.2.4. In this thesis we adopt the LSQR algorithm (Appendix A). Besides all the described advantages of sparse system solutions it furthermore tends to solve parameter directions with high eigenvalues first. As most of the problems tend to be ill-conditioned the method shows numeric regularization when only a limited number of iterations is computed. Hence stable updates are determined in a fast way.

6.3 Forward modeling of seismic traveltimes

Full wave-equation inversion to find the underlying earth model which produced the output of the seismic experiments is not only a cumbersome task but under determined as well. Generally, in seismic exploration a background velocity model is determined in which the Green's functions are computed using the wave-equation to image contrasts through migration. The global shape of the move-out of arrival traveltimes mainly depend on the propagation velocities of the background model as function of spatial coordinates, whereas the amplitudes of reflection events depend on contrast differences and geometrical spreading. Since, within the dynamic approach to near-surface problems in this thesis, we need to know the position of the datum to correctly redatum and perform a layer replacement, the reflector depth will be determined from the traveltimes. For this purpose, the ray-method is adopted, which looks for a solution of the coupled wave-equation, Equation 4.2.1a and Equation 4.2.1b, valid for high frequencies. The solution is accomplished by the ray-series solution, see Cerveny (1985),

$$p(\mathbf{x},\omega) = \exp\left[-j\omega(t-\tau(\mathbf{x})\right] \sum_n \frac{A_n(\mathbf{x})}{(-j\omega)^n}, \tag{6.3.10}$$

in which $\tau(\mathbf{x})$ represents the Eikonal or phase function and $A_n(\mathbf{x})$ the amplitude of the nth order. In the high frequency approximation we can suffice with the zeroth order approximation of the ray-series. By substituting the zeroth order approximation in the homogeneous wave-equation, we arrive at,

$$\left[|\nabla\tau|^2 - \frac{1}{c^2(\mathbf{x})}\right]\omega^2 A_0(\mathbf{x}) + \nabla^2 A_0(\mathbf{x}) + j\omega\left[2\nabla\tau(\mathbf{x})\nabla A_0(\mathbf{x}) + A_0(\mathbf{x})\nabla^2\tau(\mathbf{x})\right] = 0, \tag{6.3.11}$$

in which the velocity is defined by $c^2(\mathbf{x}) = \rho(\mathbf{x})\kappa(\mathbf{x})$. Under the assumption that $\omega^2 >> \frac{\nabla^2 A_0(\mathbf{x})}{A_0(\mathbf{x})}$, the real part of Equation 6.3.11 leads to the Eikonal equation,

$$|\nabla\tau(\mathbf{x})|^2 = \frac{1}{c^2(\mathbf{x})}, \tag{6.3.12}$$

and the imaginary part to the transport equation,

$$2\nabla\tau(\mathbf{x})\nabla\log\left(A_0(\mathbf{x})\right) + \nabla^2\tau(\mathbf{x}) = 0. \tag{6.3.13}$$

The Eikonal equation is closely linked to Fermat's principle, since the traveltime, τ_{AB} along an arbitrary path between A and B,

$$\tau_{AB} = \int_A^B \frac{\partial\tau}{\partial s}\mathrm{d}s, \tag{6.3.14}$$

is minimum along a ray path, which can be verified by inserting the Eikonal of Equation 6.3.12 into the traveltime expression of Equation 6.3.14. The Eikonal equation and Fermat's principle can advantageously be used to determine the traveltime between two locations within a known velocity field.

Two major families of algorithms are recognized: raytracing and grid-based methods. Raytracing computes the traveltimes within a potential field by integrating the system of so called ray equations, which will be explained in the next section. Grid-based methods are an attempt to directly evaluate the Eikonal equation on a gridded velocity field. The algorithms are based on shortest path calculations using graph methods, direct finite differencing on the velocity grid or a combination of the two, which will be explained in Sections 5.3.2 and 5.3.3.

■ 6.3.1 Raytracing

From the Eikonal equation, the system of ray equations can be deduced by introducing the ray-parameter σ which is related to the arclength and traveltime along the ray by $\mathrm{d}\sigma = c\mathrm{d}l = c^2\mathrm{d}\tau$,

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\sigma} = \mathbf{p}, \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\sigma} = \frac{1}{2}\nabla\frac{1}{c^2}, \tag{6.3.15}$$

in which $\mathbf{p}(\mathbf{x})$ is the slowness vector, having a length equal to the reciprocal of the velocity, $c(x)$ Instead of choosing the independent ray-parameter σ other choices are possible as well, as pointed out by Cerveny (1985). The ray-trajectory is found by numerical integration of the ray-system using for instance the Runge-Kutta method (Abramowitz and Stegun, 1970) and the proper initial conditions,

$$\begin{aligned} \sigma =& \sigma_0, \\ \mathbf{x}(\sigma_0) =& \mathbf{x}_{src}, \\ \mathbf{p}(\sigma_0) =& \frac{1}{c(\mathbf{x}_0)}[\cos(\phi)\cos(\psi), \sin(\phi)\cos(\psi), \sin(\psi)], \end{aligned} \tag{6.3.16}$$

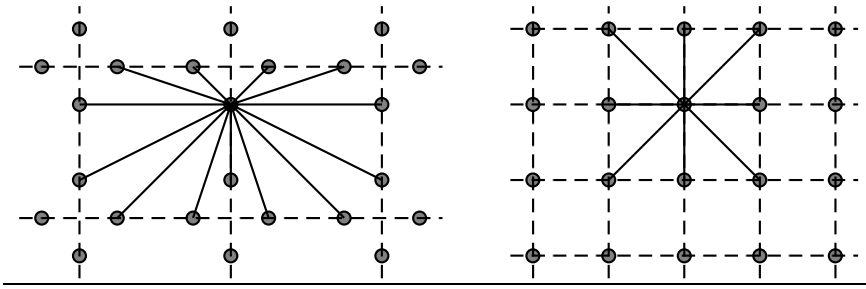

Figure 6.2: *Two possible stencils that can be used in finding the shortest path. The dotted lines denote the cell boundaries. The solid lines denote the connections between one particular node and its neighbors. Along the connections the penalties are computed, being equal to the traveltime along the connecting path.*

in which ϕ and ψ are respectively the azimuth and dip of the shooting direction. The traveltime is found by additional integration along the ray,

$$\tau(\sigma) = \int_{\sigma_0}^{\sigma} \frac{1}{c^2(\sigma)} \mathrm{d}\sigma. \tag{6.3.17}$$

Ray-tracing algorithms have some drawbacks and advantages.

- Ray-tracing represents a boundary value problem. A drawback with boundary value problems is that one does not know on beforehand what the initial values should be to let the ray end in the desired end-point. This can be overcome by raybending, by iteratively changing the initial parameters through an optimization scheme and through dynamic raytracing which extrapolates traveltimes between rays toward the desired end point by keeping track of the wavefront curvature along the ray. When the medium becomes complexer these methods tend to become expensive and unstable due to, for instance, caustics.
- By solving the ray-tracing system through numerical integration, besides the traveltime, also the exact raypath is known between the two locations. If a raytracing method is used as forward modeler within an inversion problem, the known raypaths can advantageously be used to define the partial derivatives of the traveltimes with respect to the model parameters. This will save the inversion algorithm from additional forward modeling to determine the derivatives.

■ 6.3.2 Shortest path methods

Since the Eikonal equation, Equation 5.3.14., resembles Fermat's principle, which states that the traveltime along the ray connecting two points should be stationary

and minimal, a suite of traveltime algorithms exist which are based on Dijkstra's shortest path algorithm (Dijkstra, 1959). The shortest path algorithm is based on graph methods in network theory, finding the minimal traveltime without evaluating the differential equations directly. In Moser (1991) the shortest path algorithm is outlined for computing the first arrival traveltimes in seismology. Within the shortest path method, the earths velocity field is defined in cells. Nodes, in which the first arrival times will be computed, are place along the boundaries of a cell.

A weight matrix d_{ij} is defined, containing penalty values for the connection between n_i and node n_j. Penalties are only defined for nodes that are directly adjacent to each other in the sense that the connection is not crossing any cell-borders. All other penalties are set to infinity. These penalty values define the cost of traveling from node n_i to node n_j, in our case the traveltime, given the velocity of the cell in between the two nodes and the length of the connection. The connections and penalties follow so called stencils, of which two are shown in Figure 6.2, depending on the model set-up.

Once the weight matrix d_{ij} is defined, the shortest path between source node n_s and all other nodes can be computed through Bellman's equations (Moser, 1991),

$$\tau(n_i) = \min_{j \neq i}[\tau(n_j) + d_{ij}]. \tag{6.3.18}$$

$$\tau(n_s) = 0. \tag{6.3.19}$$

These equations can be iteratively and simultaneously solved. For each iteration the traveltimes at each node will be updated, if needed, until in the last iteration no more updates are required. It can be shown that final minimum traveltimes will be certainly acquired after a finite number of iterations, depending on the algorithm used. For instance the original algorithm of Dijkstra (1959) requires at most N iterations, with N being the total number of nodes.

More advanced methods divide the set of nodes into two subsets, subset P contains the nodes for which the minimum traveltime has already been computed, subset Q contains the nodes for which a minimum traveltime has not been computed yet. Initially, only the source node n_s resides in P. In a first step all nodes around n_s (which are all in Q) will be updated by using the connections with n_s and the weight matrix d_{ij} as defined by the used stencil. In a next step the node with minimum traveltime in Q will be moved from Q into P and all the nodes in Q around the node that just has been moved will be updated according to the used stencil and the process will be repeated until all nodes have been moved from Q to P.

The most expensive part of the algorithm is sorting the nodes in Q according to their traveltimes to be able to select the node with minimum traveltime. This can be done through a so called heap-stack algorithm (Moser, 1991).

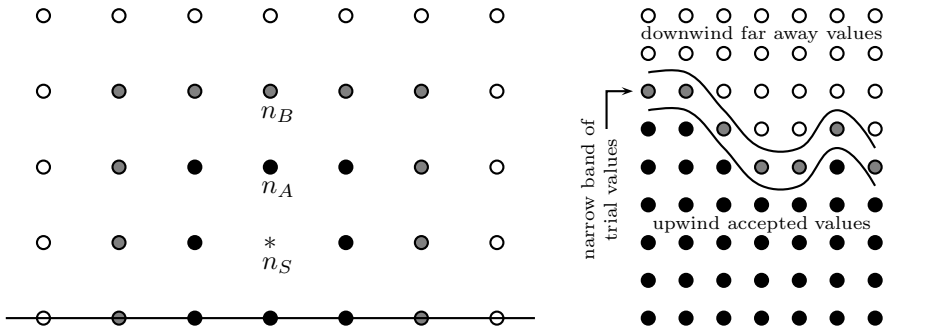

Figure 6.3: *A) Shows the principle of a finite difference scheme, using expanding rings. Starting from source location n_S at some point the rings have expanded such that the times at the solid black nodes have already been determined and the next ring to be evaluated consists of the rings with the gray nodes. If n_A is the node with the smallest traveltime from the previously evaluated ring, then n_B will be the first node to be updated, and so on. B) Shows the three regions as defined within the fast marching method (courtesy Sethian and Popovici, 1999)*

The order of operations to compute the first arrival times throughout the grid using these more advanced algorithms is $O(n\log n)$.

■ 6.3.3 Finite difference methods

Opposite to shortest path methods, finite difference methods, like the algorithm developed by Vidale (1988), evaluate directly the differential Eikonal equation on a regularly spaced grid. The subsurface velocity model is defined on the grid nodes, on which also the first arrival traveltimes will be computed.

The finite difference computations will be evaluated on expanding rings, starting at the source position. Assume a number of rings have been expanded, as displayed in Figure 6.3A; the black dots have already been evaluated, the open dots will be evaluated during later advancement and the gray dots will currently be evaluated. In the current evaluation, the traveltimes of the gray dots are determined through forward differencing using the traveltimes of the previously evaluated nodes.

The nodes to be updated in the current cycle are ordered according to the traveltimes of the nodes that are behind the current nodes in the previous cycle. The node that is connected to the node with the smallest traveltime in the previous cycle is updated first. In Figure 6.3A for instance, node n_A has the smallest traveltime of the nodes in the previous cycle, therefore node n_B will be evaluated first within the

current cycle.

According to the pattern of surrounding nodes for which already first arrival times have been computed, different types of finite differencing templates are being used (see for instance Abramowitz and Stegun, 1970).

Again, the ordering of nodes can be done extremely efficient through a heap stack algorithm. As the grid is evaluated through advancing and expanding rings, these type of methods are also referred to as upwind finite differencing schemes.

Opposite to shortest path algorithms, finite differencing schemes suffer from instabilities when cusps and caustics are being developed within the traveltime field. On the other hand, shortest path methods are restricted in their raypath definition due to the fixed structure of the stencils through the straight line connections accompanied with weight values between nodes.

The method proposed by Sethian and Popovici (1999), the fast marching method, combines the advantages of the shortest path methods with the finite differencing methods.

Like the shortest path methods, the nodes in the finite differencing grid are assigned to three sets, as indicated in Figure 6.3B; one set contains the nodes that have already been evaluated, the *upwind accepted values*, one set contains the nodes that still need to be evaluated, the *downwind far away values*, a third set contains the nodes within the so called *narrow band of trial values*. Then the heart of the algorithm reads,

[1] From the *narrow band of trial values*, the node with the smallest value is selected.

[2] The selected node is removed from the *narrow band of trial values* and added to the *upwind accepted values*.

[3] All neighbors of the selected node that are not yet part of the *upwind accepted values* or the *narrow band of trial values* are added to the *narrow band of trial values*.

[4] The values of the neighbors that have just been added to the *narrow band of trial values* are updated by an upwind finite differencing scheme.

[5] The scheme will continue at step one.

Crucial is the upwind finite differencing scheme. The scheme proposed by Sethian and Popovici (1999) assures that information propagates from small values to larger values and thus satisfies entropy and expanding waveforms. The selection of the smallest value within the narrow band to be evaluated first is consistent with the

upwind finite difference scheme, is equivalent to the shortest path method and can can be interpreted to be consistent with Huygen's principle. Again, the speed of the algorithm is given by the heap-stack method to select the nodes to be updated.

Most of further improvements to finite differencing schemes are related to a revised differencing template and to the allowance of a system re-entering. Finite differencing templates can be revised such that locally a curved wavefront or a plane wavefront is modeled. Also, higher order finite difference schemes can be used. Re-entering of the system could allow better for head waves or turning waves.

In Hegge (2001) an extensive overview and comparison of available methods is made. Based on his conclusions and the fact that near-surface problems are of a complex nature, in this thesis the Eikonal solver based on the algorithm of Vidale (1988) is preferred for the full traveltime inversion. The available implementation has a fast performance and exhibits a second order accuracy.

6.4 Parameterization of the velocity model

In the previous section the available forward modeling schemes were discussed. Maybe more important than the modeling scheme, is the parameterization of the velocity model, which is searched for by traveltime inversion. Within conventional refraction statics procedures two types of model parameterization are distinguished.

Cell-type models

Cell-type models split up the velocity model into cells. These cells can be rectangular or triangular for 2D models and cubic or tetrahedral for 3D models. When the model is build by using rectangular cells, each cell will have a constant velocity, which will be updated within each iteration. When the model is build by using triangular cells, each triangle has either a constant velocity, in which case the cell velocity is updated, or within each cell the velocity is determined by interpolation of the velocities as defined at the nodes. In the latter case the velocities at the nodes are updated. The model parameterized by nodal velocities has the advantage that the velocities are continuous across cell interfaces, whereas a velocity model built with constant velocity cells has a discontinuous velocity field across the cell interfaces.

Layered models

Instead of using cells as building blocks, the model can be build up by constructing layers. Often the earth, and especially the near surface can be approximated by layers. These layers can be very irregular, have pinch-outs and velocity gradients. Thus, layer i can be described by a set of geometrical nodes in combination with an interpolation function and a set of velocity nodes combined with an interpolation

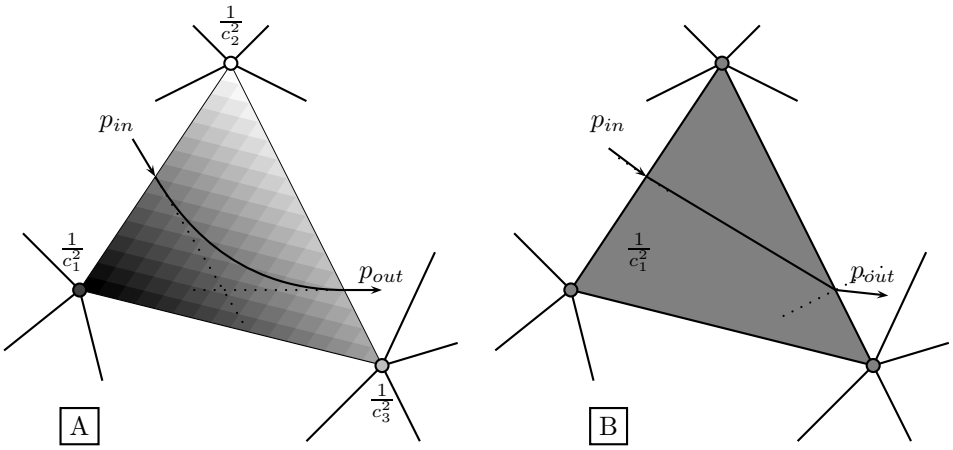

Figure 6.4: *Cell type models can be described by nodal velocities (A) in which case the velocities are continuous across boundaries and rays are describe by a curved trajectory within the cell, or the models are described by constant cell velocities (B) in which case the velocities are discontinuous across boundaries and the rays are described by straight lines within the cell.*

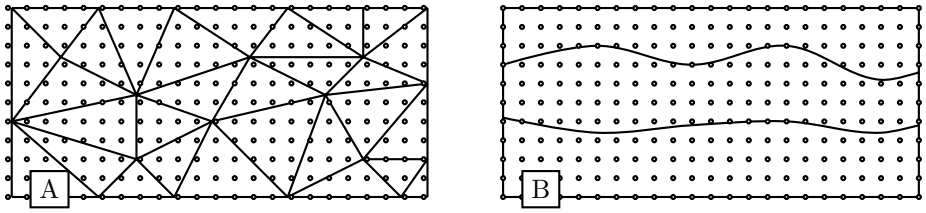

Figure 6.5: *Whether the velocities are described by cell-type (A) models or layered models (B), in case traveltimes are computed by an Eikonal solver, the models are first transformed into a gridded model.*

function,

$$z_i(x) = \sum_n z_n \psi_n(x), \tag{6.4.20}$$

$$c_i(x) = \sum_n c_n \chi_n(x). \tag{6.4.21}$$

Traveltime modeling and parameterization

Especially when a ray-tracing method is used, cell-based model building has the big advantage that the ray within a cell can be analytically traced, and the position of the ray at the cell edge is thus easily obtained. When the velocity within the cell is described by an interpolation of the velocities at the nodes, at the edge of the cell the ray parameters are continuous such that the rayparameters can be directly determined in the next cell (see Figure 6.4A). The computational effort has to be made in determining the exit point as the ray describes a curved path through the cell. In case the velocity within a cell is described by a single value, the velocity will be discontinuous at the edges of the cell. Therefore, the rayparameters within the new cell need to be computed by Snell's law at the exit point (see Figure 6.4B). As the velocity is constant within the cell, computing the ray within the cell will be less complex. For large models, the number of cells will be larger than the number of nodes. Thus, describing a cell model with nodal velocities will lead to less parameters compared to a cell model described by cell velocities. Cox (2004) has developed a general tomographic inversion method for operator traveltimes based on cell-type methods which uses a criterion based on ray-densities to determine how much detail the velocity model needs locally.

In case of a layered model the determination of the location where a ray crosses the layer boundary is not straight forward. The location needs to be searched for iteratively. A fixed geometrical step is taken along the ray. When the step length takes the ray into the next layer, the step length is halved and taken again until the ray is within an acceptable distance from the interface at which point the boundary conditions are applied to take the ray into the next layer. During the ray tracing, the algorithm continuously needs to keep track in which layer the ray is traveling to be able to apply the correct velocity description. For two-way raytracing, however, the layered model has the advantage that reflection points can be easily incorporated. To include reflection points within cell-based models, additional bookkeeping needs to be taken into account.

Since in this thesis, a grid-based method is used to compute the traveltimes, we do not need to consider the effects the parametrization has on tracing rays. The

velocity model is continuously defined by either two methods, cell based or layer based, and discretized into a gridded velocity model to be used for computing the traveltimes.

By parameterizing the velocity model through a layered system, a less smooth model can be build with the same amount of parameters, such that we are better able to cover the higher frequencies within the traveltime operators. The disadvantage is that a larger a priori imprint is introduced by the choice of the number of layers and velocity nodes within the layers. Dispite this disadvantage the layered parameterization will be adopted for our inversion algorithm. As interpolation function the cubic spline is used.

6.5 Comparison of raytracing and grid-based methods

As explained, forward modeling of traveltimes using grid-based methods is fast and very efficient. Compared to ray tracing, the traveltimes computed by the grid-based methods are directly available at the desired receiver locations, whereas for raytracing the starting or shooting parameters have to be optimized to assure the endpoint of the rays to reach the desired locations. As becomes clear from Figure 6.6, finding the proper endpoints can become a highly nonlinear and unstable task. This is especially the case for complex model problems in which anomalous features can rapidly change; a slight change in the shooting direction could lead to a big change in the endpoint location, which in system theory could be described as a poorly conditioned system.

On the other hand, as the traveltime-inversion problem is non-linear in itself, multiple linearized iterations have to be made in order to find the proper parameters describing the near-surface velocity model and the proper values of the CFP-locations. For each iteration the Jacobian, $\mathbf{J}$ has to be evaluated. The Jacobian comprises the traveltime derivatives with respect to the model parameters. In case the traveltimes are modeled by using a raytracing algorithm, the ray connecting CFP location i and surface location j is known throughout the model and, therefore, in case of a cell-velocity model, the time-derivative with respect to the velocity of the k^{th} cell of in total K cells the ray is traveling through can be easily accomplished,

$$t_{ij} = \sum_{k=1}^{K} l_k p_k, \tag{6.5.22}$$

$$\frac{\partial t_{ij}}{\partial p_k} = l_k, \tag{6.5.23}$$

in which p_k is the slowness, reciprocal of velocity, of the k^{th} cell and l_k the length of the ray within the k^{th} cell. For more advanced nodal-velocity model descriptions, similar type of expressions can be found.

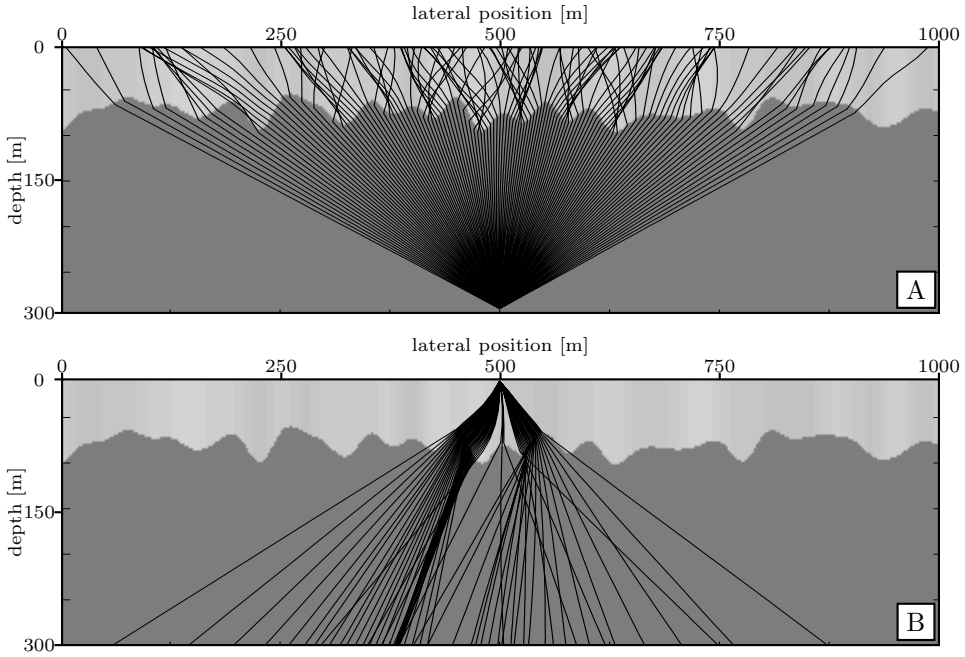

Figure 6.6: *Due to the fast variations in the near surface ray-tracing suffers from caustics, resulting in unstable two-point ray-tracing solution. Figure (A) shows rays modeled through a complex near surface from a common focus point to surface locations. Figure (B) shows rays through the same model from one surface location to a set of common focus points.*

In case of using grid-based methods to solve the Eikonal equation, the raypath is not known on forehand and, therefore, the time-derivatives have to be computed by finite differencing, for which additional forward runs are necessary,

$$\frac{\partial t_{ij}}{\partial c_k} \approx \frac{t_{ij}(c_k + \Delta c_k) - t_{ij}(c_k)}{\Delta c_k}, \frac{\partial t_{ij}}{\partial z_k} \approx \frac{t_{ij}(z_k + \Delta z_k) - t_{ij}(z_k)}{\Delta z_k}, \qquad (6.5.24)$$

in which Δc_k are small perturbations to the velocity value and depth values at node k, respectively (as defined by Equation 6.4.20 and Equation 6.4.21). These perturbations are usually taken to be about 5% of the current value.

Alternatively, the raypaths can be determined through one additional forward modeling; by combining the traveltimes throughout the gridded model of a forward

modeling step with a source at CFP location i and of a forward modeling step with a source at surface-location j, the raypath between location i and location j is found to be the path following the minimum traveltime through the model of combined traveltimes. This follows through combining Fermat's principle and reciprocity, as the raypath connecting the two points should be minimum and since reciprocity dictates that it should not matter whether the ray-path is computed from location i to location j or vice versa. In practice, the traveltimes are computed by placing a source at each surface-location and extracting the traveltimes from the model at the current CFP locations. To be able to deduce the raypaths additional forward models have to be computed using each CFP location as source, which have to be stored in memory. In case we have to evaluate traveltimes between 500 surface locations and 500 CFP locations, within a model that is roughly 12.5 km wide and 500 m deep, we need an order of at least 125Mb of memory storage for each iteration apart from cumbersome computations to extract correctly the raypaths and model dependencies, if a grid spacing of 5m is used within the forward modeling (500 sources $\times(12500/5)\times(500/5)$). This can become impractical with current single node computer specifications. In case of 3D evaluations, the amount of needed memory can become an order of 10^6 larger.

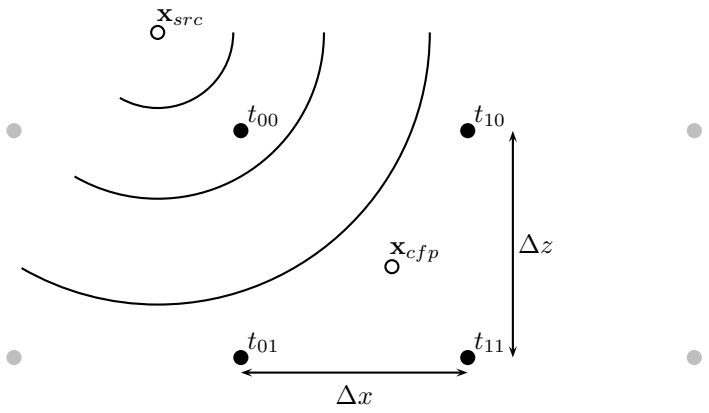

Figure 6.7: *Within a gridded velocity model, the time-derivatives with respect to the CFP-coordinates, $\mathbf{x}_{cfp}$ can be determined with the aid of the traveltimes t_{00}, t_{00}, t_{00} and t_{00}, as computed on the surrounding nodes by using a source at location $\mathbf{x}_{src}$.*

Computing the traveltime derivatives with respect to the CFP locations, using gridded methods, is less cumbersome and time consuming. As the traveltimes are known throughout the gridded model, the time-derivatives can be easily obtained through finite differencing in the current traveltime-grid, when traveltimes are modeled at

the CFP locations by putting sources at the surface locations. According to Figure 6.7 the traveltime derivatives with respect to the CFP locations can be determined through,

$$\frac{\partial t(\mathbf{x}_{src}, \mathbf{x}_{cfp})}{\partial x_{cfp}} \approx \frac{t_{11} - t_{01} + t_{10} - t_{00}}{2\Delta x}, \tag{6.5.25}$$

$$\frac{\partial t(\mathbf{x}_{src}, \mathbf{x}_{cfp})}{\partial z_{cfp}} \approx \frac{t_{11} - t_{10} + t_{01} - t_{00}}{2\Delta z}, \tag{6.5.26}$$

in which Δx and Δz are the grid spacings along the x-direction and z-direction respectively, within the gridded velocity model.

Given the instabilities and poor conditioning of the raytracing methods within complex near surface models, in this thesis the gridded methods are adopted, taking the additional forward modeling steps for granted.

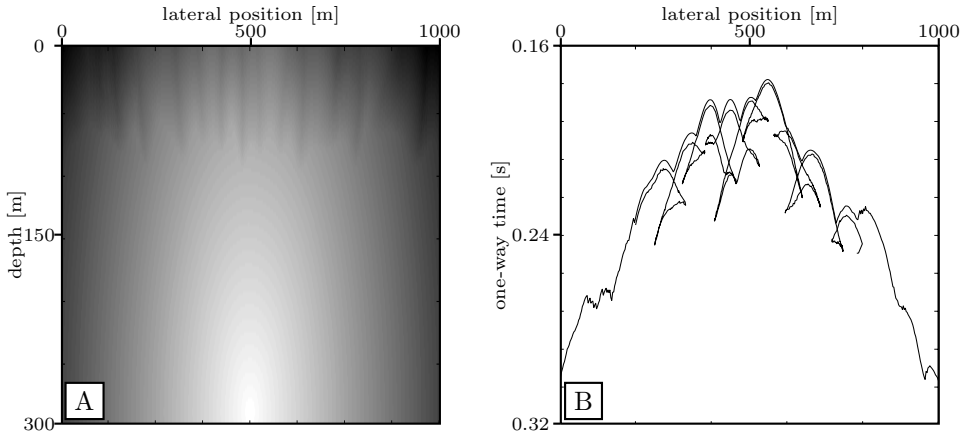

Figure 6.8: *(A) shows first arrival times computed by finite differencing. (B) displays first arrival times at the receiver locations computed by finite differencing and by using a ray-tracer. Ray-tracers can handle triplications but are unstable, where Eikonal solvers are stable but only are able to compute first arrival times. It is noted that the Eikonal solver is not exact, due to large discontinuous velocity contrasts.*

6.6 Travel-time inversion for a complex near surface, an example

The traveltime inversion for CFP-operators determined in a complex near surface is demonstrated with the aid of small example. Figure 6.10 displays a near surface model, consisting of two laterally undulating layers, having a laterally changing

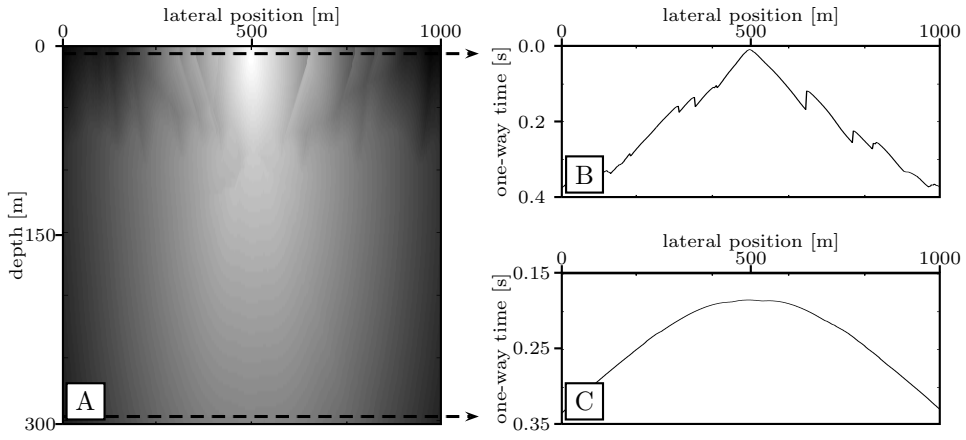

Figure 6.9: *By computing first arrival times from surface location to common focus point locations (C), using a finite difference Eikonal solver within a gridded velocity model (A), first break times are obtained as a byproduct (B).*

velocity as well. Around $x = -550$, a velocity inversion is noticed, the velocity in the top layer is higher than the velocity in the layer below. Within the model CFP-responses were modeled, arising from 101 CFP locations placed at a depth of 500 m. and equally spaced along the width of the model. The times were recorded at 101 surface locations, equally spread along the width of the model as well. The traveltimes were modeled using the Eikonal solver within a gridded model.

After applying the inversion algorithm, the model of Figure 6.11A was obtained within 6 iterations, using a cut-off error of 5%. It should be noted here that both the creation of the measured traveltimes as well as the forward computations within the inversion algorithm were established by using the same Eikonal solver. As can be clearly noticed, the CFP locations are resolved reasonably well although some structure is observed where it should be absent. However, the velocity profile is far off and no velocity inversion is observed at all.

In a second experiment, the first-break times were also included in the inversion, to further constrain the inversion process. The resulting model after again 6 iterations is displayed in Figure 6.11B. Again, the CFP-locations are resolved reasonably well at their correct locations. Compared to the inversion result which was obtained without the additional constraint by using the first-break times, the CFP locations do not display any structure at first sight.

Figure 6.11C compares the CFP locations obtained by both experiments in detail. The CFP locations obtained by inverting only the CFP operator times are denoted

by dots, whereas the CFP locations obtained by the combined inversion of CFP operator times and first-break times are denoted by stars. The exact lateral locations are indicated by the vertical dotted lines and the correct depth should be 500m. The combined inversion performs better in terms of resolving both the vertical locations as well as the lateral locations. Furthermore, the combined inversion resolves both near-surface layers better, although it still doesn't resolve the velocity inversion.

Figure 6.12 displays the resolution matrices, for both exercises. The formulation of the resolution matrix is given in Appendix B, and displays how well and independently the parameters are resolved. The model was described within the inversion algorithm by splines defining the two boundaries of the layers and the 2 lateral velocity profiles of the layer. Each of the in total 4 splines was determined by 10 vertices. Within the symmetric resolution matrices, the 40 parameters describing the two layers are found first along the 2 axes, next the single velocity defining the deeper layer is found, followed by the 101 lateral CFP locations and the 101 depth-values of the CFP locations. As can be concluded from Figure 6.12A, in case of inverting only for CFP operator-times a large number of crossterms is found, indicating that the parameters are not completely independently solved, a depth-velocity ambiguity might well be present. Figure 6.12B shows a better parameter resolution, when CFP operator times and first-break times are inverted in a combined fashion. Still crossterm energy is found, especially for the parameters describing the two layers. However, the cross-terms between the depth location and the near surface layers have strongly reduced, putting more confidence in the estimated depths.

It should be noted here that our main interest lies in properly determining the CFP-locations rather than estimating a correct velocity model as the idea is to describe the near surface in terms of propagation characteristics rather than by a velocity-depth model. In that sense we have established our task.

In literature, see e.g. Hampson and Russell (1984), it seems to be a matter of choice on how many layers and whether we should change velocity laterally or the layer outline. In most static procedures a layered model is adopted, which seems to describe near surface velocity variations quite well in most cases. The majority of conventional statics procedures however is based on *refraction* data, whereas our method is based on *reflection* data. Still questions arises as how many layers should be parameterized, should velocities vary lateral, should boundaries vary in the lateral direction or should both vary.

Tests with this type of flexible parameterization have already shown that the solutions tend to be unstable, among others due to the well known depth-velocity ambiguity (Lines, 1993; Al-Chalabi, 1997). As shown in the second experiment, by combining the CFP approach and inversion of first break times or other a priori information on the model, the solution can be stabilized. Through Bayes theorem different experiments are easily combined within an inversion procedure.

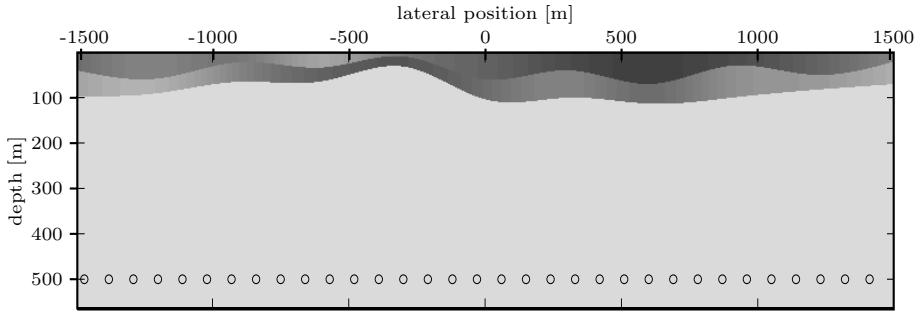

Figure 6.10: *Complex near surface model consisting of two layers with lateral varying velocity. CFP operators were modeled at equally spaced lateral positions at 500 m. depth. Note the velocity inversion around* $x = -500m$.

6.7 Conclusions

As shown in the previous chapters, the Common Focus Point method is well capable of determining the propagation characteristics of the near surface. However, no information is gained on the actual depth and lateral locations of the focal point. For that matter the focusing operators should be input to an inversion procedure which determines the focal point locations in space.

The inversion scheme requires a forward modeling step to compute traveltimes. This is achieved by solving the Eikonal equation. To solve the Eikonal equation one can either use a raytracing algorithm or a gridded Eikonal solver. Ray tracing has the advantage that rays through the model are directly available to use for computing the traveltime derivatives needed to populate the Jacobian within the inversion scheme. When use is made of a gridded Eikonal solver, additional forward traveltime evaluations need to be made to compute the traveltime derivatives. This makes the algorithm more expensive.

On the other hand, a ray tracer can become unstable in the presence of complex anomalies whereas an Eikonal solver will behave more stable. For that reason an Eikonal solver has been used. Through further research one could improve in deriving derivatives by using a gridded Eikonal solver.

Parameterizing the subsurface model, is not a straight forward tasks. It remains a classic question on how many layers should be used, on how the velocity should behave within a layer, etc. This has been tackled in the work of Cox (2004). In

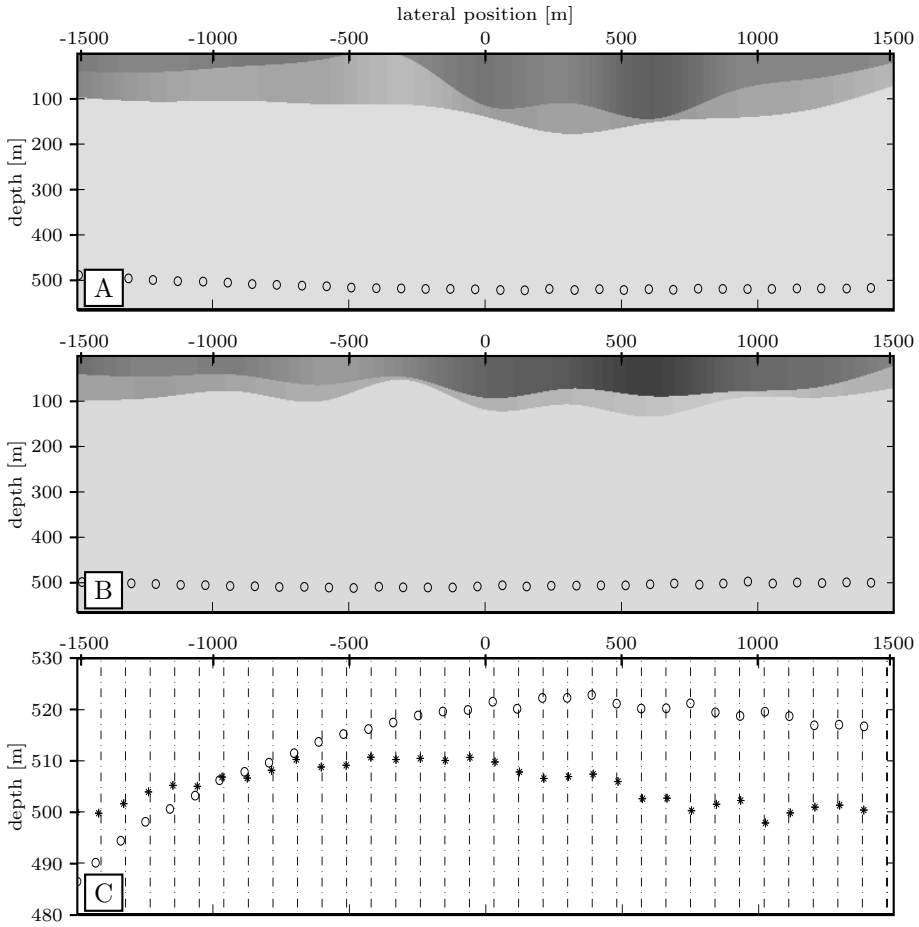

Figure 6.11: *(A) Model estimated by tomographic inversion of CFP operators. The dots denote the estimated CFP locations. (B) Model estimated by combined tomographic inversion of CFP operators and first break traveltimes. The dots denote the estimated CFP locations. (C) Estimated CFP locations. The dots denote locations retrieved by inversion of CFP operators only. The stars denote the locations retrieved by combined inversion of CFP-operators and first break travel-times. The vertical lines denote the correct lateral positions of the CFP's, and the correct depth should be* $500m$*. Note the exaggerated depth axis in (C).*

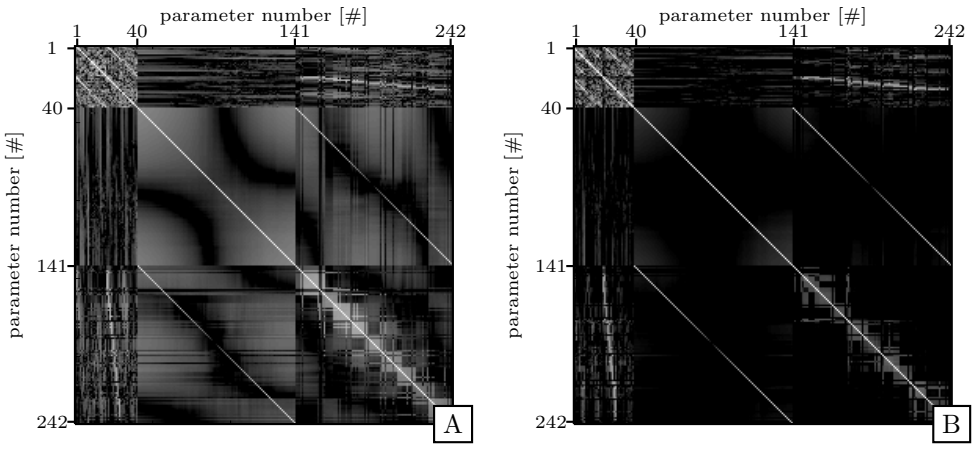

Figure 6.12: *(A) shows the resolution matrix of the inversion of CFP operators only. (B) shows the resolution matrix of the combined inversion of CFP operator times and first break times. The first 40 parameters correspond to the velocity and boundary nodes, the next 101 parameters correspond to the x-coordinates of the CFP locations and the last 101 parameters correspond to the z-coordinates of the CFP-locations.*

addition, especially when one needs to deal with low velocity subsurface models, the depth velocity ambiguity will play a role. Through Bayes' theorem multiple independent experiments can be combined such as both reflection data as well as refraction data that might be able to partially constrain the problem further.

7

True amplitude redatuming

7.1 Introduction

In this chapter the estimated traveltime operators are used to redatum the surface data to a level underneath the complex near surface. For this two strategies can be considered.

The first strategy uses the operators to redatum the data to the focused reflector. Since the focus-point locations can be unequally distributed on the reflector and since the reflector in general might not be a flat horizon at all, a second datuming step would be necessary from the reflector toward a flat datum with a user-defined background velocity. An estimated near surface model, as outlined in Chapter 6, would be needed to determine the CFP-locations, which are input to the second datuming step.

The second strategy uses the estimated velocity-model to precompute operators to a known, user-defined flat datum. Although the redatuming result would be directly interpretable, the redatumed data and associated resolution will be stronger model-driven as the original data-driven operators are no longer used. Any feature found within the data and data-driven updating procedure that could not be recovered during the model estimation will not be accounted for during redatuming.

We will mainly focus on the first strategy, under the assumption that the CFP-locations are known through, for instance, the inversion procedure outlined in Chapter 6.

For correct redatuming not only the propagation characteristics as given by the traveltimes are necessary, but also the energy characteristics, governed by the oper-

ator amplitudes. In this chapter a redatuming strategy will be outlined that correctly takes into account all amplitude characteristics.

From the operator time-tables we are able to determine the so-called background amplitudes. These are the amplitudes which are related to the geometrical spreading, or spherical divergence of the related wavefield. However, the operators only describe the first arrival primary energy related to reflector under consideration to which the data will be redatumed. As such, any amount of reflection energy arising from reflectors and scatterers in between the datum level and the surface is neglected. Thus we will not have a complete wavefield description of the near surface.

Reflection energy arising from reflectors will be mapped to the anti-causal part of the resulting data after redatuming and can therefore be easily removed.

Scattered energy arising from near surface scatterers, which is superposed on the recorded reflection energy, needs to be removed prior to redatuming by means of techniques as described by Blonk and Herman (1994), Ernst (1999), Campman et al. (2001).

By using the principle of data construction we are finally able to estimate transmission effects of the near-surface characteristics from the data by a linear inversion of a parameterized amplitude function.

Once the data has been corrected we can redatum the recorded data in a least-squares sense, by using the estimated background amplitudes and the transmission effects as operator amplitudes.

In this chapter, first the need of true amplitude redatuming is demonstrated, followed by the general outline of least-squares redatuming. In the following sections a step-by-step analysis is made of the different amplitude aspects, after which the full redatuming sequence is demonstrated with the aid of a synthetic example.

7.2 Least-squares redatuming

After subsequent iterations in the operator updating procedure, we end up with the traveltime tables, describing the kinematics of the near-surface. The aim of this thesis is to determine the wavefield as if it was excited by sources and recorded by receivers at a datum underneath the complex near surface. For that matter the goal is to redatum the data in a true amplitude sense to the datum for which the focusing operators were estimated kinematically. In this section we will discuss least-squares true amplitude redatuming. In the next three sections the several steps needed to fill the redatuming scheme are described.

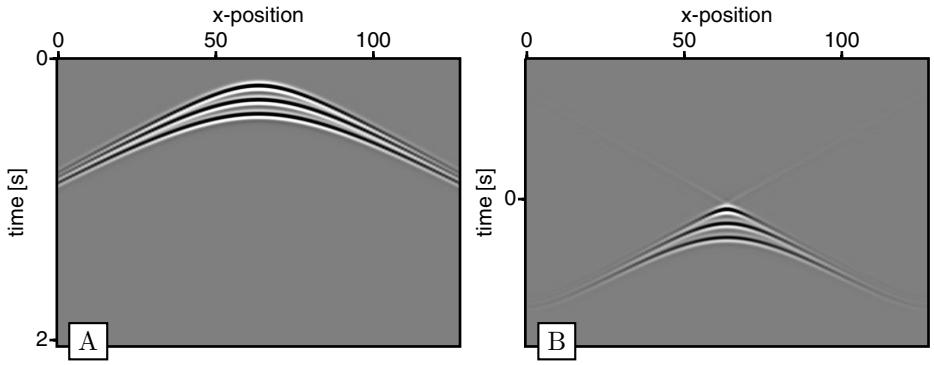

Figure 7.1: *If the redatuming operators are well known in terms of phase and amplitude we are able to perfectly redatum the recorded 3 reflection events of Figure A) toward the first reflector as displayed in Figure B)*

■ 7.2.1 Redatuming as two focusing steps

The principle of operator updating was based on the fact, that the operator used to construct a CFP-gather should align with the event under consideration in the CFP-gather. It was argued that the operator can be considered as the *incident wavefield* of the anti-causal impulse response resulting from a point source at the focusing location, measured at either the receiver locations or the source locations, depending on whether focusing in emission or focusing in detection was used. The constructed CFP-gather then comprises of traces which resemble the focusing of a receiver gather or source gather (again depending on the focusing type) and can be considered as the earth's response of the total wavefield recorded at the surface locations as result of a point source at the focus point. Thus, the CFP-gather can be considered as a first step in the redatuming procedure, where a virtual source is constructed at one particular focus point from a complete measurement at the surface. By defining a lateral set of focus points along the event of interest, a complete source distribution at the datum is constructed.

The second redatuming step propagates the receivers to the same common focus points, thus constructing a complete receiver distribution at the datum level for each virtual source. Note, that at time $t = 0$, for the redatumed shot-gathers, at zero-offset, the image of the particular focus-point is found. Furthermore, as a consequence of the redatuming procedure, the direct arrival found in the redatumed shot-gathers resembles the bi-focal image of the focus points (Berkhout, 1997b; van Wijngaarden, 1998), which contains the reflection properties of that focal point.

Therefore, the Green's functions, or focusing operators, should account for all prop-

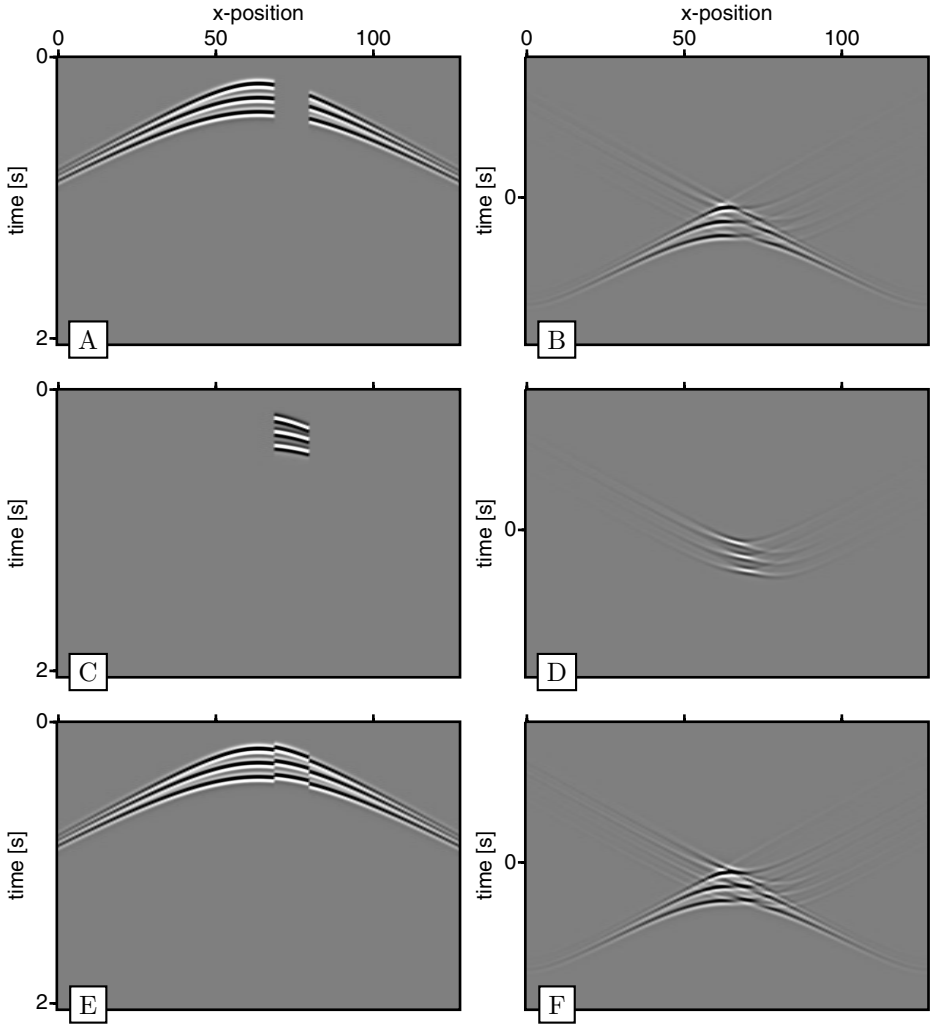

Figure 7.2: *The effect of amplitude- and time-perturbations act as wavefield-deficiencies. A) Shows the three events of Figure 7.1a, but now with an amplitude deficiency at traces 70-80. B) shows the result after redatuming of A). C) Shows the difference between Figure 7.1a and A). D) Shows the redatuming of C). E) Displays the data of Figure 7.1a with a time-deficiency at traces 70-80. F) Displays the data of E) redatumed to the first reflector.*

agation effects in the near surface. However, in practice the focusing process only accounts for the primary wavefield and any reflecting event in between the surface and datum level, as well as any scattering effect, is neglected, although they have their effect on the primary wavefield as well, e.g. energy losses, interference etc.

The focusing process was analyzed by inverse extrapolation, using the adjoint of the forward operator. For traveltime analysis this yielded correct results, based on the principle of equal traveltime. However, generally this approach will not hold for the amplitudes when inhomogeneous media are considered:

- For homogeneous media, the high frequency approach is exact. Amplitudes are correctly recovered.
- For slowly varying media, the high frequency approach will be exact when the medium parameters are correctly incorporated at the source and the receiver side.
- For inhomogeneous media, the focusing analysis within the redatuming procedure, as described in this thesis, only accounts for the most energetic arrivals. Thus, when for instance caustics or triplications are present, not all energy is properly redatumed.

Besides errors introduced due to the high frequency approximation of the focusing operators and the negligence of multipathing, redatuming will lead to incorrect results due to:

- Usage of the Rayleigh operators, when acquisition was done along a rugged terrain. Correct redatuming will only occur when the complete wavefield is available, thus including the normal derivative field, such that the full Kirchhoff integral is evaluated (see e.g. Wapenaar and Berkhout, 1989).
- By avoiding the construction of a subsurface model, any scattering of the focusing operator between datum and acquisition surface is neglected. Thus, the complete Green's function is unknown resulting in energy loss due to transmission, which is not taking into account due to the negligence of the scattered part.

■ 7.2.2 Effect of wavefield deficiencies on redatuming

Amplitude and time-perturbations, which are not accounted for within the operators used for redatuming, act as wavefield deficiencies.

In Chapter 3 it has already been shown how residual statics can severely degrade the resolution of the final stack and image. It has also been shown that if operators are available that correctly describe the near surface in principle we are able to reconstruct, at least kinematically, the wavefield at the user defined datum. If,

however, the operators do not fully describe the measured wavefield in terms of time and amplitude wavefield distortions will arise during redatuming. This is clearly demonstrated by the example as displayed in Figure 7.1 and Figure 7.2.

In Figure 7.1A a data-set is displayed consisting of 3 reflecting events embedded within an homogeneous background medium. As long as the operators are correctly defined in terms of time and amplitude we are able to obtain a perfectly redatumed data-set at the first reflector as shown in Figure 7.1B.

In Figure 7.2A the same dataset is displayed with an amplitude-disturbance between trace 70-80 (amplitudes are simply zeroed out). Using operators that do not account for these wavefield disturbances will return a redatumed data-set as displayed in Figure 7.2B, clearly displaying a degraded result, both in time as well as in amplitude. Taking the difference between the original clean data and the disturbed data, as shown in Figure 7.2C as input to a redatuming scheme results in the redatumed result of Figure 7.2D.

The disturbances in the redatuming result of Figure 7.2B compared to the desired result are exactly equal to the redatuming result of Figure 7.2D; the focusing operators used for redatuming have in fact been defocussing the wavefield disturbances which were not accounted for by the operator.

The same effects occur if deficiencies occur in terms of small time-disturbances as shown by Figure 7.2E and Figure 7.2F. Such effects have been noted before by Kabir and Verschuur (1995), which used this analogy as a base for their radon-domain based algorithm to fill-in missing near offset. Here it is emphasized that any deficiency not accounted for within redatuming will act as a secondary source which will defocus during imaging, migration or redatuming.

■ 7.2.3 Least-squares implementation of redatuming

For the moment it is assumed that the operators are fully described by the traveltime information and the derived amplitudes. In Chapter 4 the total surface free wavefield

before redatuming was described by Equation 4.4.9,

$$\mathbf{P}^{-}(z_0, z_0) = \mathbf{X}(z_0, z_0)\mathbf{S}^{+}(z_0), \tag{7.2.1}$$

$$\mathbf{X}(z_0, z_0) = \sum_{m=1}^{M} \mathbf{W}^{-}(z_0, z_m)\mathbf{R}(z_m, z_m)\mathbf{W}^{+}(z_m, z_0), \tag{7.2.2}$$

$$\begin{aligned}\mathbf{X}(z_0, z_0) = &\sum_{m=1}^{d} \mathbf{W}^{-}(z_0, z_m)\mathbf{R}^{+}(z_m, z_m)\mathbf{W}^{+}(z_m, z_0) \\ &+ \mathbf{W}(z_0, z_d)\mathbf{X}(z_d, z_d)\mathbf{W}(z_d, z_0).\end{aligned} \tag{7.2.3}$$

The overburden response contains all energy arising from reflections between the surface and the reflector ($\mathbf{R}^{+}(z_d, z_d)$) we have assigned to redatum to. Note that after redatuming these events will arise within the anti-causal part of the data and can therefore be easily separated and removed from the part of the data we are interested in.

Then the causal part of the redatumed data can be written as,

$$\mathbf{P}(z_d, z_d) = \left[\mathbf{R}^{+}(z_d, z_d) + \mathbf{X}(z_d, z_d)\right] \mathbf{S}^{+}(z_0), \tag{7.2.4}$$

under the assumption that $\mathbf{S}^{+}(z_0)$ is a diagonal matrix. The redatumed data of Equation 7.2.4, contains a reflection event $\mathbf{R}^{+}(z_d, z_d)$. This event is the event we were targeting our focusing operators on, as within the original data this target event would manifest itself as,

$$\acute{\mathbf{P}}(z_0, z_0) = \mathbf{W}(z_0, z_d)\mathbf{R}^{+}(z_d, z_d)\mathbf{W}(z_d, z_0)\mathbf{S}^{+}(z_0). \tag{7.2.5}$$

After redatuming, this particular event can be interpreted as energy traveling along the boundary with both a causal as well as an anti-causal part of which the zero-offset term (after redatuming) resembles the imaged energy for this particular sub-surface location. The remaining lateral cross-terms resemble the so called bifocal images (Berkhout, 1997a). From this pseudo first-arrival event the Amplitude Versus Ray-parameter can be derived (Berkhout et al., 2001; van Wijngaarden, 1998; Winthaegen and Verschuur, 2003).

As demonstrated in Chapter 4, Section 4.5.2, the operator used for redatuming, $\mathbf{W}(z_0, z_d)$, will contain amplitude anomalies caused by any change in transmission (giving rise to reflections as contained within the overburden response) between surface and datum-level. If the matched filter version of $\mathbf{W}(z_0, z_d)$ is used for redatuming, the amplitude distortions within the final results can be significant.

Therefore, instead of the matched filter approach which is sufficient for the purpose of deriving the traveltimes, we apply a least-squares approach with the aid of

the previously defined forward operators,

$$\mathbf{X}(z_d, z_0)\mathbf{S}(z_0) = [\mathbf{W}^H(z_d, z_0)\mathbf{W}(z_d, z_0) + \epsilon\mathbf{I}]^{-1}\mathbf{W}^H(z_d, z_0)\mathbf{P}(z_0, z_0), \tag{7.2.6a}$$

$$\mathbf{X}(z_d, z_d)\mathbf{S}(z_0) = [\mathbf{W}^H(z_d, z_0)\mathbf{W}(z_d, z_0) + \epsilon\mathbf{I}]^{-1}\mathbf{W}^H(z_d, z_0)\left[\mathbf{X}(z_d, z_0)\right]^T\mathbf{S}(z_0). \tag{7.2.6b}$$

In the first redatuming step, Equation 6.2.6a, the receivers are redatumed to the focused reflector, resulting in measurements of a wavefield excited at the surface and measured at the datum. In the second redatuming step, Equation 6.2.6b, the sources are redatumed to the focused reflector, resulting in measurements of the wavefield excited by sources at the datum and recorded at the datum as well.

Often, computing the inverse is unstable due to the filtering effect of the operator and thus loss off high angle information (resulting in a large null space). Similar to least-squares migration, a fair and stable estimate is

$$\mathbf{X}(z_d, z_0)\mathbf{S}(z_0) = \left[\mathrm{DIAG}(\mathbf{W}^H(z_d, z_0)\mathbf{W}(z_d, z_0) + \epsilon\mathbf{I})\right]^{-1}\mathbf{W}^H(z_d, z_0)\mathbf{P}(z_0, z_0), \tag{7.2.7a}$$

$$\mathbf{X}(z_d, z_d) = \left[\mathrm{DIAG}(\mathbf{W}^H(z_d, z_0)\mathbf{W}(z_d, z_0) + \epsilon\mathbf{I})\right]^{-1}\mathbf{W}^H(z_d, z_0)\mathbf{X}^T(z_d, z_0), \tag{7.2.7b}$$

with the notion that $\mathbf{W}^H(z_d, z_0)$ is close to the inverse of $\mathbf{W}(z_d, z_0)$ in terms of phase. In case of a homogeneous medium, for which $\mathbf{W}^H(z_d, z_0) = \mathbf{W}^{-1}(z_d, z_0)$, indeed Equation 6.2.7a and Equation 6.2.7b reduce to the matched filter solution.

■ 7.2.4 Example of least-squares redatuming

The importance of including amplitudes during redatuming through a strongly varying near surface is demonstrated in the following example. In Figure 7.3A a shot-gather is displayed which shows the response of three horizontal reflectors. The horizontal reflectors have been modeled as unit valued, angle independent reflection coefficients. Through a laterally varying inhomogeneous near-surface layer the forward propagator, $\mathbf{W}(z_0, z_d)$, was modeled by first arrival ray-tracing. To obtain the data, first at $z = z_d$ the response was modeled of two deeper horizontal reflectors and the dipole response of z_d itself. As such we have,

$$\mathbf{X}(z_d, z_d) = \mathbf{R}^+(z_d, z_d) + \mathbf{W}(z_d, z_1)\mathbf{I}\mathbf{W}(z_1, z_d) + \mathbf{W}(z_d, z_2)\mathbf{I}\mathbf{W}(z_2, z_d), \tag{7.2.8}$$

which is displayed in Figure 7.3B. Using the forward modeled operators, the forward modeled surface data is obtained through

$$\mathbf{P}(z_0, z_0) = \mathbf{W}(z_0, z_d)\mathbf{X}(z_d, z_d)\mathbf{W}(z_d, z_0)\mathbf{S}(z_0), \tag{7.2.9}$$

and it is assumed that the first arrival as well as any other scattering has been perfectly removed. Through this forward modeling procedure it is assured that the

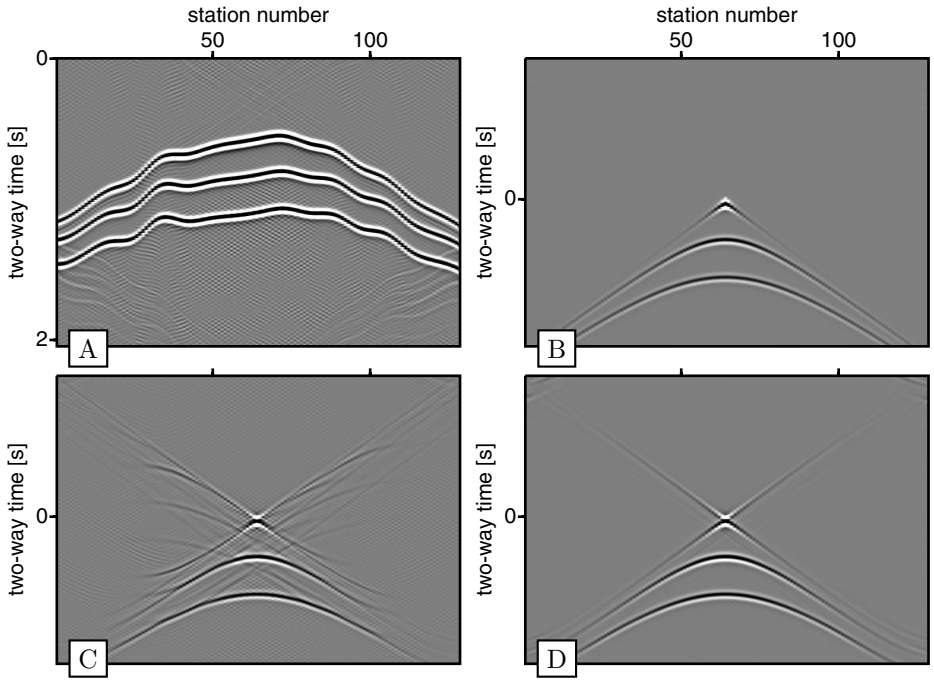

Figure 7.3: *Redatuming of noise free data. The amplitudes of the measured data in A) only rely on spherical divergence and are consistent with the inverse operators. Artifacts occurring after matched-filter redatuming in C) are resulting from non-hyperbolic move-out only whereas least-squares redatuming in D) renders the desired data of B).*

measured data and available operators are consistent. The desired data after redatuming equals the forward modeled reflections of Figure 7.3B. Through a matched filter inverse extrapolation the data as displayed in Figure 7.3C is obtained. Although the redatuming result is kinematically accurate, the near surface inhomogeneities result in amplitude and phase artefacts. If, however, the least-squares approach as described in Equation 7.2.7a is used to redatum the data, clearly the data is well recovered at z_d, as can be observed in Figure 7.3D.

If additional amplitude deficiencies are added to the measured data, in this case surface consistent random variations, as displayed in Figure 7.4A, the matched filter approach for redatuming becomes a little bit worse, if these deficiencies are not accounted for during redatuming within the used operators as displayed in Figure 7.4C. If a least-squares solution is used for redatuming without accounting for these deficiencies, the result will become even worse as displayed in Figure 7.4D[1]. The

[1]It should be noted here that significant dip-filtering and fk-filtering had to be applied to obtain the displayed result, which has not been the case for the other results.

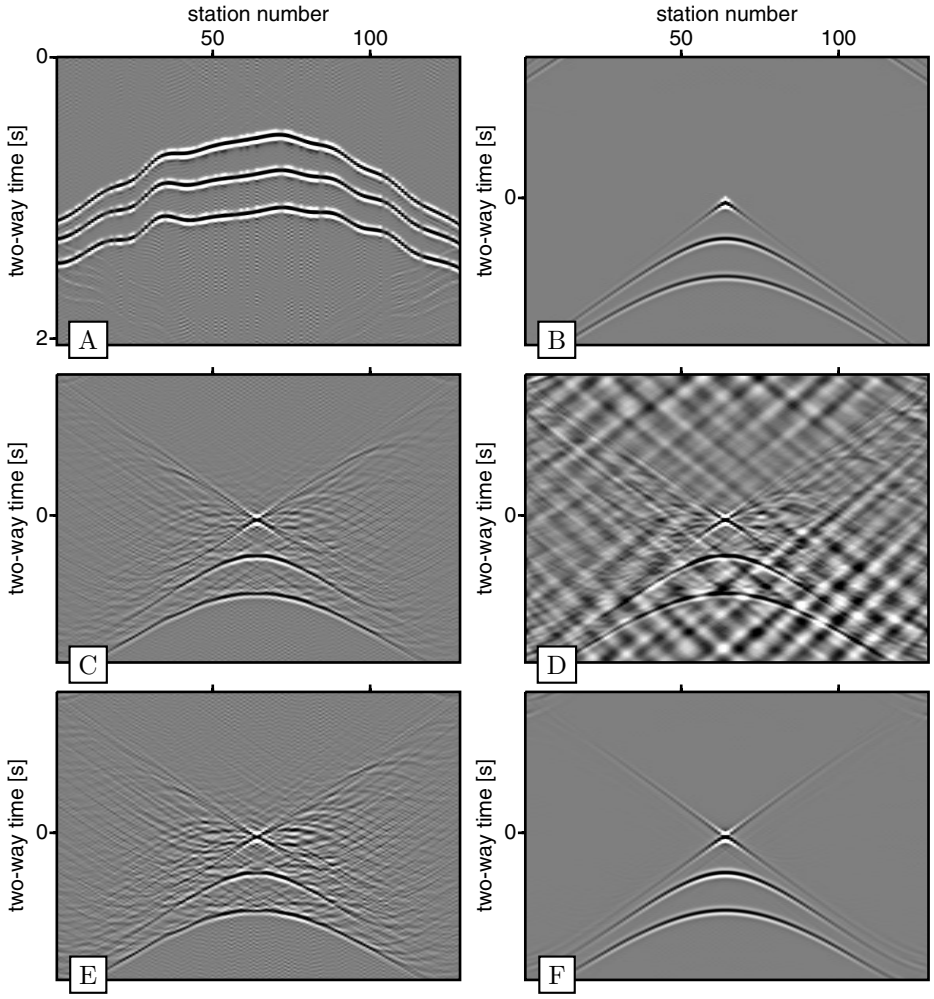

Figure 7.4: *Redatuming of data with amplitude noise included (A). C) shows the result after matched filter redatuming without attempting to include the amplitude anomalies within the operators. D) shows the result after least-squares redatuming without including the amplitude anomalies. E) shows the result after matched-filter redatuming taking into account the amplitude anomalies. F) displays the result after least-squares redatuming including amplitudes. F) clearly renders the desired data of B).*

matched filter solution incorporating the amplitude deficiencies within operators does also not give a better result as displayed in Figure 7.4E; the artifacts are more pronounced compared to Figure 7.4C. However, through least-squares redatuming using operators that do account for the deficiencies will render a clean result as displayed in Figure 7.4F.

The example clearly shows the importance of including the correct amplitudes when redatuming through a complex near surface. However, if amplitudes are treated incorrectly the redatumed result will be worse compared to a redatuming without taking amplitudes into account.

7.3 True amplitude redatuming strategy

Within this section, a step by step procedure is outlined to treat the amplitudes in a correct way, in which we start with a number of preprocessing steps,

- retrieve from the data the amplitudes of the reflector for which the traveltime operators were defined. This can be accomplished by using *Fermat modeling* as outlined in Chapter 4, Section 4.6.2
- Remove constructive near surface scatter energy from the primary data as outlined within Campman et al. (2001).
- Remove any residual surface consistent time anomalies through time decomposition routines as proposed by for instance Cambois and Stoffa (1993).
- Remove any residual surface consistent amplitude anomalies through amplitude decomposition routines as proposed by for instance Cambois and Stoffa (1992).

The preprocessed data will be decomposed into 2 sources of amplitude contributions,

- Determination of the background amplitudes, arising from spherical divergence.
- Estimation of operator amplitudes, which can not be explained by spherical divergence. Any remaining energy in the data has to be attributed to losses due to for instance transmission effects and should be included in the focusing operators. These contribution will be derived through parameterized inversion from the primary data.
- Applying a least-squares redatuming process to the amplitude corrected data, using the focusing operators including their estimated amplitudes. For this purpose the procedures as described in Section 6.2.3. will be used.

As the residual surface consistent amplitude decomposition has a close resemblance to the operator amplitude estimation procedure, they will be discussed together in Section 6.3.2.

■ 7.3.1 Preprocessing for least-squares redatuming

before we are able to estimate the amplitudes of the redatuming operators, we first need to preprocess the data. This includes extracting the measured amplitudes from the recorded data, which is used in a cost function to be minimized in order to retrieve the transmission effects to be used in the operators. Also, before we can successfully carry out a redatuming scattered multiplicative noise which obscures the primary energy of the events of interest, as well as any surface consistent time and amplitude residual should be removed from the data.

Detection of measured amplitudes

To correct the amplitudes within the measured data and to estimate the operator amplitudes we need to define the input amplitudes which are subject to the proposed algorithms. As already outlined within section 6.2 it is of importance that any processing applied to define the operator amplitudes should be consistent with the measured data. Through Fermat modeling we are able to produce the traveltimes of exactly those amplitudes within the data that are influenced by the redatuming operators.

Since the phase characteristics of the focusing operators are already defined through the outlined updating procedures, we are able to construct the two-way traveltimes of the reflector under examination through a stationary phase analysis, see Chapter 4, Section 4.6.2.

From the one-way operator times which were derived through updating as displayed in Figure 7.5A the two-way traveltimes can be derived which are shown in Figure 7.5B. If these two-way traveltimes are plotted within two of the 301 shot-gathers that were used for the focusing analysis, clearly the two-way traveltimes overlay the event for which the focusing operators have been estimated, as can be observed in Figure 7.5C and D. By subsequent muting around the event and correcting the data with the computed two way traveltime we are able to easily subtract and filter the amplitude data of interest as displayed in Figure 7.5E and F.

Removal of backscattered noise

One of the contributions to near surface amplitude anomalies arises from near surface scatterers. During downward propagation from source to target level, energy arising from these scatterers will map within the anti-causal part of the data after redatuming. However a significant scatter energy distribution is introduced during upward propagation toward the receivers; scattered energy will be induced by the reflected wavefield just before arriving at the recorders. In Campman (2005), near surface effects are fully subscribed to the scattering effects of small near surface inhomogeneities in terms of both phase as well as amplitudes. Thus, the total wavefield,

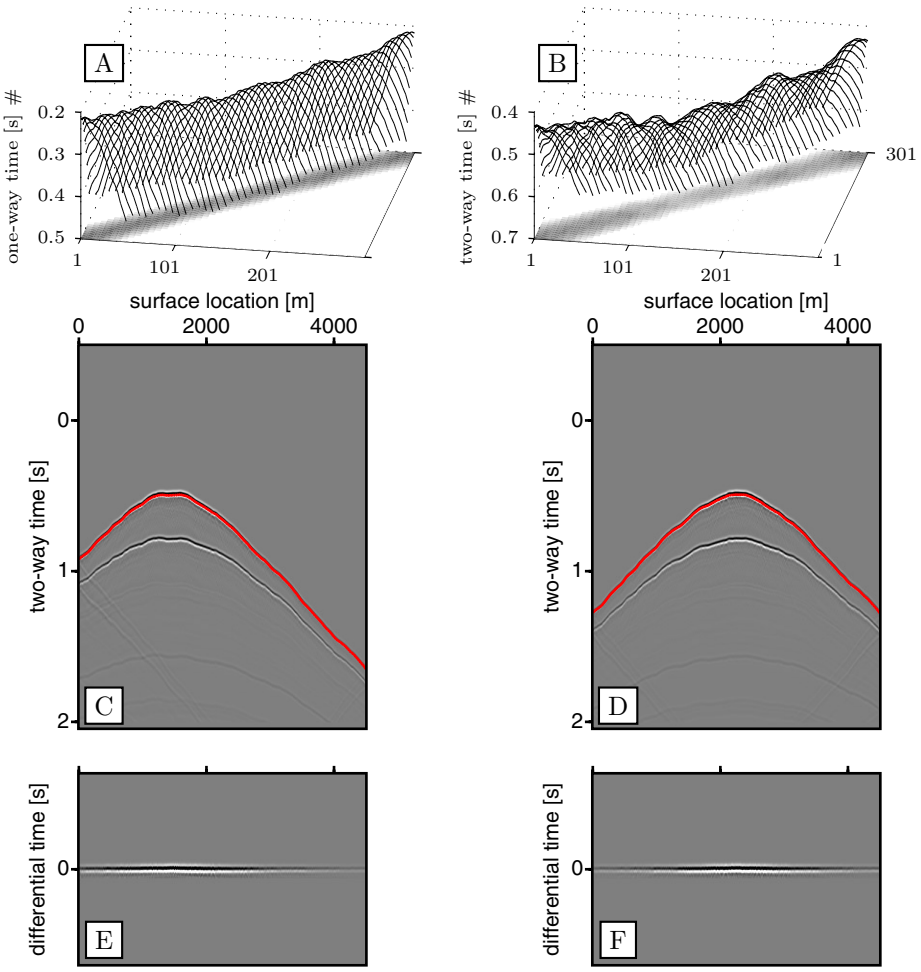

Figure 7.5: *From the one-way traveltimes as displayed in A), obtained through operator updating, the two-way traveltimes can be obtained through the principle of data reconstruction as displayed in B). These two-way times perfectly overlay the event of interest within the shot-gathers as displayed in C) and D). By correcting with the computed two-way traveltime and subsequent muting the amplitudes for this particular event can easily be retrieved for subsequent filtering and processing as displayed in E) and F).*

$\mathbf{P}_t(\mathbf{x}_{rcv}, \mathbf{x}_{src})$, can be described by Equation 4.3.72 as,

$$\begin{aligned}\mathbf{P}_t(\mathbf{x}_{rcv}, \mathbf{x}_{src}) = \mathbf{G}(\mathbf{x}_{rcv}, \mathbf{x}_{src})\mathbf{S} + \int_{\partial\mathbb{D}} \mathbf{G}(\mathbf{x}_{rcv}, \mathbf{x}')\mathbf{P}_t(\mathbf{x}', \mathbf{x}_{src})\mathrm{d}\mathbf{x}' \\ + \int_{\mathbb{D}} \mathbf{G}(\mathbf{x}_{rcv}, \mathbf{x}')\mathbf{B}(\mathbf{x}')\mathbf{P}_t(\mathbf{x}', \mathbf{x}_{src})\mathrm{d}\mathbf{x}',\end{aligned} \tag{7.3.10}$$

in which $\mathbf{G}$ is the Green's function and $\mathbf{B}$ is the contrast operator as described in Chapter 4, related to the near surface inhomogeneities. The first term on the right-hand side resembles the direct wave-field, the second term describes the near surface scatter free reflection data and the last term resembles the energy that arises from near surface scattering of the reflected data. In his method it is assumed that the inhomogeneities are lying within the vicinity of sources and receivers, such that the wavefield impinging on the near surface scatterers can be approximated by the measured wavefield. The measured scattered wavefield is derived by filtering one strong event in the wavenumber domain. If only the wavefield scattered from the near surface inhomogeneities is considered, the direct wavefield, $\mathbf{GS}$ and the part of the wavefield not affected by the inhomogeneities traveling through the background, denoted by the surface integral, vanish under a first order approximation. Under these considerations, the scattered wavefield, measured at the surface is described by

$$\mathbf{P}_s(\mathbf{x}_{rcv}, \mathbf{x}_{src}, z = z_0) = \int_{\mathbb{D}} \mathbf{G}(\mathbf{x}_{rcv}, \mathbf{x}')\mathbf{B}(\mathbf{x}')\mathbf{P}_t(\mathbf{x}', \mathbf{x}_{src}, z = z_0)\mathrm{d}\mathbf{x}'. \tag{7.3.11}$$

By choosing the coupling part of the medium operator for the Greens state equal to zero and the propagating part equal to the true medium operator, Equation 7.3.11 can be written as,

$$\mathbf{P}_s(\mathbf{x}_{rcv}, \mathbf{x}_{src}, z = z_0) = \int_{\mathbb{D}} \mathbf{G}(\mathbf{x}_{rcv}, \mathbf{x}')\mathbf{\Phi}(\mathbf{x}')\mathbf{P}_t(\mathbf{x}', \mathbf{x}_{src}, z = z_0)\mathrm{d}\mathbf{x}'. \tag{7.3.12}$$

Within this formulation the only unknown is the contrast operator, if the background medium is assumed to be known. After discretization, the integral can be rewritten as,

$$\mathbf{D} = \mathbf{K}\sigma, \tag{7.3.13}$$

in which $\mathbf{D}$ is a column vector in which all the seismic experiments are contained and σ is contains the unknown contrasts. Campman (2005) then estimates the scattering distribution through a least-squares inversion. The used Green's operators are modeled in a laterally homogeneous medium, which allows for a fast computation of the forward model in the wavenumber number for a limited number of depth levels on which the scatter distributions are projected. In this way, for each surface location a scatter image can be constructed, which is frequency dependent and can be employed for a limited number of depth locations. After imaging the scatterers for the selected event, the images are used to predict the scattered energy for the remainder

of the recorded data, by using the forward model of Equation 7.3.12 on the total recorded wavefield. Due to the assumption that the scatterers are residing in the close vicinity of the recording stations, the method is surface consistent and thus has a close resemblance with residual statics and surface consistent deconvolution. The difference lies in the fact that opposite to standard residual statics and surface consistent deconvolution, the method described is a multi-trace procedure. It should be

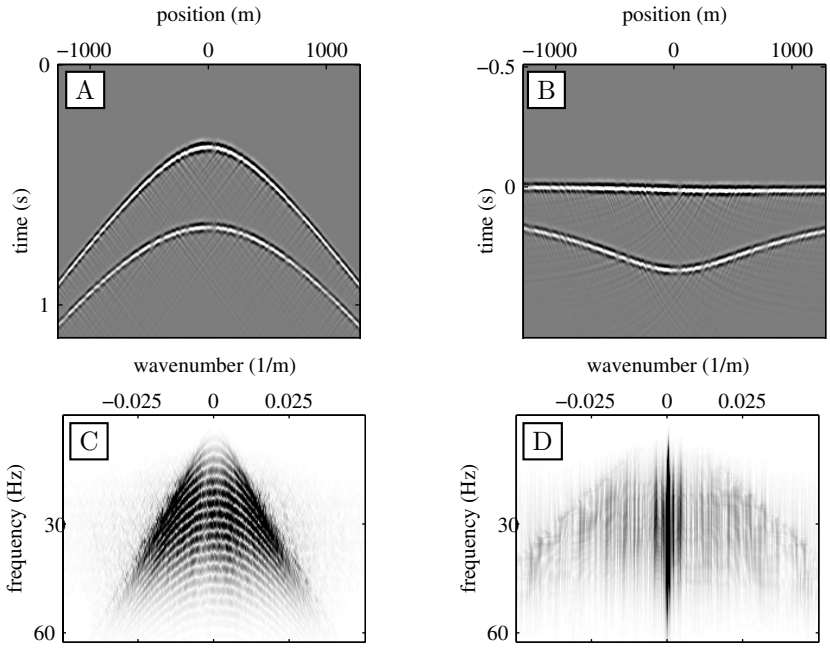

Figure 7.6: *The original shot-gather in A) is contaminated by near surface diffraction energy. Its wavenumber spectrum in C) clearly is clearly distorted by the near surface diffraction energy. By correcting the first event with the precomputed two-way traveltimes, as displayed in B) we are able to separate the reflection energy, aligned around wavenumber zero, from the near surface diffraction energy as observed in D).*

noted here, that the assumption of a 1D medium does not hold for longer wavelength near surface effects as is dealt with in this thesis. But, since the scatter distribution is projected at or close to the surface, the assumption can still hold if near surface irregularities can mainly be described by an undulating base of the near surface layer.

Furthermore, the estimation of the scattered field used in the scatter distribution prediction step is retrieved by filtering a selected event in the wavenumber domain, under the assumption that the selected event has a linear move-out, for instance the

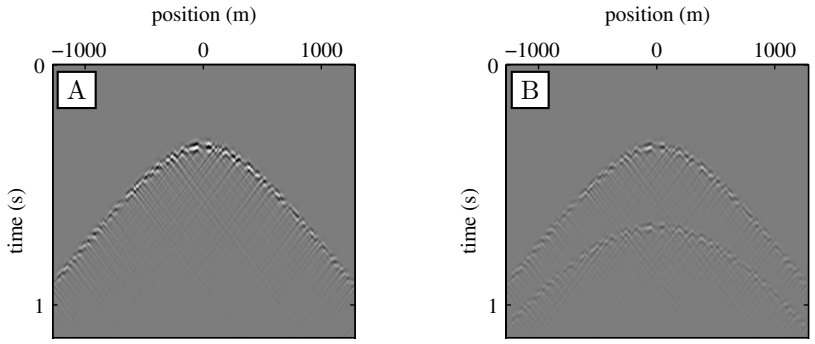

Figure 7.7: *The separated near-surface diffraction energy in A) is used to estimate a near surface scatter distribution which can be used to estimate the diffraction energy for all other arrivals (B).*

Rayleigh wave or spatially dense sampled experiments with a small maximum aperture. Within the problems described in this thesis, the move-out is far from close to linear. However, using the found operators, the total move-out of the selected event can be synthesized to fully flatten the event under investigation. In this sense the methodology of this thesis and the methodology of Campman (2005) can be fully complementary to tackle both the long wavelength as well as the short wavelength near surface effects.

A drawback is that the spatial sampling has to be designed such that the data is not aliased to properly separate the scattered data from the total data; within the proposed method, the scattered noise is band-limited through the use of the Green's operators in the background medium. Near surface problems, observed in the field are generally not band-limited, due to, among others, randomly behaving coupling effects of sources and receivers, extreme low velocity anomalies, mode conversions, etc.

The following example demonstrates the near-surface backscattering method. In Figure 7.6A, the input data is displayed, consisting of two events measured at 256 receivers for one seismic experiment, contaminated with near surface scattering effects, mainly influencing the amplitude characteristics. The wavenumber spectrum of the input data is clearly degraded, due to the near surface scatterers. It is also clear from the spectrum that both signal as well as noise are band-limited due to the fact that the noise and the clean data are both modeled according to the wave-equation with a background velocity of 1500 m/s. By computing the move-out of the first event, which could be done by using common focus-point operators, the data can be horizontally aligned as displayed in Figure 7.6B. Since the event is horizontally aligned, the noise free event can be removed by suppressing the energy around

$k_x = 0$, which was done by applying a hanning-window on the spectrum of Figure 7.6D. Next, the two-way traveltime corrections can be removed again.

The noise contaminating the first event, as displayed in Figure 7.7A, can then be used together with the full measured data to estimate a scatter distribution which was used to derive the full noise distribution, displayed in Figure 7.7B. By subtracting the estimated noise field in a least-squares sense, noise contaminating both events is suppressed, resulting in clean data with a proper wavenumber spectrum (Figure 7.8).

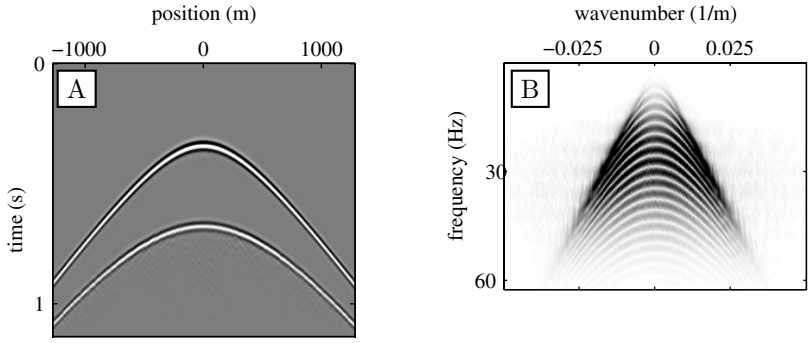

Figure 7.8: *The estimated near surface diffraction energy from Figure 6.7B can be adaptively subtracted from the data (A), which leads to a clean wavenumber frequency spectrum (B).*

Removal of residual time shifts

The operator updating procedure tends to stack out fast varying traveltime perturbations, conventionally known as residual statics. In practice these perturbations can originate from a varying range of sources, which are not necessarily consistent for each source and receiver pair. Therefore, besides loss of resolution, they can never be corrected for within the updating procedure. If not handled correctly, these residuals can leak through the spatial frequency band and hence degrade the quality of the data after focusing and redatuming.

Through the stationary phase approach of the previous section, the stationary two-way traveltimes can be extracted for each source and receiver pair. This stationary time can than be used to apply a move-out correction to each source and receiver pair. Besides, the two-way traveltime can be used as a time-window, to determine stack optimizing time shifts through procedures as described by Ronen and Claerbout (1985).

■ 7.3.2 Operator amplitude determination for least-squares redatuming

As was demonstrated in the previous section, for proper least-squares redatuming, good knowledge on the operator amplitudes is important. To account for these amplitudes we consider two types of amplitude contributions:

- Background amplitudes; these are equivalent to spherical divergence and will be incorporated directly within the operators.
- Residual amplitudes; these are caused by unaccounted transmission losses and coupling effects among others. Part of these effects will be removed prior to redatuming (surface consistent parts such as coupling effects), part will be incorporated within the operators (offset dependent transmission effect).

Background amplitudes

For slowly changing scattering-free media, the high frequency approximation of the propagators $\mathbf{W}^+$ and $\mathbf{W}^-$ and their inverses, $\mathbf{F}^+$ and $\mathbf{F}^-$ under the modified matched filter approach would be valid to use, reading

$$W_{ij}^{+} = \left[\frac{\rho_i c_i}{\rho_j c_j}\right]^{\frac{1}{2}} A_{ij} \exp\left(-j\omega\tau_{ij}\right), \qquad F_{ij}^{+} = \left[\frac{\rho_j c_j}{\rho_i c_i}\right]^{\frac{1}{2}} A_{ij} \exp\left(+j\omega\tau_{ij}\right), \quad (7.3.14)$$

$$W_{ij}^{-} = \left[\frac{\rho_j c_j}{\rho_i c_i}\right]^{\frac{1}{2}} A_{ij} \exp\left(-j\omega\tau_{ij}\right), \qquad F_{ij}^{-} = \left[\frac{\rho_i c_i}{\rho_j c_j}\right]^{\frac{1}{2}} A_{ij} \exp\left(+j\omega\tau_{ij}\right). \quad (7.3.15)$$

The operators are depending on the medium parameters at the source and receiver locations, the traveltime τ and an amplitude factor A_{ij}. Within the redatuming procedure, which is established by applying the inverse propagators at both sides of the data-equation of Equation 6.2.5, the effects of the medium at the source and receiver side are canceled and the amplitude factors assure orthogonality between the forward propagators and their inverses. The amplitude factors are also known as the geometrical spreading, which follows from solving the transport equation of Equation 5.3.13 and depend on the curvature in time of the wavefronts. Within ray-tracing schemes the geometrical spreading can be computed through paraxial ray-tracing (Cerveny, 1985) around the centrally traced rays.

Since, within the updating procedure a full set of traveltimes was determined between all surface locations and all focus-point locations, the spherical divergence can also be determined directly from the traveltimes, without any paraxial ray-tracing. An expression for the geometrical spreading can be found in Tygel et al. (1992),

$$\mathcal{A}(\mathbf{x}_s, \mathbf{x}_{cfp}) = \frac{1}{c(\mathbf{x}_s)} \sqrt{\frac{\cos(\alpha_s)\cos(\alpha_{cfp})}{\|N\|}}, \quad (7.3.16)$$

in which $\cos(\alpha_s)$ and $\cos(\alpha_{cfp})$ are the emergence and receiving angle respectively and $\|N\|$ is the determinant of the second order mixed derivatives of the traveltime

with respect to the source coordinates and the CFP-coordinates, projected at the acquisition plane and the reflector respectively, which reads in 2D space,

$$N_{ij} = \frac{\partial^2 t_{ij}}{\partial x_{cfp;i} \partial x_{src;j}}. \tag{7.3.17}$$

In Vanelle and Gajewski (2000) the determinant is related to a hyperbolic approximation around the arrival time around the central ray, called the hyperbolic paraxial time,

$$t^2(\mathbf{s},\mathbf{g}) = \left[t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) - \mathbf{ps} + \mathbf{qg}\right]^2 + t0(\mathbf{x}_{src},\mathbf{x}_{cfp}) \left[\mathbf{s}^T\tilde{\mathbf{N}}\mathbf{s} + \mathbf{g}^T\bar{\mathbf{N}}\mathbf{g} - 2\mathbf{s}^T\mathbf{N}\mathbf{g}\right], \tag{7.3.18}$$

in which $\tilde{\mathbf{N}}$ and $\bar{\mathbf{N}}$ are the second order derivatives of the traveltime with respect to $\mathbf{x}_{src}$ and $\mathbf{x}_{cfp}$ respectively, projected on the source and CFP sheet. $\mathbf{s}$ and $\mathbf{g}$ are the projected locations of the source and CFP-locations $\mathbf{x}_{src}$ and $\mathbf{x}_{cfp}$ respectively, which are used to define the paraxial approximation. $\mathbf{p}$ and $\mathbf{q}$ are the apparent slownesses around $\mathbf{x}_{src}$ and $\mathbf{x}_{cfp}$ respectively.

Through the updating sequence, traveltimes are derived between a full lateral set of CFP-locations and a full lateral set of surface locations. Thus, the matrices $\tilde{\mathbf{N}}$ and $\bar{\mathbf{N}}$ can be determined by comparing $t_0(\mathbf{x}_{src},\mathbf{x}_{cfp})$ with paraxial times for source and receiver locations centered around $\mathbf{x}_{src}$ and $\mathbf{x}_{cfp}$. With $\mathbf{s} = (0,0)^T$ and $\mathbf{g} = (g_1,0)^T$ for instance, Equation 7.3.18 reads,

$$\tilde{\mathbf{N}} = \frac{t^2(\mathbf{s},\mathbf{g}) - (t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) + q_1 g_1)^2}{t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) g_1^2}. \tag{7.3.19}$$

An equivalent relation can be found for $\bar{\mathbf{N}}$. The slownesses $\mathbf{p}$ and $\mathbf{q}$ are found by central differencing of the found traveltimes around the central rays. The obtained values for $\tilde{\mathbf{N}}$ and $\bar{\mathbf{N}}$ are then substituted into Equation 7.3.18 to arrive at an expression for $\mathbf{N}$, for which element N_{11} would read for instance,

$$N_{11} = \frac{\left[t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) - p_1 s_1 + g_1 q_1\right]^2 + t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) \left[s_1^2 \tilde{N}_{11} + g_1^2 \bar{N}_{11}\right] - t^2(\mathbf{s},\mathbf{g})}{2 t_0(\mathbf{x}_{src},\mathbf{x}_{cfp}) s_1 g_1}. \tag{7.3.20}$$

In Kästner and Buske (1999) it is stated that the resulting determinant, occurring in Equation 7.3.18, is less accurate when compared to calculations of the geometrical spreading by direct finite differencing schemes as described in Abramowitz and Stegun (1970). It should be noted here, however, that the traveltimes should then be known on the full subsurface grid, which can be accomplished by using Vidale's Eikonal solver over the known velocity field. The method of Vanelle and Gajewski (2000), however, relates the spherical divergence to properties projected on the receiver and source sheets, which is actually what we derive directly from the operator

updating procedure, without estimating a background velocity model.

In case the operators have been used to estimate a near surface model, the method of Kästner and Buske (1999) is preferred, otherwise the method of Vanelle and Gajewski (2000) could be used. It should be noted though, that if no near surface model has been estimated, some assumptions should be made on the CFP-locations. The fluctuations in depth are still unknown and as demonstrated in Chapter 4, the lateral locations could have drifted during the updating procedure. If it is assumed that the reflector we are redatuming to only varies marginally in depth and drift is small, then the errors in $\mathbf{s}$ and $\mathbf{g}$ will be small.

Since the second order derivatives are related to the projected slownesses, derived through central differencing of traveltimes, the obtained slownesses are sensitive to fast fluctuations of the measured traveltimes. Since these fluctuations will in general be present within the operators due to the near surface problems, it is suggested to laterally filter the measured traveltimes, e.g. by means of a alpha-trimmed mean filter, before applying the outlined procedure.

The amplitudes obtained through subsequent application of Equation 7.3.20 and Equation 7.3.16 are referred to as the background amplitudes to be applied in the redatuming procedure.

Residual amplitudes

In the past several authors, such as Lavielle (1991), Spagnolini (1993), Berlioux and Harlan (1997), Kirchheimer and Ferber (2001), Guo and Zhou (2001), Denisov et al. (2001) and Chemingui et al. (2001), have investigated methods to decompose measured amplitudes in order to correct for near surface effects. The work conducted by the cited authors is all based on the principle papers of Cambois and Stoffa (1992) and Cambois and Stoffa (1993), in which the amplitudes are linearly decomposed by taking the logarithm of the measured amplitudes in the temporal Fourier domain. These methods are based on the surface consistency of the measured near surface amplitude anomalies and are, therefore, assumed to be applicable to the full time-length traces. Here we would like to take the methodologies one step further; instead of analyzing the full trace, the amplitude decomposition will be operated only on the event of interest, which can be described by the full lateral set of traveltime operators as derived in the previous chapters. Using Fermat modeling we can construct the datum reflection event through a proper combination of focusing operators. Through a stationary phase analysis we are able to determine the stationary location x_s along the reflector we have estimated the focusing operators for, such that we can suffice by evaluating the following expression,

$$P(x_{rcv}, x_{src}, \omega) = W(x_{rcv}, x_s) R(x_s, x_s) W(x_s, x_{src}, \omega), \tag{7.3.21}$$

under the assumption of a locally reacting medium. Following Cambois and Stoffa (1993), in the high frequency approach (stationary phase), each trace is written in the time domain as a temporal convolution of three terms,

$$p_{ij}(t) = s_i(t) * r_j(t) * \gamma_{ij}(t), \tag{7.3.22}$$

in which $s_i(t)$ and $r_j(t)$ describe surface consistent source and receiver characteristics respectively and $\gamma_{ij}(t)$ propagation characteristics through the overburden. In the temporal frequency domain Equation 6.3.22 can be written as,

$$P_{ij}(\omega) = S_i(\omega)R_j(\omega)\Gamma_{ij}(\omega). \tag{7.3.23}$$

Cambois and Stoffa (1993) assume validity of CMP-processing such that Equation 7.3.23 can be written in the common midpoint offset domain as,

$$P_{ij}(\omega) = S_i(\omega)R_j(\omega)\Gamma_{kl}(\omega), \tag{7.3.24}$$

in which k is the CMP-number, $k = i + j$, and l is the offset index, $l = i - j$. Under the assumption of valid CMP processing an operator Q is considered which transforms the CMP-responses of each offset index to a normal incidence response,

$$Q(P_{ij}(\omega)) = S_i(\omega)R_j(\omega)Y_k(\omega), \tag{7.3.25}$$

in which Y_k is the normal incidence response of CMP k. Operator Q contains processes such as field statics, normal move-out correction and correction for geometrical spreading. Furthermore, operator Q is considered to be linear in time and frequency, which is often not the case. For instance normal move-out introduces stretching and muting for the far offsets. For that purpose an additional offset term is introduced, which corrects for non-linear offset dependent artifacts.

Instead of CMP-oriented decomposition, here we consider CFP-oriented decomposition. From Chapter 4, Section 4.6.2, we derived by stationary phase analysis a CFP oriented response for each source-receiver pair,

$$P_{ij} = A_{ik}A_{kj}R_k \exp\left(-j\omega\tau_{ik}\right)\exp\left(-j\omega\tau_{kj}\right)\frac{\sqrt{2\pi}}{\sqrt{\mu}}\exp\Big[-j\frac{\pi}{4}\text{sign}\mu\Big], \tag{7.3.26}$$

$$\mu = \sqrt{\omega\big[\frac{\partial^2(\tau_{ik}+\tau_{kj})}{\partial x^2}\big]_{x=x_k}}, \tag{7.3.27}$$

in which the amplitude factors A_{ik} and A_{kj} describe the amplitude effects of propagation from respectively receiver i and source j to CFP-location k and are determined by transmission effects, T_i and T_j, and spherical divergence, $\mathcal{A}_{ik}$ and $\mathcal{A}_{kj}$ as defined by Equation 7.3.16. The factor R contains the angle dependent reflection coefficient. The stationary location at the reflector is denoted by x_k. Spherical divergence can be estimated from the derived traveltime curves as explained in the previous section

and serves as the background amplitude. Down-going transmission effects manifests itself by a surface consistent source component and up-going transmission by surface consistent receiver components. Thus the amplitude can be described by,

$$\|P_{ij}\| = T_i^- T_j^+ R_k \mathcal{A}_{ik} \mathcal{A}_{kj} Y_{i-j}, \tag{7.3.28}$$

in which Y_{i-j} is a factor which corrects for nonlinear factors such as angle-dependent reflectivity. Here we have a similar expression as derived by Cambois and Stoffa (1993). If we correct the data with the obtained two-way traveltime then the residual data is described by,

$$P_{ij} = \left[\mathcal{A}_{ik} \mathcal{A}_{kj} Y_{i-j}\right] T_i^- T_j^+ R_k \exp\left(j\omega \Delta t_{ij}\right), \tag{7.3.29}$$

in which Δt_{ij} are the residual traveltimes and the term between square brackets accounts for propagation effects approximately equal to propagation through a homogeneous near surface and the rest terms account for near surface scattering effects in terms of transmission losses. The phase part can be accounted for by means of a near surface decomposition as well, equivalent to Taner et al. (1974),

$$\Delta t_{ij} = \Delta t_i + \Delta t_j + \Delta t_k + \Delta t_{i-j}. \tag{7.3.30}$$

Taking the logarithm of Equation 7.3.29 results in,

$$\begin{aligned} \log P_{ij} = & \log T_i^+ + \log T_j^- + \log R_k + \log Y_{i-j} \\ & + \log \mathcal{A}_{ik} + \log \mathcal{A}_{kj} \\ & - j\left[\omega \Delta t_i + \omega \Delta t_j + \omega \Delta t_k + \omega \Delta t_{i-j}\right]. \end{aligned} \tag{7.3.31}$$

Thus, we can solve the set of equations for both the real part and the imaginary part, from which the real part describes the amplitude decomposition and the imaginary part the phase decomposition. To solve the amplitude decomposition we have to solve the system of equations described by,

$$\log \|P_{ij}\| - \log \mathcal{A}_{ik} - \log \mathcal{A}_{kj} = \log T_i^+ + \log T_j^- + \log R_k + \log Y_{i-j}, \tag{7.3.32}$$

or in matrix notation

$$\mathbf{p} = \mathbf{G}\mathbf{m} \tag{7.3.33}$$

For which, according to Appendix B, a least-squares solution reads,

$$\mathbf{m} = \left[\mathbf{G}^T \mathbf{W} \mathbf{G} + \mathbf{F}^T \mathbf{F}\right]^{-1} \mathbf{G}^T \mathbf{W} \mathbf{p}, \tag{7.3.34}$$

in which the weighting matrix depends on the properties of the signal and $\mathbf{F}$ is determined by a priori information on the parameters and noise. As pointed out by Cambois and Stoffa (1993), although the system seems overdetermined, it still

contains a null space. Thus Cambois and Stoffa (1993) suggest to equip the system with an additional set of a priori equations,

$$\sum \log T_i - \frac{\sum \log T_i}{N} = 0, \qquad \sum \cos(2\pi \frac{i-1}{N}) \left(\log T_i - \frac{\sum \log T_i}{N} \right) = 0, \tag{7.3.35}$$

$$\sum \log T_j - \frac{\sum \log T_j}{N} = 0, \qquad \sum \cos(2\pi \frac{j-1}{N}) \left(\log T_j - \frac{\sum \log T_j}{N} \right) = 0, \tag{7.3.36}$$

$$\sum \log R_k - \frac{\sum \log R_k}{N} = 0, \qquad \sum \sin(2\pi \frac{i-1}{N}) \left(\log T_i - \frac{\sum \log T_i}{N} \right) = 0, \tag{7.3.37}$$

$$\sum \log Y_{i-j} - \frac{\sum \log Y_{i-j}}{2N} = 0, \qquad \sum \sin(2\pi \frac{j-1}{N}) \left(\log T_j - \frac{\sum \log T_j}{N} \right) = 0, \tag{7.3.38}$$

$$\sum (-1)^k \left(\log R_k - \frac{\sum \log R_k}{N} \right) = 0, \tag{7.3.39}$$

in which N equals to the number of surface locations. The a priori equations assure that the average residual corrections should equal zero to remove biases and can be described as,

$$\mathbf{Fm} = 0, \tag{7.3.40}$$

which provides the a priori part of Equation 7.3.34. What we are left with is used to compensate the data for. Thus the surface consistent parts are subtracted from the data, while the spherical divergence and offset dependent parts are used to construct the amplitudes of the focusing operators with. Note that we have to deconvolve for the whole data set and not only the reflector we have selected.

Estimation of offset dependent transmission losses

Intuitively, we would like to introduce an offset dependent transmission term for both source and receiver. This term should correct for angle dependent transmission. Naturally we would like to estimate the full transmission function, but then the number of parameters would become too large and the data is no longer redundant, to obtain an overdetermined system. Therefore a three-parameter model of the offset dependent transmission effect is proposed, describing a translated Gaussian shape,

$$\mathcal{T}(x_{src}, x_{cfp}) = T(x_{src}) \exp\left(-\frac{((x_{src} - x_{cfp}) - x_{0,src})^2}{\sigma_{src}^2}\right), \tag{7.3.41}$$

in which $T(x_src)$, $x_{0,src}$ and σ_{src}^2 are the three parameters (see Figure 6.9), which vary as function of x_{src}. When the data is transformed to the log/Fourier domain, the parameterized transmission function of Equation 7.3.41 would read in the

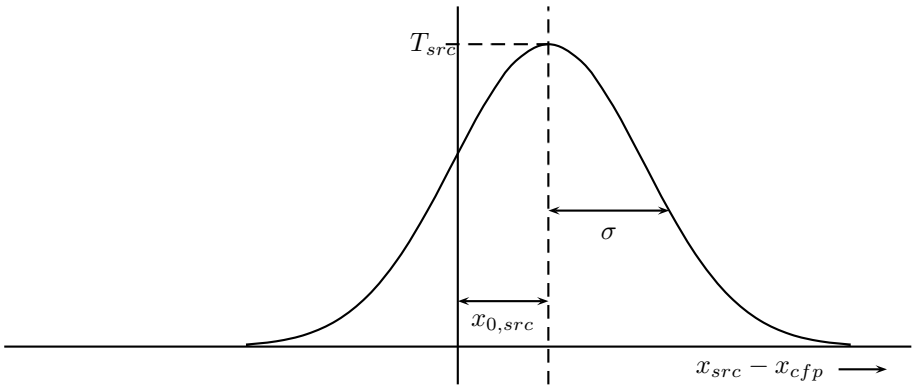

Figure 7.9: *Schematic representation of the parameterized transmission function.*

log/Fourier domain,

$$\log \mathcal{T}(x_{src}, x_{cfp}) = \log T(x_{src}) - \frac{((x_{src} - x_{cfp}) - x_{0,src})^2}{\sigma_{src}^2}, \tag{7.3.42}$$

$$\begin{aligned} =& \frac{1}{\sigma_{src}^2}(x_{src} - x_{cfp})^2 - 2\frac{x_{0,src}}{\sigma_{src}^2}(x_{src} - x_{cfp}) \\ &+ \left[\frac{x_{0,src}^2}{\sigma_{src}^2} + \log T(x_{src})\right]. \end{aligned} \tag{7.3.43}$$

As such we can extend the parameterization of the amplitudes as described by Equation 6.3.32 through,

$$\begin{aligned} \log \|P_{ij}\| - \log \mathcal{A}_{ik} - \log \mathcal{A}_{kj} =& \log R_k + \log Y_{i-j} + T_{0;i} + T_{0;j} \\ &+ T_{1;i}(x_i - x_k) + T_{1;j}(x_j - x_k) \\ &+ T_{2;i}(x_i - x_k)^2 + T_{2;j}(x_j - x_k)^2, \end{aligned} \tag{7.3.44}$$

with,

$$\sigma_i^2 = \frac{1}{T_{2;i}}, \tag{7.3.45}$$

$$x_{0,i} = -\frac{1}{2}\frac{T_{1;i}}{T_{2;i}}, \tag{7.3.46}$$

$$\log T_i = \frac{1}{4}\frac{T_{1;i}^2}{T_{2;i}} - T_{0;i}. \tag{7.3.47}$$

Thus, by introducing higher order surface consistent terms in the log/Fourier domain, we can account for directional dependent transmission effects.

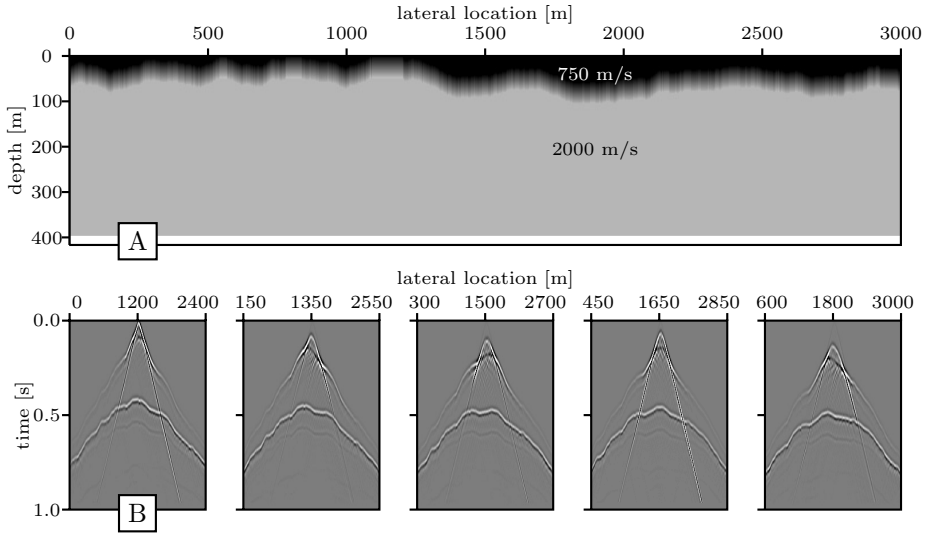

Figure 7.10: *A) The synthetic model used to demonstrate true amplitude redatuming. At 400 m. depth a reflecting event is placed to which the data will be redatumed. B) Five of the shot gathers, modeled over the model in A), using an acoustic finite difference algorithm.*

7.4 Example

The following example illustrates the results of amplitude estimation using surface consistent decomposition and using the adjusted decomposition. Once the amplitudes have been estimated using the adjusted decomposition, the least-squares redatuming results are analyzed in comparison to redatuming through matched-filter extrapolation.

Figure 7.10A shows the used model and Figure 7.10B shows five shot-gathers, modeled with the aid of an acoustic finite differencing algorithm. The model contains only one reflecting event, located at 400m depth. The reflections of this event are caused by a contrast difference only, therefore the recorded reflections contain no dependency on illumination angle. The recorded data will be redatumed to this particular event. It is assumed that the operator times, to be used for redatuming, are perfectly estimated. After operator updating we are able to retrieve the amplitudes of the focused event. Figure 7.11A shows the amplitude map (average over all frequencies) for all sources and receivers. From the operator times, the amplitudes related to spherical divergence can be estimated. Figure 7.11C shows the recorded amplitude map in the log/Fourier domain, corrected for spherical divergence. The correct amplitudes of the redatuming operators were computed by finite differencing within the same model, Figure 7.10A, by placing sources at the reflector and recording their energy at the surface. The operator amplitudes are displayed in Figure

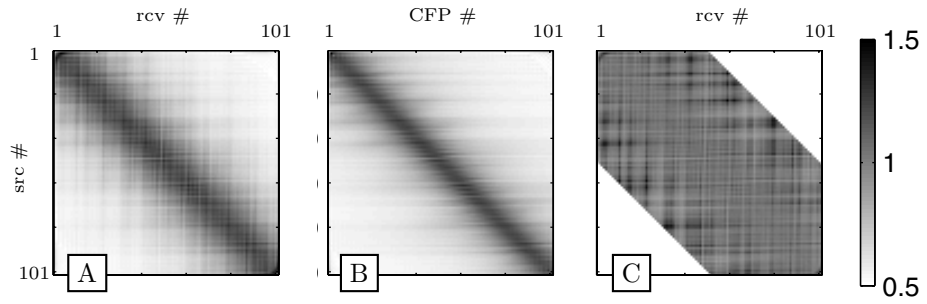

Figure 7.11: *A) Amplitudes extracted from the measured data, along the two-way traveltimes. B) Amplitudes of the forward modeled CFP operators. C) The logarithm of the measured amplitudes, corrected for geometrical spreading.*

7.11B.

In Figure 7.12 the results are displayed of the inversion process to obtain the near surface amplitude effects. Results of both the surface consistent decomposition as well as the adjusted decomposition are displayed. As the data could be distorted by the direct arrival, remnants of near surface reflection and other correlated noise events, the amplitudes were estimated by using a hybrid norm solution within the estimator (see Appendix B). This makes the inversion scheme non-linear and, therefore, the solutions were found iteratively. It should be noted that the first iteration within the inversion scheme resembles a Gaussian norm solution.

Through surface consistent deconvolution, surface consistent source and receiver amplitude terms are found. These were used within the forward model to forward compute the spherical divergence corrected amplitudes of the recordings. Figure 7.12A displays these forward modeled amplitudes after one iteration of the hybrid norm solution and Figure 7.12B displays the forward modeled amplitudes after 9 iterations, which shows some improvement. The assumption of surface consistency nicely shows up as vertical (source term) and horizontal (receiver term) striping within the forward modeled data. However, the residuals, computed by subtracting the measured amplitudes of Figure 7.11C from the forward modeled amplitudes of Figure 7.12B, shown in Figure 7.12C, are quite large.

In Figure 7.12D the forward modeled amplitudes resulting from the Gaussian norm solution of the adjusted amplitude decomposition are displayed and in Figure 7.12E the forward modeled amplitudes resulting from the hybrid norm solution are displayed. Again some improvements are observed between the Gaussian norm and the hybrid norm. More remarkable is the observation of skewed diagonal striping within the forward modeled amplitudes. Apparently the adjusted decomposition was capable of determining the near surface amplitude effects better. The good performance

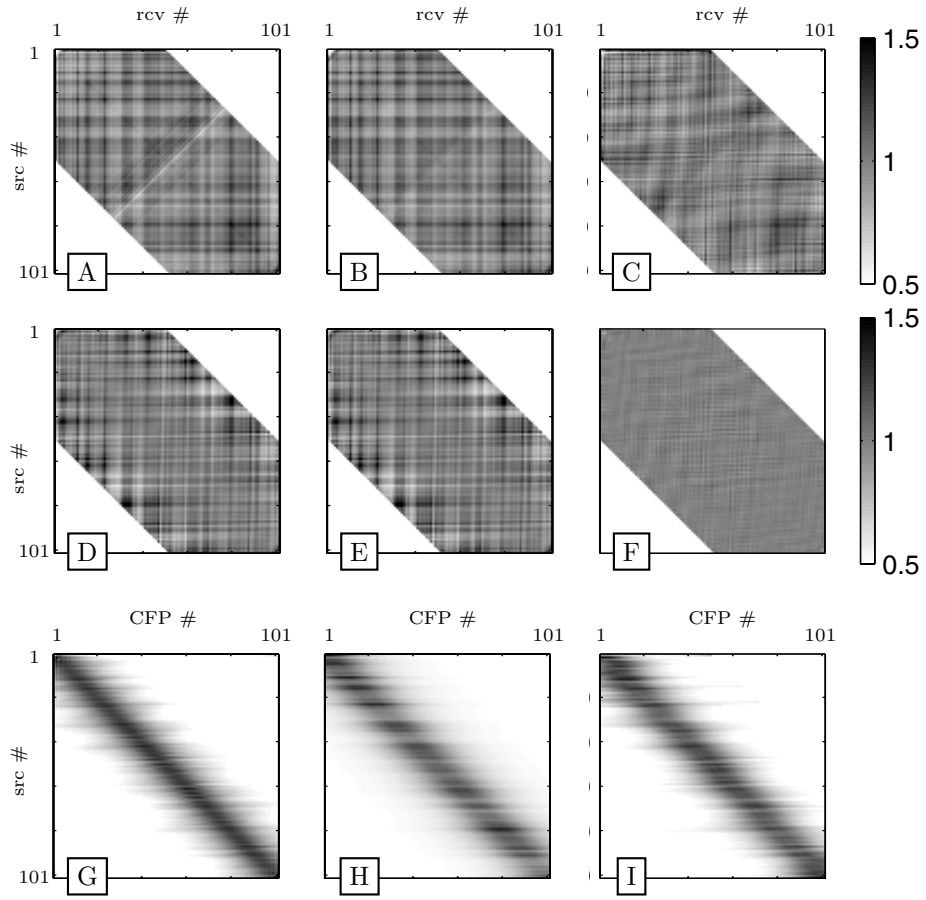

Figure 7.12: *Logarithmic results of the amplitude estimation procedure. A) displays the first iteration of the hybrid-norm estimation based on pure CMP processing. B) shows the same results after reaching the cut-off error norm. C) displays the difference with the measured amplitudes. D) to F) show the same results as A) to B), but here an efficient dependent factor was included for CFP processing. G) shows the forward modeled operator amplitudes, H) shows the estimated operator amplitudes after CMP processing, I) shows the operator amplitudes after CFP processing.*

is confirmed by Figure 7.12F, displaying the residual amplitudes obtained through subtracting the measured amplitudes from the forward modeled amplitudes. Clearly, by introducing an offset dependent transmission effect, the inversion procedure to retrieve the operator amplitudes is able to recover the near surface amplitude effects quit well.

In Figure 7.12G the correct, forward modeled amplitudes of the focusing operators are displayed. Figure 7.12H displays the amplitudes retrieved from the surface consistent amplitude decomposition whereas Figure 7.12I displays the result of the adjusted amplitude decomposition. The amplitudes estimated by including an offset dependent transmission show the same features as the original amplitudes in a filtered sense. The surface decomposition, however, renders erroneous results, especially for the CFP numbers between 30 and 70.

As by now both the amplitudes as well as the times have been estimated for the redatuming operators, a redatuming can be applied. It is preferred to correct the data for all amplitude components which are surface consistent (independent of offset) to improve stability of the least-squares redatuming. The data within the example contains only reflection event, the event we are redatuming the surface recorded data to. As such the result after redatuming contains the imaged energy of this particular event. The reflection energy arose from a density contrast only, therefore the imaged energy should not vary as function of incident angle. Thus, if the redatuming result is analyzed in the time and spatially transformed domain,

$$\hat{\mathbf{P}}(z_d, z_d) = \mathcal{F}_x^+\{\mathbf{P}(z_d, z_d)\}, \tag{7.4.48}$$

then the resulting amplitude spectrum, $\|\hat{\mathbf{P}}(z_d, z_d)\|$, should be uniform in value within the bandwidth determined by the maximum frequency and the minimum apparent velocity.

Figure 7.13A displays the spectrum of a redatumed shot-gather, after least-squares redatuming without using amplitudes. The operators have amplitudes equal to one, regardless the offset and traveltime. As such operators are completely inconsistent with the forward model, the spectrum of the resulting redatumed data is far off from the result we would have expected. Figure 7.13B displays the spectrum of a redatumed shot-gather, after redatuming through matched filter extrapolation, using amplitudes based on spherical divergence only. As was stated in Chapter 4, Section 4.5.1, indeed the matched filter approach performs reasonably well although effects arising from the near surface can clearly be observed.

Figure 7.13C displays the spectrum of a redatumed shot-gather, after redatuming through the matched filter approach, using the estimated amplitudes. Also stated in Chapter 4, Section 4.5.2, when the focusing operators are equipped with the estimated amplitudes attributed to the near surface, the matched filter redatuming will

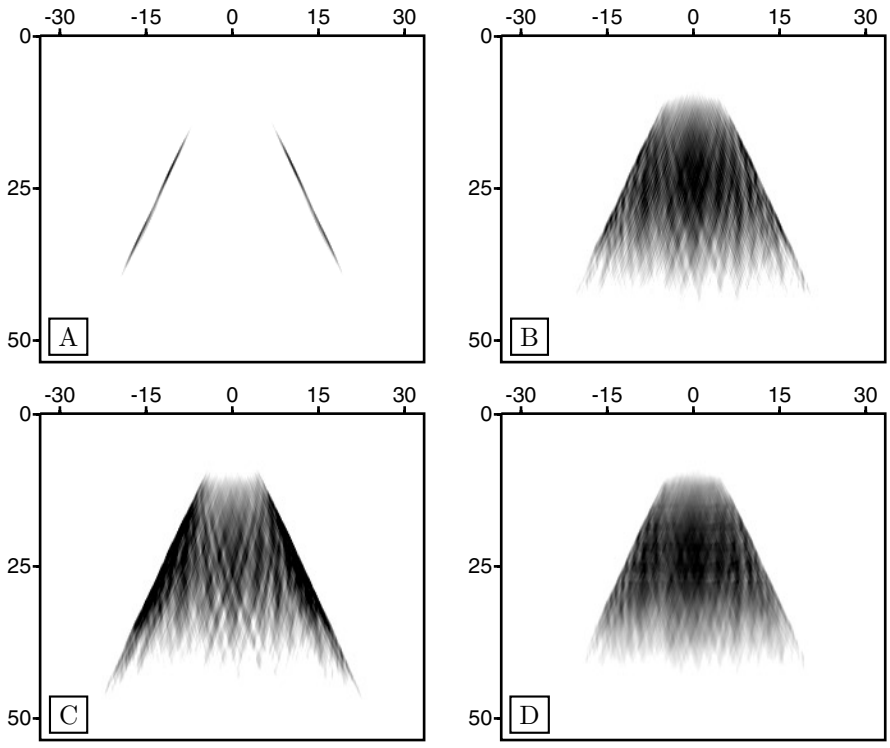

Figure 7.13: *Spectra after redatuming. A) shows the result after least-squares redatuming without any attempt of including amplitudes. B) shows the result after matched-filter redatuming, using amplitudes based on spherical divergence only. C) Shows the results after a matched filter redatuming, using the estimated amplitudes. D) Shows the results after least-squares redatuming using the estimated amplitudes.*

amplify these effects. Comparing Figure 7.13C to Figure 7.13B clearly shows that the near surface effects have a stronger imprint within the spectrum. Figure 7.13D displays the spectrum of a redatumed shot-gather, after least-squares redatuming using the estimated amplitudes. Clearly the least-squares redatuming renders the cleanest spectrum.

An alternative domain to analyze the data in, is the Radon-domain, obtained through,

$$\tilde{\mathbf{P}}(z_d, z_d) = \mathcal{R}\{\hat{\mathbf{P}}(z_d, z_d)\}. \tag{7.4.49}$$

The resulting gathers after redatuming only contain the imaged energy of the recorded reflector. By transforming the redatumed data to the Radon domain, along $t = 0s$ the angle-dependent reflectivity can be found as function of ray-parameter for all CFP locations along the reflector. Again, if the data would have been perfectly redatumed, these so-called AVP-functions would render unitary amplitudes everywhere.

Figure 7.14A displays the AVP function of a redatumed shot-gather, after least-squares redatuming without using amplitudes. Figure 7.14B displays the AVP function of a redatumed shot-gather, after redatuming through matched filter extrapolation, using amplitudes based on spherical divergence only. Figure 7.14C displays the AVP function of a redatumed shot-gather, after redatuming through the matched filter approach, using the estimated amplitudes. Figure 7.14D displays the AVP function of a redatumed shot-gather, after least-squares redatuming using the estimated amplitudes. Similar effects can be observed as in Figure 7.13. The AVP curves are printed over their values in gray scale. Clearly the least-squares redatuming renders an almost perfect unitary response. Note that the grayscale values of Figure 7.14, the result retrieved after redatuming through matched-filter extrapolation with near surface amplitude effects included, show peaks and troughs (light vs. dark) which have a similar pattern as the interface between the near surface and the deeper subsurface present in the model of Figure 7.10A. This confirms again the statement made in Chapter 4, Section 4.5.2, that the matched-filter extrapolation with the correct operator would amplify near surface transmission effects.

7.5 Conclusions

Especially when we are dealing with a complex near surface problem, amplitude distortions can become quite severe. Synthetic examples show that correct handling of amplitudes can improve the redatuming results.

In this chapter we have outlined a sequence of steps to determine the amplitudes of the focusing operators, such that an optimal least-squares redatuming result is obtained.

First, we correct the data with the two-way traveltimes, found by using Fermat

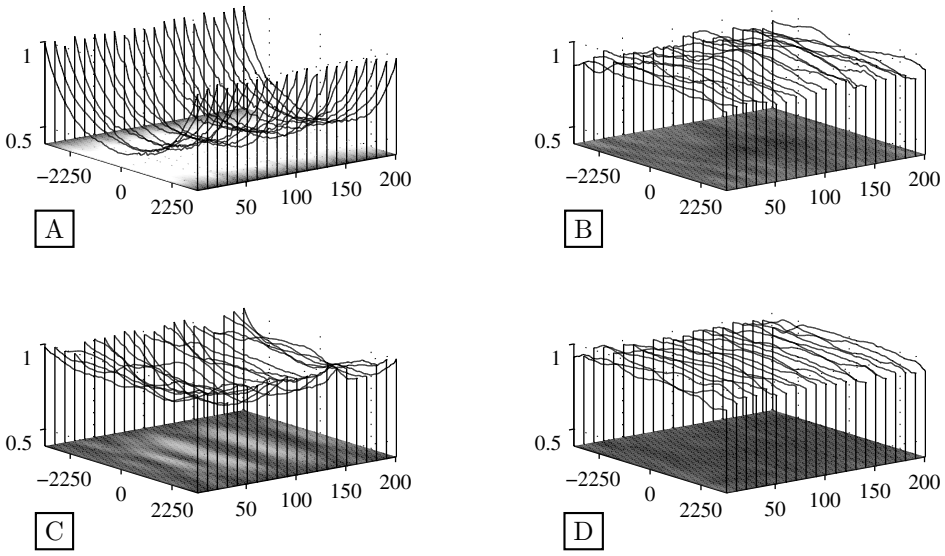

Figure 7.14: *Results of redatuming for data with one reflector of unit reflection amplitude, displayed as amplitudes in the ray-parameter domain. A) Shows the result after least-squares redatuming without any attempt of including amplitudes. B) Shows the result after matched filter redatuming, using amplitudes based on spherical divergence only. C) Shows the results after a matched filter redatuming, using the estimated amplitudes. D) Shows the results after least-squares redatuming using the estimated amplitudes.*

modeling. From the corrected data we can easily retrieve the measured amplitudes.

In a pre-processing phase we are able to remove scattering noise from the data. By using Fermat modeling we can extend the method developed by Campman (2005) to all reflections that are present in field data, without the restriction of local plane wave events.

Determining the operator amplitudes through modeling has no future. In this thesis this problem is solved by proposing a constraint data driven method.

A synthetic example shows that the proposed sequence of procedures is very well capable of rendering an improved least-squares redatuming result, using the focusing operators, obtained through the data-driven analysis as explained in the previous chapters.

8

Examples

8.1 Introduction

In this chapter the methodologies as outlined within this thesis are demonstrated on two realistic synthetic models. As the performance of near surface technology is most probably extremely sensitive to 3D effects, we refrained from applying the methodologies on real data. For that matter further research to the effects of and applicability to 3D near surface effects need to be done.

The first example is made using a synthetic model and data generated by Saudi-Aramco. It resembles a typical desert situation, suffering from intermediate wavelength statics.

The second example uses synthetic data that has been modeled by BP. The considered model is a subset of a larger model, containing all types of near surface effects that could occur in a real world. The data was generated in order to test all available static methodologies and algorithms that are currently available in the market. The subset we have chosen to analyze resembles a shorter wavelength static problem and is regarded to be one the most difficult parts of the model, according to BP.

8.2 The Middle East model

Figure 8.1A displays a typical model of a Middle East desert situation (courtesy Saudi-Aramco). Common near-surface features and anomalies encountered in desert situations are low velocity sand dunes, wadi's with and without infill material and cliffs at the edges of the wadi's, which are all present within this model. The Rus-formation, as depicted in the model, is a common strong high velocity reference marker which is present almost throughout the complete Arabian peninsula. The

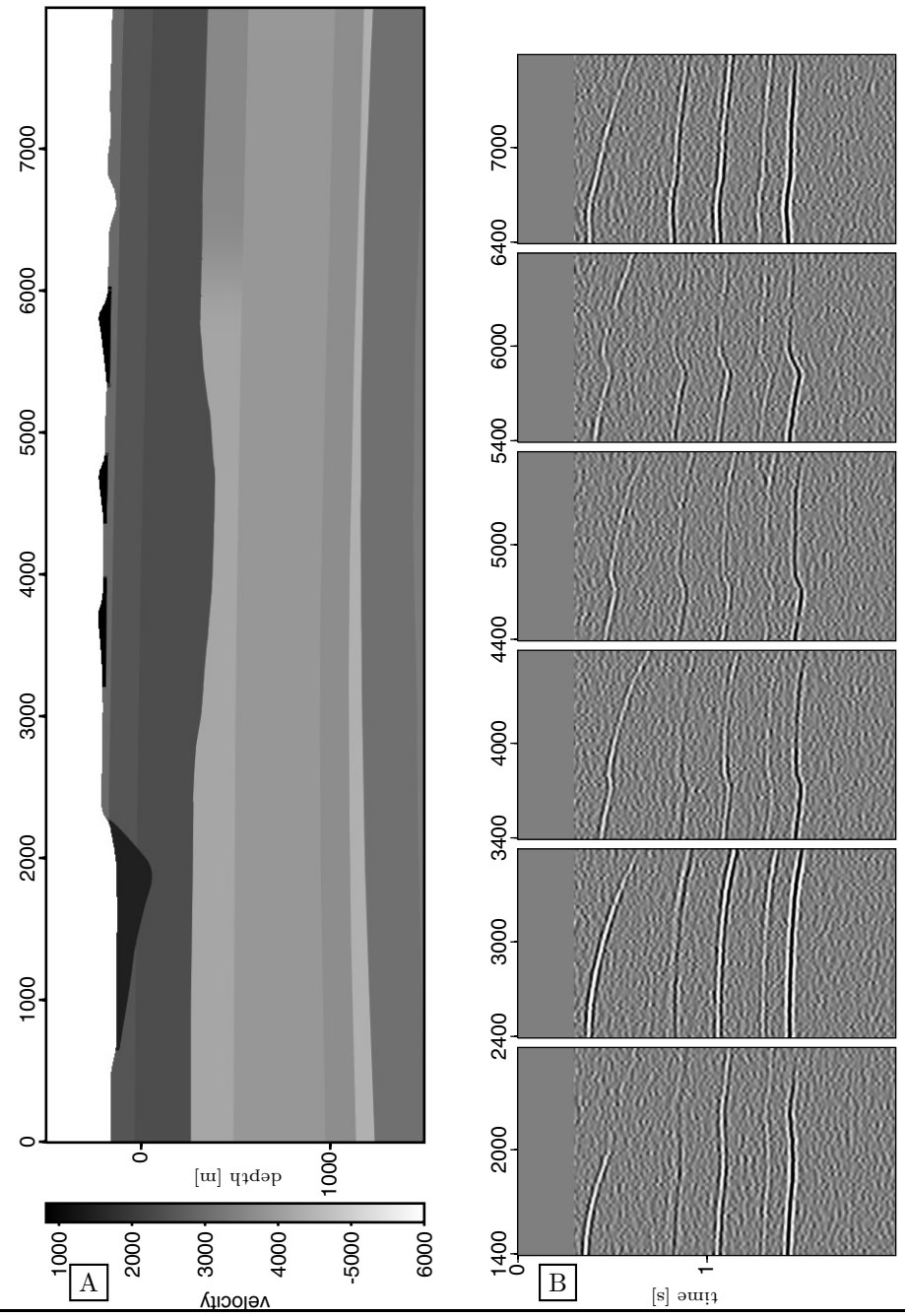

Figure 8.1: *(A) The synthetic model of a desert environment (courtesy Saudi-Aramco). (B) Five selected of the forward modeled shots, using paraxial raytracing .*

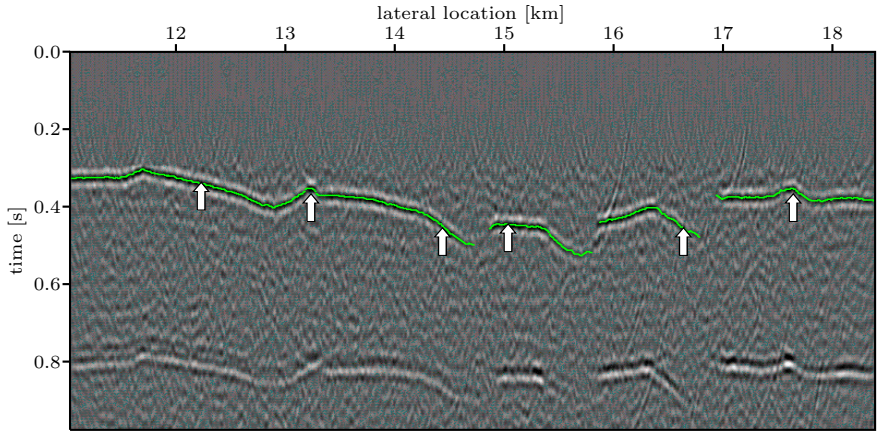

Figure 8.2: *Operating window of the Delphi CFP-workbench. In the window the time-migrated zero offset section with the tracked datum event is visible. The arrows indicate at the CFP locations for which the updating will be illustrated.*

depicted model was especially designed to test the performance of near surface algorithms over longer wavelength anomalies. Data was generated over the model by using paraxial raytracing for reflection energy only. Paraxial raytracing was used to define reasonable amplitudes as well. Although not typical for land-surveys, the data was generated as end-on data with a spread-length of 1100 m (end-on refers to a survey with positive offsets only, typical for a seismic vessel towing a source just behind the boat followed by receiver streamers). Figure 8.1B shows a set of modeled shot-records. The number of modeled shots is 360, having 60 recording groups each. The shot spacing was taken to be equal to the group-spacing of 20m.

■ 8.2.1 CFP processing

The first step in defining focus point operators is determining the initial operators from a near offset combined with a first pass estimate of the stacking velocities. Figure 8.2 displays the near offset stack on which the automatically tracked zero-offset times of the reflector of interest are indicated. Also note the dim spots due to the presence of the sand dunes and the structure caused by the wadi. In these spots we were not able to track coherent energy and the zero-offset times will be interpolated. Figure 8.3 displays the final set of zero-offset times.

Two sets of initial operators were constructed. The first set describes the travel-

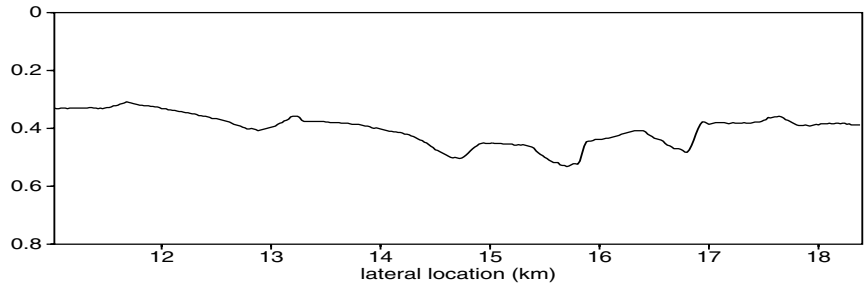

Figure 8.3: *Interpolated times obtained from tracking the two-way traveltimes in the time section of Figure 8.2.*

times from each focus point toward all surface locations by means of a zero-offset time and a single move-out velocity, as depicted in Figure 8.4A. The second set describes the traveltimes from each surface location to all anticipated CFP-locations by a zero-offset time and a single move-out velocity. By reordering the times of the second-set we obtain the initial operators as depicted in Figure 8.4B. Note that the second set of operators already contains a first estimate of the near surface imprint, which was also discussed in Chapter 5.

Both sets were subsequently used to construct CFP gathers and update the operators iteratively. The updating was done by using an automatic tracker on the full volume of constructed DTS-gathers in each update iteration.

Figure 8.5 displays six of the initial DTS-gathers constructed with the smooth initial operators of the first set. Note the near surface effects existing within the initial DTS gathers. Going through the update sequence, we arrive after 3 iterations at the DTS-gathers as depicted in Figure 8.6. Note that some gathers are nicely flattened, however, others still show some near-surface imprint, dim spots and energy missing for a range of offsets. Clearly not all near surface effects were detected well during the updating sequence. This is confirmed by the tracked surfaces in Figure 8.7A and Figure 8.7C. These figures display the tracked differential times for the initial DTS-gathers and the final DTS-gathers respectively. Where the figures did not render any data, it means that the autotracker was unable to track coherent energy. This is especially the case in the areas where the dunes were encountered and refer to a non-optimal fresnel zone-construction as explained in Chapter 5. Although the final operators of Figure 8.4C have changed with respect to the initial operators, they are still behaving relatively smooth which indicates that the near surface effects were hardly detected. The final operators were also used to redatum the data through a matched filter extrapolation of the sources and the receivers. Figure 8.8 displays 10 shot-gathers after redatuming. The shot-gathers, especially gather 210 and 270,

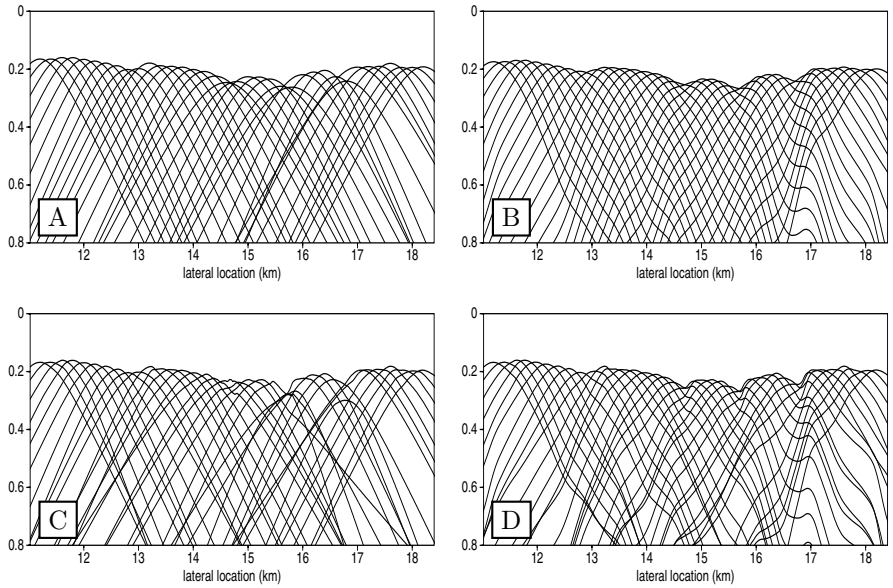

Figure 8.4: *Result of CFP operator updating. A) Displays initial operators with a hyperbolic move-out. B) displays the initial operators transposed, thus including a first estimate of the near surface effects. C) Displays the updating results after three updates of the hyperbolic operators. D) Displays the updating results after three updates of the transposed operators.*

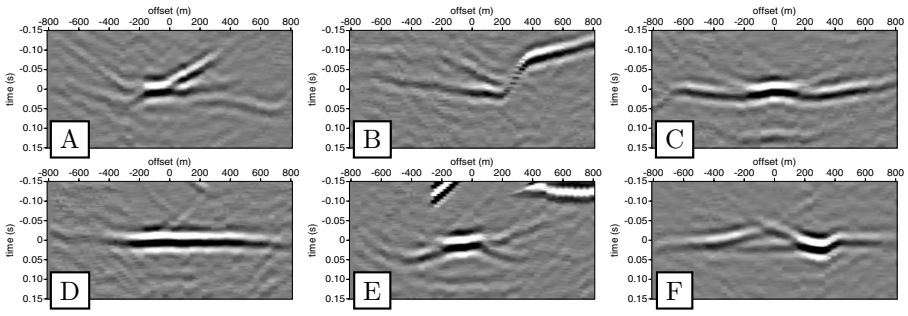

Figure 8.5: *DTS panels with the original initial hyperbolic operators (Figure 8.4A) at shot locations A) 60, B) 110, C) 170, D) 200, E) 280 and F) 330.*

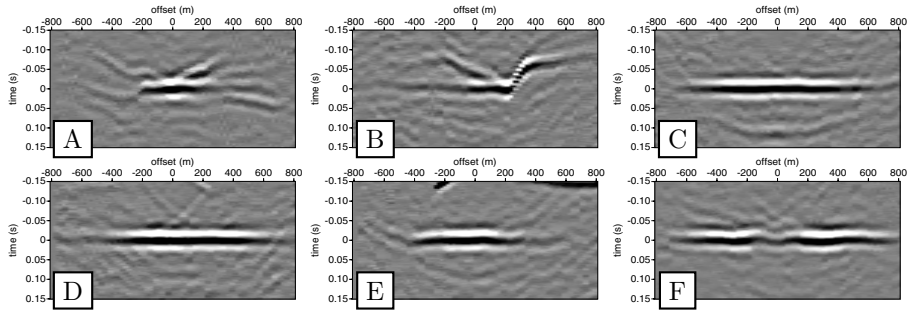

Figure 8.6: *DTS panels after three updates, starting the updating from the initial hyperbolic operators (Figure 8.4C) at shot locations A) 60, B) 110, C) 170, D) 200, E) 280 and F) 330*

clearly show features which can by no means be attributed to the deeper subsurface. The final operators are not sufficiently well defined to remove the near surface effects within the redatumed shot-gathers.

Using the second set of operators, which contains a first estimate on the near surface effects as displayed in Figure 8.4B, results in the initial DTS-gathers as displayed in Figure 8.9. Here we notice, that the initial DTS-gathers behave by far smoother and render more energy for the larger offsets. After three iterations we arrive at the final DTS-gathers of Figure 8.10. Clearly we have been able to flatten the DTS-gathers for all locations. The differential time surfaces as tracked by the autotracker confirm the better performance. As depicted by Figure 8.7 the autotracker was able to track coherent data within DTS-volumes for both the initial gathers of Figure 8.7B as well as the final gathers of Figure 8.7D, even at the locations in the neighborhood of the sanddunes. Figure 8.4D shows the final CFP-operators.

In a third exercise, the initial operators as displayed in Figure 8.4B were subject to the automatic updating algorithm based on the non-linear genetic evolution algorithm as discussed in Chapter 5. This algorithm tends to flatten the event in the shot-gathers, based on the principle of data-construction. Figure 8.12 shows the flattened shot-gathers after applying the genetic algorithm. Note that after the final step, the event is fairly well flattened. This result does not yet assure flattening of the DTS-gathers according to the principle of equal traveltimes. The algorithm performed reasonably well in terms of defining operators that assure the principle of equal traveltimes. In a last pass, the operators could be further updated by picking differential times within the DTS-gathers. If we compare the final operators from the automatic updating algorithm, 8.11C, they are almost equal to the operators

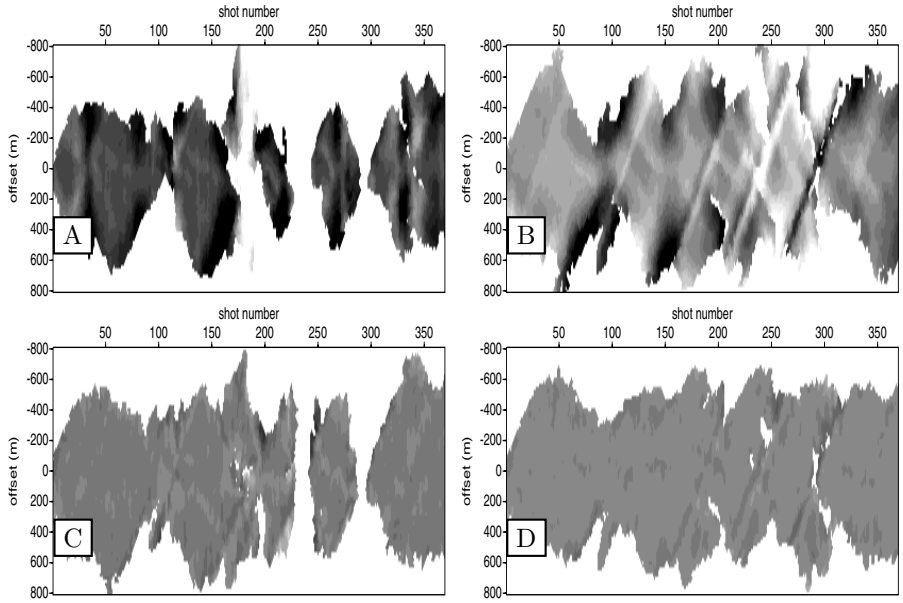

Figure 8.7: *A) Traveltime error surface obtained by automatic tracking of the datum reflection response in the initial DTS volume constructed when hyperbolic CFP operators are used. The gray-value indicates the differential traveltime value, with gray being zero. B) The tracked traveltime error surface within the initial DTS volume constructed by using transposed operators. C) The final tracked traveltime error after three updates using initially hyperbolic operators. D) The final tracked traveltime error after three updates using initially transposed operators. The white areas indicate that the tracker could not find a correlated event there. Note that around shot 240 and 290 no updating could be achieved when initally hyperbolic operators are used; including some a priori knowledge on the near surface is crucial.*

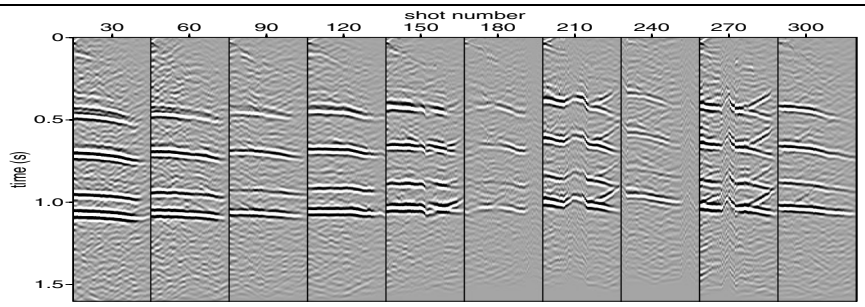

Figure 8.8: *A selection of shot records after redatuming using the estimated CFP operators of Figure 7.4B. For the CFP updating the original hyperbolic operators were used.*

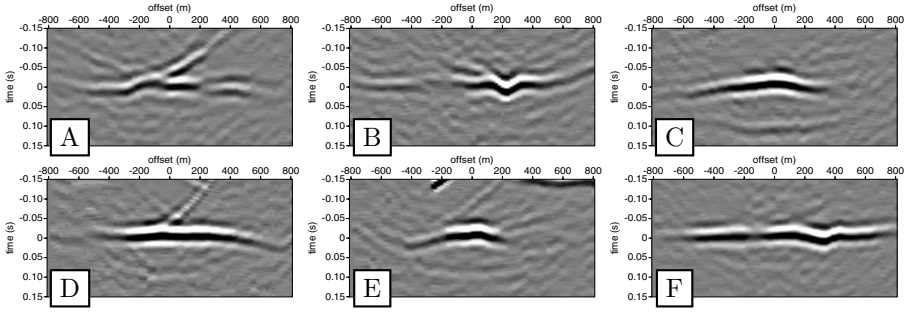

Figure 8.9: *DTS panels for the transposed initial operators (Figure 8.4B) at shot locations A) 60, B) 110, C) 170, D) 200, E) 280 and F) 330.*

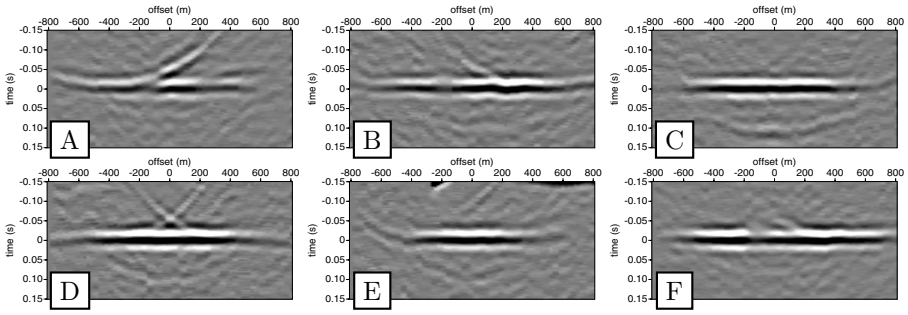

Figure 8.10: *DTS panels after three updates of the transposed initial operators (Figure 8.4D) at shot locations A) 60, B) 110, C) 170, D) 200, E) 280 and F) 330.*

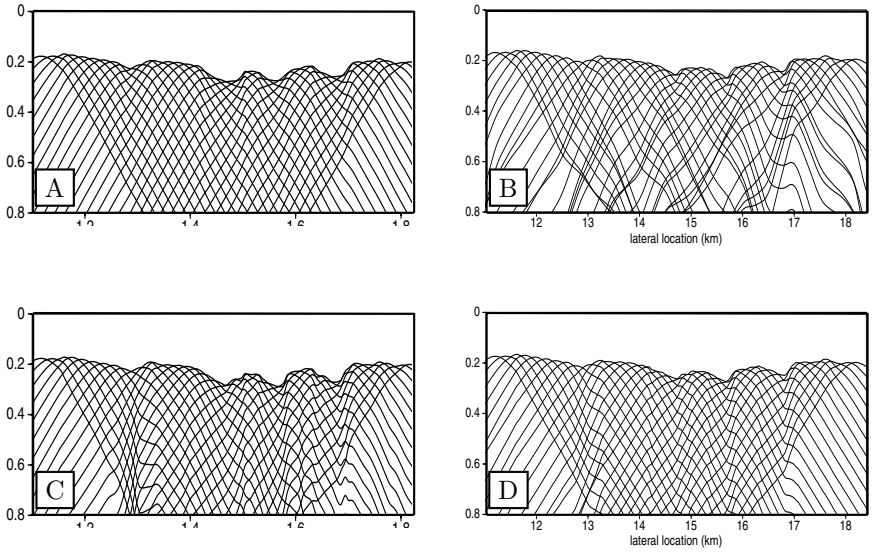

Figure 8.11: *Operators automatically determined through the principle of data reconstruction by using a genetic algorithm. A) displays the initial operators. C) Displays the final operators after sufficient convergence. For comparison, B) displays operators determined by manual updating after 3 iterations, D) displays operators modeled within the true model by using an Eikonal solver.*

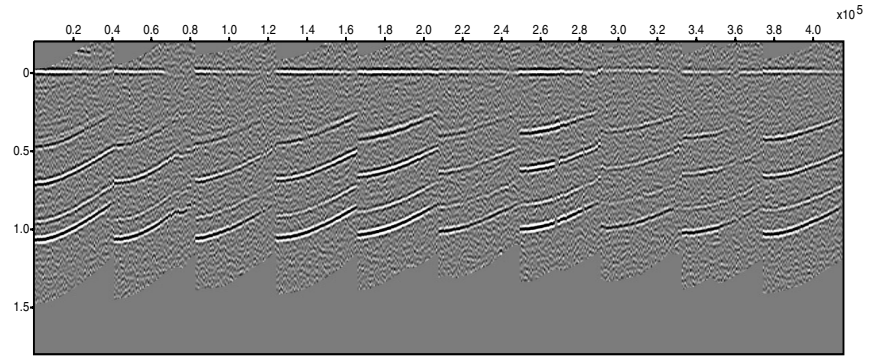

Figure 8.12: *Shot records at the surface corrected for the estimated two-way traveltimes obtained with an automatic updating procedure.*

computed by using an Eikonal solver in the original model, 8.11D, resembling the true operators. It also seems that the automatic updating algorithm outperforms manual updating for the far offsets, which can be concluded by comparing the automatically updated operators of Figure 8.11C with the manual updated operators of Figure 8.11B. This puts confidence in the deployed automatic updating procedure.

■ 8.2.2 CFP redatuming

The resulting operators, from the updating procedure are used for redatuming. The redatuming consists of two steps. First, all shots and receivers are redatumed to the Rus, the reflector used to focus the operators on. Within the next the data is forward extrapolated to a flat reference datum at zero meters depth. For this step we need both a substitute homogeneous velocity for the near surface as well as the correct depth and lateral locations of the Common Focus Points. The substitute velocity can have an arbitrary value, or a value can be used which was estimated as the refractor velocity during a previously conducted refraction exercise. The actual depth of the reflector to which we would like to redatum our data, using the CFP operators needs to be estimated through a tomographic inversion of the focusing operators.

In Figure 8.13 the initial model is displayed together with 3 iterations, the first, the third and the fifth and final iteratio, of the tomographic inversion. The known surface topology is quite well recognized. The model was initialized using two layers. The layer boundary is described by a spline with 30 vertices. The velocity of the top layer is allowed to vary laterally. The variety of the velocity is constrained by a spline with 10 vertices.

The tracked traveltime surface of Figure 8.7D is used as a weight on all the individual traveltime measurements of the focusing operators. These weights indicate which traveltime picks are reliable or not for use within the inversion scheme. The final inversion result is displayed in Figure 8.13D. Some of the original structure of the near surface is resolved, such as the wadi fill in and the reliefs below the sanddunes, although not perfectly. Due to the velocity contrast and the truncated apertures of the focusing operators, according to the applied weights, velocity and depth of the toplayer can interchange during the inversion iteration. However, the structure of the Rus formation is well resolved, the relief is clearly recognized.

Figure 8.14 displays a set of redatumed shot-gathers. Some remnants of near surface effects are still visible but in general the redatuming clearly removed the near surface effects in a kinematic sense. The redatumed shot gathers were put subject to a stacking velocity analysis, followed by a NMO correction and subsequently stacked after sorting to the CMP-Offset domain. The remaining stack is displayed in Figure

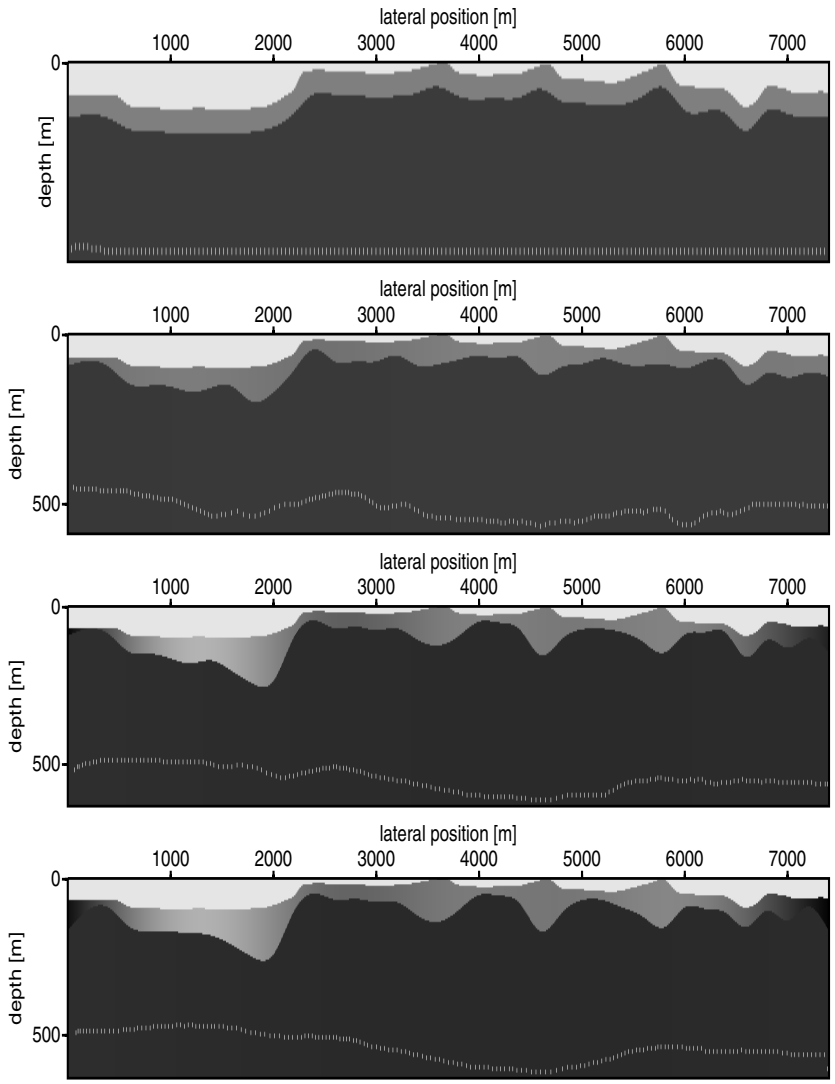

Figure 8.13: *Inversion result for inverting CFP operators. The top Figure shows the initial model, the second Figure displays the model after 1 iteration, the third Figure displays the model after 3 iterations and the bottom Figure displays the final model after 5 iterations. The CFP locations are denoted by white dots. The velocities do not fully resemble the true model, however the CFP locations are estimated quite well.*

8.15. The stack looks in general clear and smooth. The observed anticlinal structure

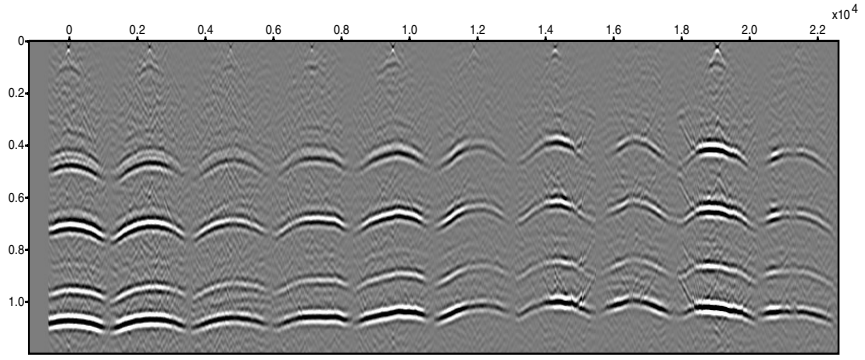

Figure 8.14: *Shot records after redatuming to the RUS formation. The reflection energy of the RUS formation has collapsed to a direct arrival.*

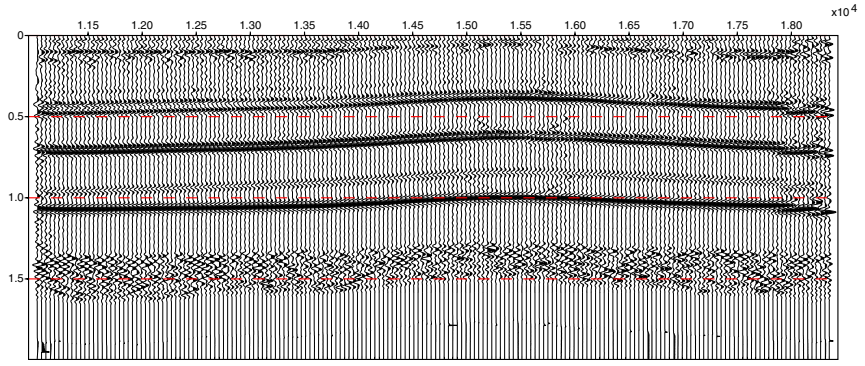

Figure 8.15: *Brute stack at RUS, after applying CFP redatuming with verified focusing operators.*

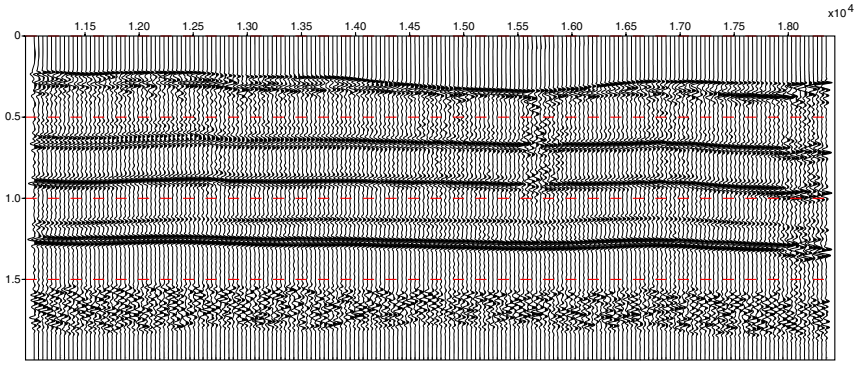

Figure 8.16: *Brute stack at floating datum, after forward extrapolation of data from RUS reflector.*

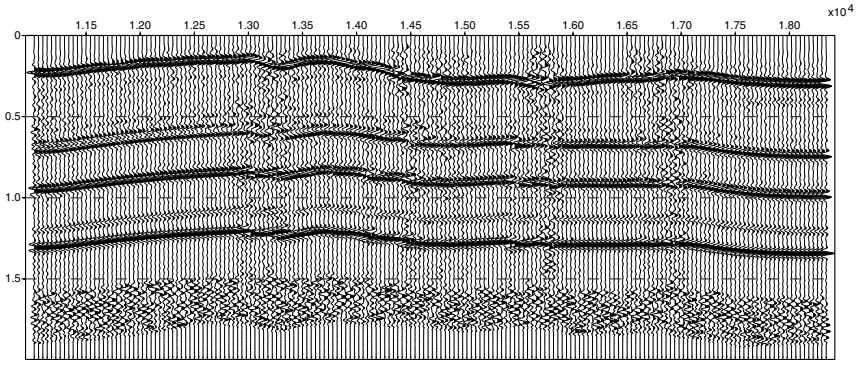

Figure 8.17: *Brute stack at floating datum after application of refraction and datum statics.*

is caused by the relief of the Rus and the fact that all sources and receivers have been redatumed to the Rus. In fact this is a false structure and can be interpreted as a pseudo topology imprint.

Using the estimated depth from the tomographic inversion an a homogeneous replacement velocity of 2500 m/s between the RUS and $z = 0$, we were able to forward extrapolate the data, as computed through redatuming to the RUS, from the RUS back to a depth of $z = 0m$. After forward extrapolated, the structure disappears within the stack and clearly the relief of the Rus is recognized. If the stacked result of Figure 8.16 is compared to the stack of Figure 8.17, for which conventional processing was used to obtain source and receiver statics to be used for a floating datum correction, we can conclude that the CFP technology has a superior performance.

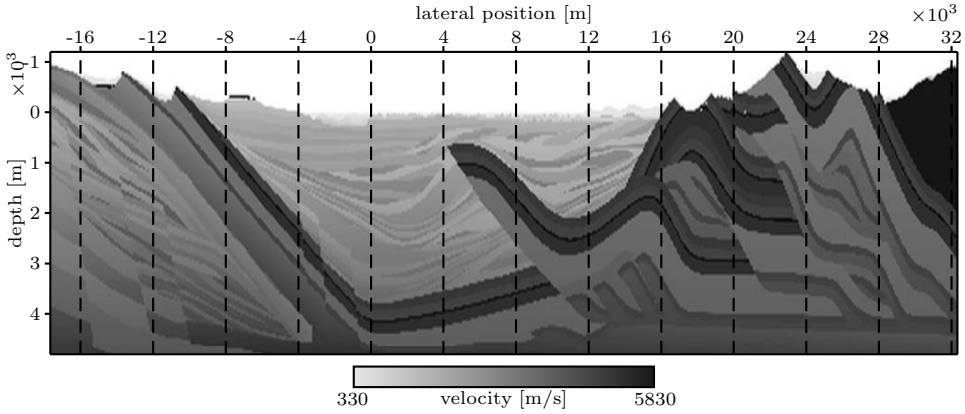

Figure 8.18: *The 2D model developed by BP to use for testing near-surface algorithms. The model contains a variety of geological features which can be encountered within real life seismic acquisition programs, including topology effects.*

8.3 The small wavelength near surface model

Within BP-Amoco a multi-purpose model has been developed of which its first purpose was to test currently available methods to solve for statics problems. For testing to be truly effective, it was necessary that the components of the model include problems which are difficult to resolve, even by the most sophisticated tools.

The full model is depicted in Figure 8.18. It includes mountainous thrust belts, a complex submarine fan system complicated by syn-depositional faulting, thrusted duplexes, blind thrusts and traps formed by unconformities and facies changes. Rock types that vary between carbonates and both hard and soft clastics were inserted in the model. As most refraction statics programs make many dubious assumptions, some of which are more appropriate for certain near-surface conditions than others, several types of anomalies were embedded in the near-surface of the model, such as buried channel systems, laterally discontinuous beds, buried volcanoes, near surface faulting and very rugged weathering profiles.

An attempt has been made to make the modeled data very similar in appearance to real data. This was done so that the processing steps would be applied under conditions similar to those encountered with real data. The model data does adhere to two key assumptions: it is strictly a 2-D model and only acoustic waves were

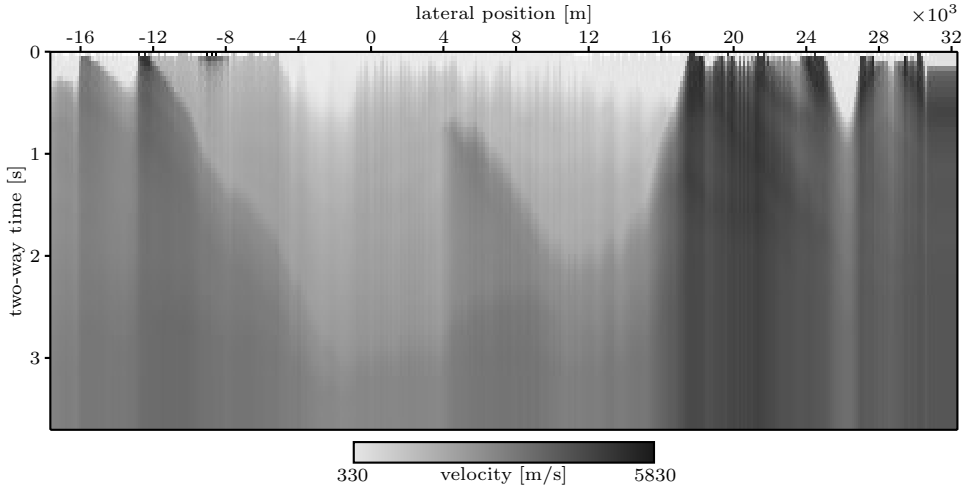

Figure 8.19: *The rms stacking velocities derived from the data modeled within the model of Figure 8.18.*

allowed to propagate through it.

The seismic data has been acquired on a 5x5m grid, with a frequency content of 0-65 Hz. The measured data has been summed into 25m receiver-arrays on an 8km cable without roll-in and roll-out. The shots were taken at the midpoint of the cable. A first passed analysis of stacking velocities is displayed in Figure 8.19. These velocities were used as input to a near offset time-migration of the full data set modeled within the full model as displayed in Figure 8.20. Note all the effects associated with the near surface anomalies, such as close to surface consistent dimming and brightening, lack of coherent stacking and hardly recognized deep reflection events.

The model and computed data is clearly sufficiently complex that the problems must be addressed in much the same way they would need to be addressed with real data.

Within this thesis we have restricted ourselves to only investigate part of the full model, running from 0-4000m. This part of the model has no change in topography, although this can be handled by CFP analysis equally well as demonstrated by the previous example. The near surface anomalies consist of a very rugged refractor profile. Figure 8.21A displays in detail the part of the full model that has been used for the CFP analysis. In Figure 8.21B a selection of shot records, modeled within this part of the model, has been displayed. Clearly, typical energy associated with land-seismics is present within the model, except for ground-roll, as shear energy

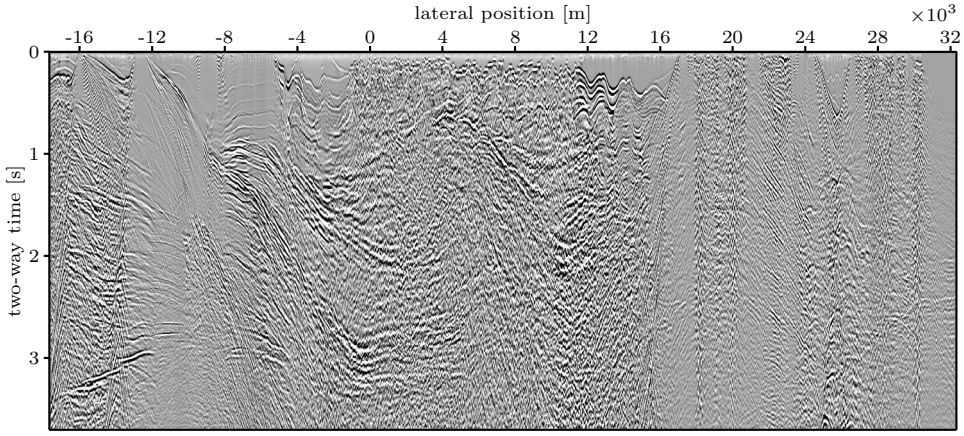

Figure 8.20: *The time-migrated near-offset section computed from the full dataset modeled over the model of Figure 8.18.*

does not travel through the model. Due to the rugged behavior of the near surface refractor, hardly any deeper reflections are clearly recognized.

■ 8.3.1 CFP processing

The first step in constructing operators to be input to the CFP-operator updating sequence is determining the stacking velocities and picking zero-offset times from the time-migrated near-offset section. Figure 8.22 displays the stacking velocities. As a consequence of presence of near-surface anomalies, these stacking velocities are not necessarily optimal, however, they only serve as an initial estimate on the move-out of the initial focusing operators. As outlined in Chapter 2, these velocities could be optimized by analyzing the crosscorrelations of the shot-records. This has not been applied within this example and would be recommended for future analysis.

Figure 8.23 displays the time-migrated near-offset section from which the zero-offset times were picked for the chosen datum reflector, as indicated in Figure 8.21A. Some muting and dip-filtering has been applied to the data to be able to better determine the zero-offset times. Remnants of the refraction energy result in the strong event as observed in the top of the section. Selecting a deeper reflector could have resulted in a better determination of zero-offset times as they would have been less obscured by refraction energy but would have resulted in a more complicated updating procedure.

Initial operators were constructed by using the zero-offset times and the stacking

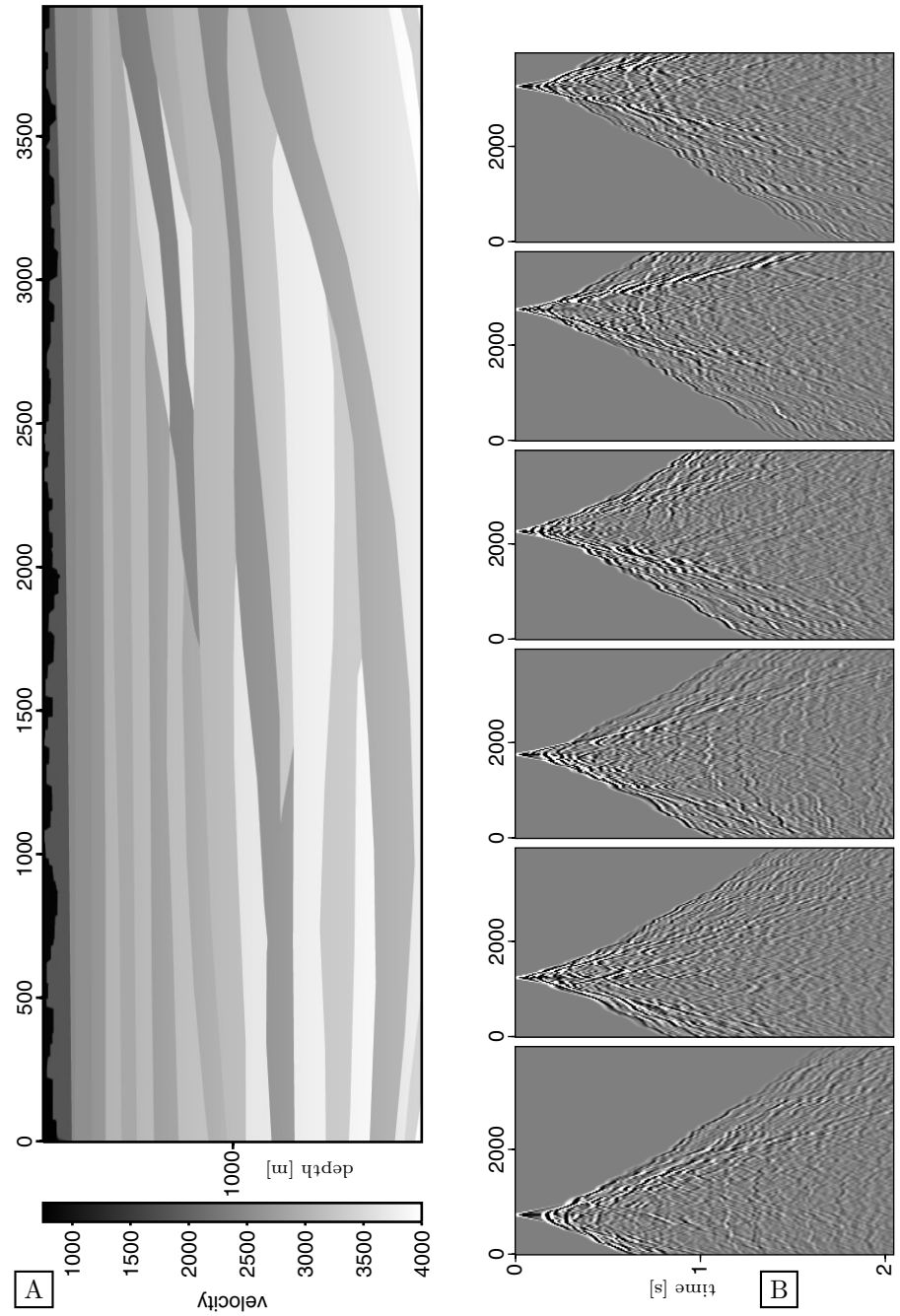

Figure 8.21: *A) The subset of the model of Figure 8.18, used for the near-surface analysis. B) Five shot-records computed by acoustic finite differencing within the model of (A).*

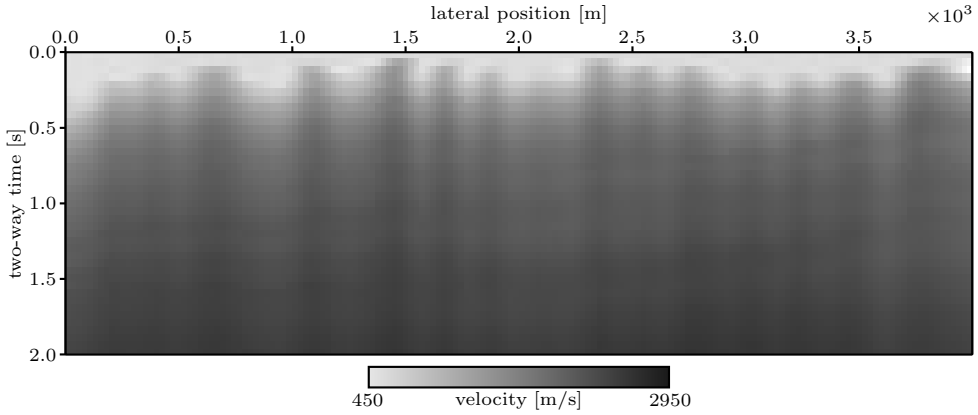

Figure 8.22: *The rms stacking velocity as determined for the subset-model of Figure 8.21.*

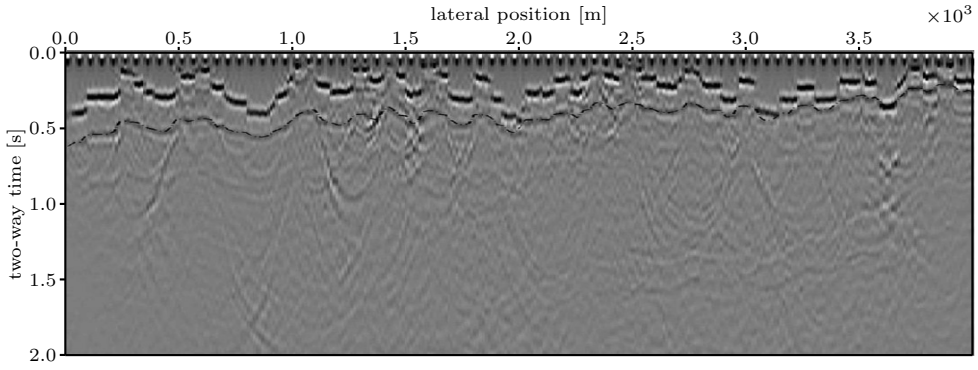

Figure 8.23: *The zero offset section for the submodel of Figure 8.21. The indicated zero offset times will be used in combination with the stacking velocities of Figure 8.22 to construct the initial operators.*

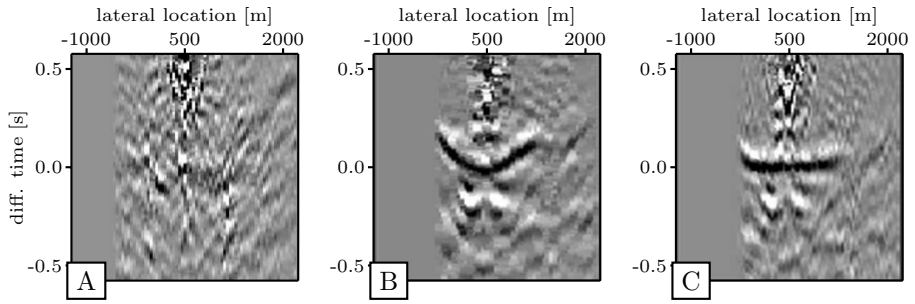

Figure 8.24: *DTS gather for lateral location x=500 constructed by using A) smooth initial operators, B) using transposed initial operators, C) using operators after 3 update iterations with transposed operators as initial operators.*

velocities. Both the focusing operators with smooth move-out as well as the transposed operators were applied to the data. Figure 8.24 displays the resulting DTS gathers for a location at x=500m. The DTS-gather resulting from the smooth operator, Figure 8.24A, hardly renders any energy that could be picked, whereas the transposed operators result in a smooth and easy to be picked DTS-gather, as depicted by Figure 8.24B. After only 3 updates this particular DTS-gather aligns almost perfectly around t=0s, as shown in Figure 8.24C. Clearly the very rugose nature of the near-surface refractor results in incoherent stacking during the CFP and DTS construction process when smooth operators are used, whereas the transposed operators already contain the major near-surface effects.

The final operators after three updating steps are displayed in Figure 8.25. These operators clearly contain fast variations caused by the small wavelengths of the near surface inhomogeneities. Within the next section, these operators will be used to redatum the data and to estimate a near surface model.

■ 8.3.2 CFP redatuming

The operators of Figure 8.25, retrieved after 3 updating steps have flattened the DTS gathers of Figure 8.26A sufficiently and are therefore considered to be the final operators that describe near surface propagation sufficiently well. In the next steps, the operators will be used to redatum both the sources as well as the receivers toward the near surface reflector just below the fast varying near surface inhomogeneities. Redatuming only the sources, is equivalent to constructing the CFP gathers. The CFP-gathers are displayed in Figure 8.26B. These gathers resemble measurements of the seismic energy at surface, excited by sources synthetically placed along the datum reflector. In a second step the receivers are redatumed as well, resulting in

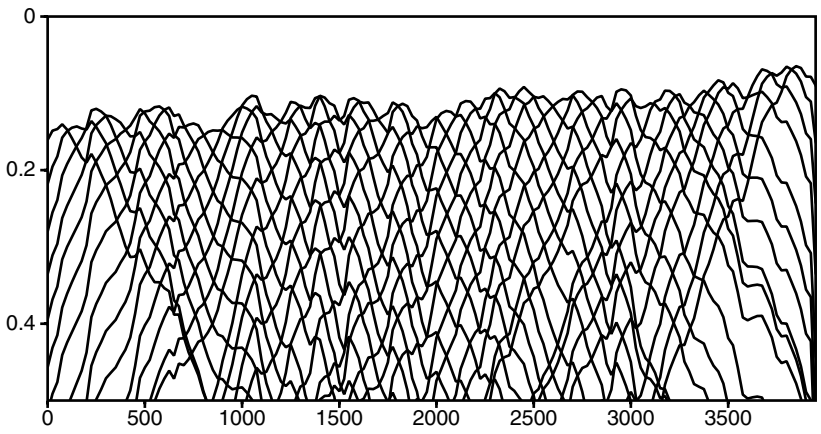

Figure 8.25: *Final operators after three updating steps.*

fully redatumed shot-gathers, of which a selection is displayed in Figure 8.26C. Note that the near surface effects have almost fully disappeared.

Next, the final operators have been subject to a tomographic inversion. In Figure 8.27 4 iterations have been displayed, the initial model, the model after two, after four and after six iterations of in total six iterations, of the non-linear model updating. The model as displayed within the final iteration closely resembles a low frequency representation of the true model. Also the position of the reflector has been resolved quite well.

After velocity analysis, application of move-out and stacking the redatumed data of Figure 8.26C, the stack of Figure 8.29 is obtained. When this result is compared to the stack of Figure 8.28, obtained by analyzing and stacking the surface data, it is clear that by using CFP technology to describe the propagation characteristics of the near-surface and using the obtained operators to redatum the data to a level below the near-surface, will render an enhanced resolution of the deeper reflectors. The smooth and sharp behavior of the reflected data within the stack of Figure 8.29 clearly resembles the geology of the model of Figure 8.21A quite well.

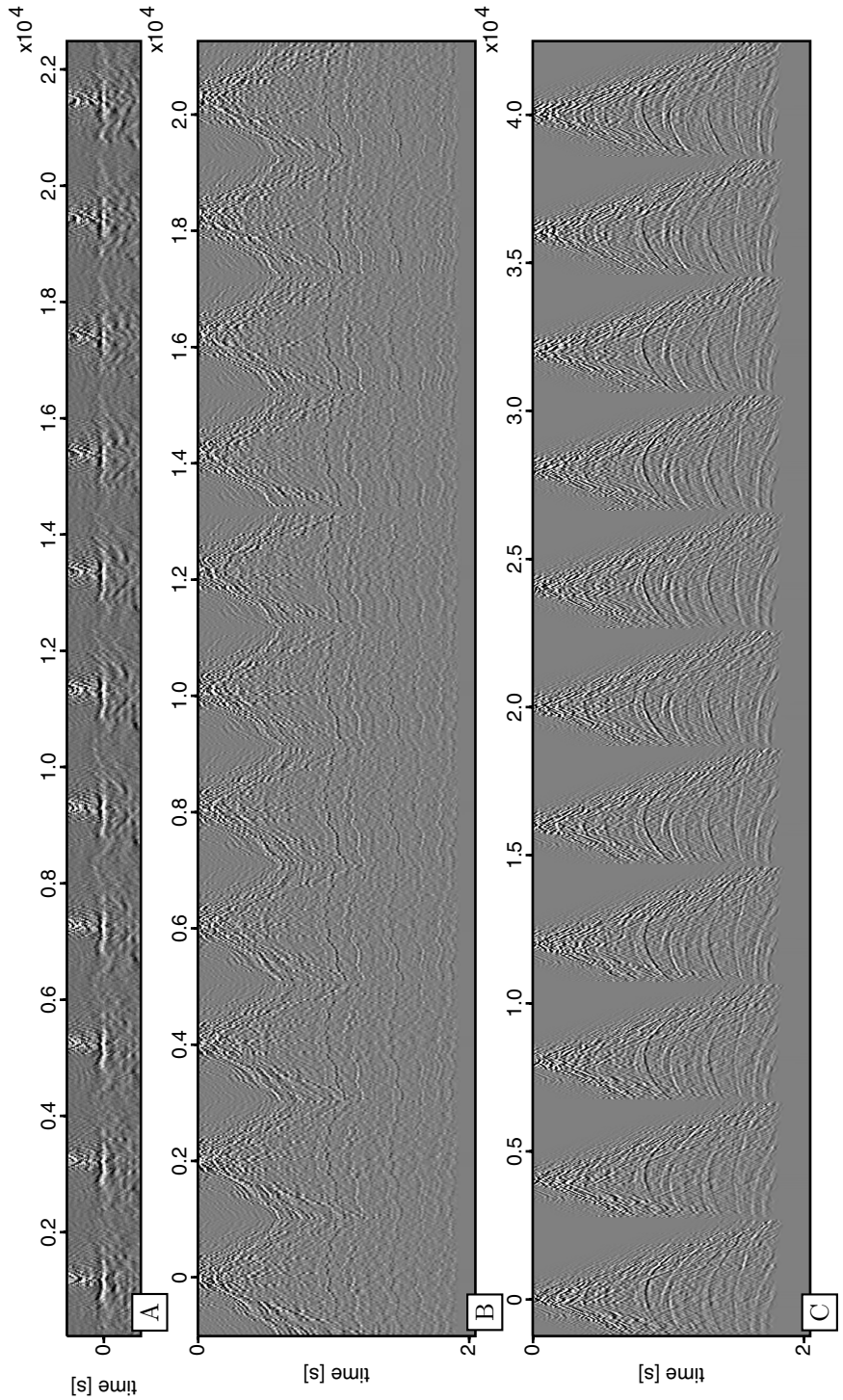

Figure 8.26: *Final CFP result after three update iteration using the transposed operators as initial operators. A) displays a set of DTS-gathers. B) Displays for the same locations the CFP-gathers. C) Displays for the same locations the synthetically redatumed shot-gathers.*

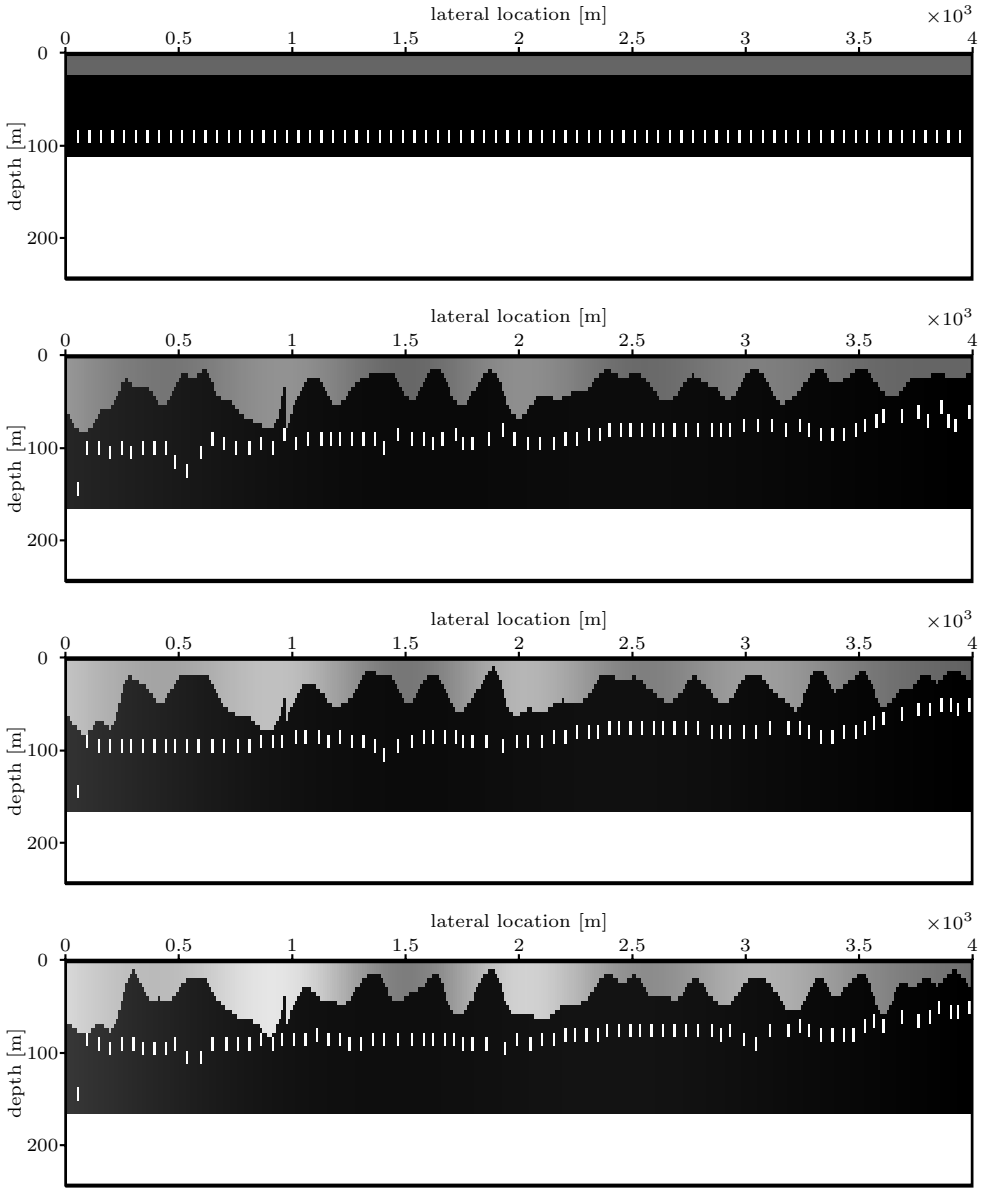

Figure 8.27: *Estimated near-surface model by inverting the final operators for four iteration within the nonlinear tomographic inversion algorithm, the initial model at the top followed by the model after two, four and six iterations at the bottom.*

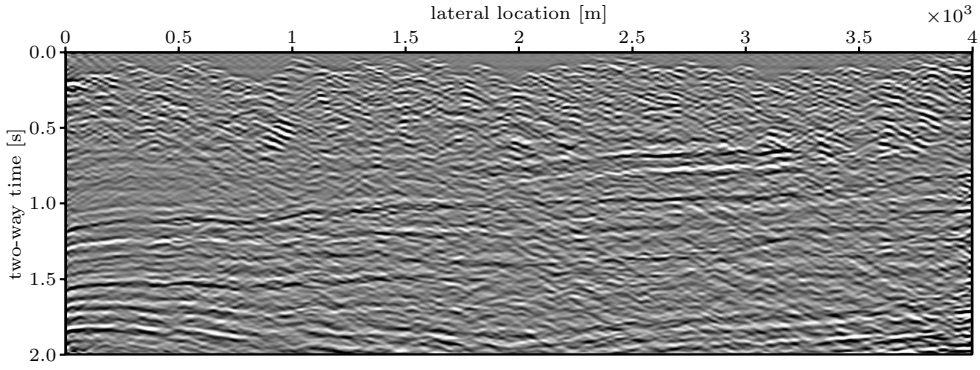

Figure 8.28: *Final stack of the data as measured at the surface.*

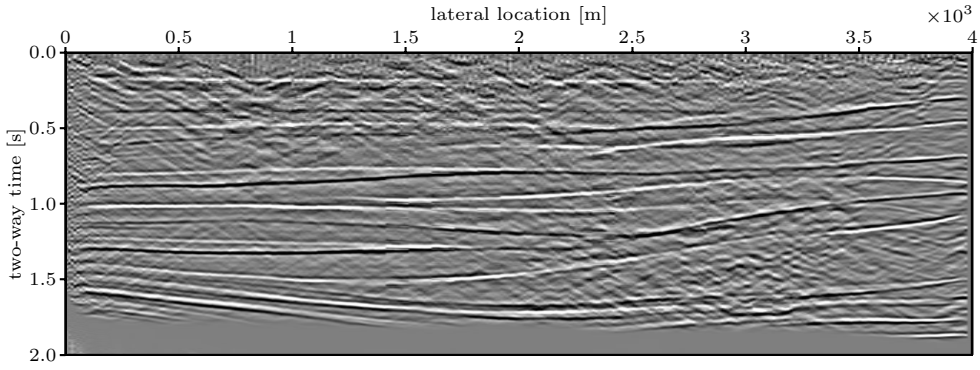

Figure 8.29: *Final stack of the data after redatuming to the shallow reflector just beneath the rugged near-surface refractor.*

9

Conclusions and recommendations

9.1 Conclusions

Conventional statics do not sufficiently solve the problem of near surface inhomogeneities. Applying such corrections may render an erroneous velocity profile and will misplace the final image during migration. Statics solutions are based on the assumption that all raypaths, regardless of originating depth or offset, will be close to the normal of the interface once entered into the low velocity near surface. This assumption will become more eroneous if the velocity contrasts between near surface anomalies and the underlying layer will become smaller. In case of velocity reversal (for instance within permafrost regions) results will be very wrong.

The near surface problem is a wave propagation problem and, therefore, redatuming should be used to solve for near surface effects. The Common Focus Point technology allows the estimation of operators that describe propagation through the near surface directly from the data, without estimating a near surface model first. As such, it is an ideal technology to remove the propagation effects of a complex near surface.

By using the concept of reciprocity, it is shown that wavefront healing can be used to our advantage when describing propagation between the surface and a datum reflector below the near surface. In this way we are able to parameterize a full set of laterally shifted operators with a sparse set of parameters. Applying a non-linear optimization algorithm, the near surface effects are recovered adequately by using a cost function based on stack power. Not only are we able able to include the near surface effects automatically, the hypothesis suggests that this method also avoids nonuniqueness.

If a conventional statics solution is available this can be incorporated in the initial operator. For this purpose, a new statics solution is proposed. The method correlates neighboring shot-gathers in time and sums the result along the receiver coordinates. The correlation removes the receiver statics. After summing along the receiver coordinates, the posistion of maximum energy shows the relative difference in shot-static when optimal normal move-out is applied to the data.

The focusing operators contain the propagation effects through the near surface but no information is obtained about the location of the new datum. Therefor, the estimated operators need to be input to a velocity model inversion. The inversion will solve for both the near surface velocities as well as the focuspoint locations. By separating the estimation of the common focuspoint operators from the near surface model determination, it is assured that the data itself will dictate the final redatuming result without being constraint by a near user provided surface model. The estimated near surface model merely explains the seismic propagation effects in the near surface.

The redatuming process is an inverse propagation problem. Therefore, not only phase information is needed in terms of traveltimes but also amplitude information. Excluding amplitude information has a similar effect as missing traces or measurements during redatuming or imaging. For near surface problems, the popular matched-filter approach appears to be insufficient. A least-squares redatuming procedure should be used, using operators that are weighted with spherical divergence, transmission losses and other near surface related amplitude effects.

The proper operator amplitudes can be derived through an optimization procedure which minimizes the difference between the forward modeled amplitudes and the measured data. This can be achieved by using Fermat modeling in combination with amplitude decomposition techniques. To properly cover offset dependent effects, the amplitude decomposition scheme has been expanded with higher order terms. If one does not adequately include the proper amplitudes one should refrain from a least-squares redatuming procedure and refer to the matched filter solution.

The method of Campman (2005), aiming at removing multiplicative noise due to near surface inhomogeneities, and the method outlined within this thesis turn out to be complementary. The method of Campman (2005) covers near surface short wavelength scattering effects whereas the method in this thesis covers the longer wavelength propagation effects.

9.2 Recommendations

Processing workflow

The method described in this thesis ties together several processing steps developed within the concept of the WRW-model. To process data acquired over a complex near surface it is recommended to use the following processing sequence:

- Optimize the acquisition as described by Al-Ali (2007).
- Removal of surface waves as described by Al-Ali (2007).
- Apply a temporal static correction as described in this thesis, Chapter 2.
- Construct initial operators as described in this thesis, Chapter 5.
- Improve operator updating as described by van der Rijzen et al. (2004).
- Estimate amplitudes as described in this thesis, Chapter 7.
- Apply a full 3D redatuming as described by Tegtmeier-Last (2007).
- Estimate a subsurface model as described by Cox (2004).

3D effects

In the real world, near surface inhomogeneities manifest themselves as three dimensional anomalies. This thesis has described the methodology for a two-dimensional situation. The next step is the extension to three dimensions. Especially for proper redatuming weights, the amplitudes should correct for three dimensional effects. With 2D assumptions severe artefacts can be introduced, as demonstrated in Chapter 7.

Field data calibration

It has been shown in Chapter 4 that the single focus point updating procedure does not necessarily result in a unique operator. Families of operators exist that satisfy the used criteria: the *principle of equal traveltime*, the match with the two-way traveltimes estimated from the one-way operators by Fermat modeling and the property that the inverse focusing operators applied to the forward operators should give a bandlimited pulse in the origin. The hypothesis is that not all of these operators are physically possible. By parameterizing the operators, as has been done in Chapter 5, a constraint is introduced. However, the problem of non-uniqueness can be solved by estimating a range of operators along a horizon simultaneously. Validation can be applied by sparsely measuring the physical operators in the field by using downhole receivers.

Computational efficiency in model estimation

When strong and relatively fast varying anomalies are expected in the near surface, it has been shown that grid-based methods are favorable over raytracing methods to compute the forward traveltimes. Derivatives of the traveltimes with respect to the model parameters have been derived through perturbation theory. This makes the inverse scheme more expensive in comparison to the use of raytracing algorithms. Especially, when the methodology is extended to three dimensions, the computational time will become a crucial factor to the applicability. It is therefore recommended to investigate alternative ways of computing the derivatives. This can be done for instance by carefully tracking the stencils used to derive connections and traveltimes when a shortest path algorithm is used.

Improving amplitude corrections

The redatuming procedure outlined in Chapter 7 was based on the determination of a single set of redatuming weights to be used for all frequencies. Moreover, the procedure was based on tracking a single event only. When the data becomes more complex and noisy, it might be worthwhile to include as much data as possible, to stabilize the procedure. This could lead to an amplitude matching algorithm applied to certain time-gates, including frequency dependency. This should be done in close relationship with the method developed by Campman (2005).

A

Two-way and one-way wavefields

A.1 Introduction

In this appendix the general theory for acoustic wavefield extrapolation and scattering is derived, which provides the base for redatuming the wavefield through the near surface. Both a two-way wavefield propagation formulation can be used as well as a one-way wavefield propagation formulation. A one-way formulation simplifies the expressions for extrapolation and scattering significantly in comparison to a two-way formulation.

First we will start with the constitutive equations, linearized under a low frequency approximation, describing acoustic wave-motion. The first equation is the equation of motion,

$$\partial_k \hat{p}(\mathbf{x},t) + \rho(\mathbf{x})\partial_t \hat{v}_k(\mathbf{x},t) = \hat{f}_k(\mathbf{x},t), \tag{A.1.1}$$

in which $\hat{p}(\mathbf{x},t)$ represents the scalar pressure field as function of space, $\mathbf{x}$, and time, t, and $\hat{v}_k(\mathbf{x},t)$ the vectorial velocity field excited by directional force given a spatially varying density $\rho(\mathbf{x})$.
The second equation is the equation of deformation,

$$\partial_k \hat{v}_k(\mathbf{x},t) + \kappa(\mathbf{x})\partial_t \hat{p}(\mathbf{x},t) = \hat{q}(\mathbf{x},t), \tag{A.1.2}$$

excited by a volume injection source $\hat{q}(\mathbf{x},t)$ given a spatially varying adiabatic compressibility modulus $\kappa(\mathbf{x})$. Equation A.1.1 and Equation A.1.2 are called the basic acoustic wave equations, which can be written in the space-frequency domain as,

$$\partial_k p(\mathbf{x},\omega) + j\omega\rho(\mathbf{x})v_k(\mathbf{x},\omega) = f_k(\mathbf{x},\omega), \tag{A.1.3}$$
$$\partial_k v_k(\mathbf{x},\omega) + j\omega\kappa(\mathbf{x})p(\mathbf{x},\omega) = q(\mathbf{x},\omega). \tag{A.1.4}$$

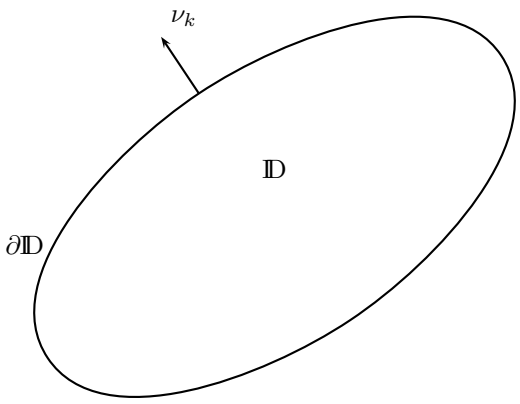

Figure A.1: *Definition of the volume used for evaluating the wave-equation and reciprocity relations, in which $\partial\mathbb{D}$ denotes the boundary of volume $\mathbb{D}$ for which ν_k is the outward pointing normal vector with unitary length.*

From the basic coupled acoustic equations, we can derive integral representations to describe wave-propagation within a known volume $\mathbb{D}$. Volume $\mathbb{D}$ is depicted in Figure A.1, with ν the unit vector along the normal of $\partial\mathbb{D}$ pointing away from the boundary of the discontinuity $\mathbb{D}$. At the boundary of volume $\mathbb{D}$ two pertaining boundary conditions can be defined. The first one reads,

$$\lim_{\epsilon \downarrow 0} \hat{p}(\mathbf{x} + \epsilon\nu, t) = 0, \qquad \mathbf{x} \in \partial\mathbb{D}, \tag{A.1.5}$$

which is the condition of a perfectly reflecting impenetrable void (pressure free surface), known as the Dirichlet condition. The second one reads,

$$\lim_{\epsilon \downarrow 0} \nu_k \hat{v}_k(\mathbf{x} + \epsilon\nu, t) = 0, \qquad \mathbf{x} \in \partial\mathbb{D}, \tag{A.1.6}$$

which is the condition of a perfectly reflecting impenetrable rigid object, known as the Neumann condition. With the above description of two-way wavefields, expressions can become complicated with the given boundary conditions. Further on expressions for simpler one-way wavefields will be derived by defining the medium parameters to be continuous and reflecting free, outside the volume V as a boundary condition.

A.2 Two-way acoustic wave equations

■ A.2.1 Two-way impulse response

Following Fokkema and van den Berg (1993), instead of considering the basic acoustic equations, Equation A.1.1 and Equation A.1.2, the homogeneous and source free

acoustic equations can be considered. Then, the volume injection source Green's states G^q and Γ_k^q are defined as the scalar impulse responses for the pressure and particle velocity,

$$\partial_k p^q(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\rho(\mathbf{x}) v_k^q(\mathbf{x}, \mathbf{x}^s, \omega) = 0, \tag{A.2.1a}$$
$$\partial_k v_k^q(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\kappa(\mathbf{x}) p^q(\mathbf{x}, \mathbf{x}^s, \omega) = q\delta(\mathbf{x} - \mathbf{x}^s). \tag{A.2.1b}$$

The volume force source Green's states G_l^f and $\Gamma_{k,l}^f$ are defined as the vector impulse responses for the pressure and particle velocity,

$$\partial_k p^f(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\rho(\mathbf{x}) v_k^f(\mathbf{x}, \mathbf{x}^s, \omega) = f_k\delta(\mathbf{x} - \mathbf{x}^s), \tag{A.2.2a}$$
$$\partial_k v_k^f(\mathbf{x}, \mathbf{x}^s, \omega) + j\omega\kappa(\mathbf{x}) p^f(\mathbf{x}, \mathbf{x}^s, \omega) = 0, \tag{A.2.2b}$$

with the Green's fields defined as,

$$\{p^q, v_k^q\} = q\{G^q, -\Gamma_k^q\}, \tag{A.2.3a}$$
$$\{p^f, v_k^f\} = f_l\{-G_l^f, \Gamma_{k,l}^f\}. \tag{A.2.3b}$$

Substituting Equation A.2.3a into Equation A.2.1a, leads to

$$\Gamma_k^q = \frac{1}{j\omega\rho}\partial_k G^q. \tag{A.2.4}$$

Substituting Equation A.2.3b into Equation A.2.1a, leads to

$$\Gamma_{k,l}^f = \frac{1}{j\omega\rho}\left[-\frac{1}{j\omega\rho}\partial_l\partial_k G^q + \delta_{l,k}\right]. \tag{A.2.5}$$

Substituting Equation A.2.3a into Equation A.2.1a, leads to

$$G_l^f = -\frac{1}{j\omega\rho}\partial_l G^q. \tag{A.2.6}$$

The Green's functions are expressed in terms of the volume injection pressure Green's function G^q, for which the following holds,

$$\rho\partial_k\left(\frac{1}{\rho}\partial_k G^q\right) + \frac{\omega^2}{c^2}G^q = -j\omega\rho\delta(\mathbf{x} - \mathbf{x}^R), \tag{A.2.7}$$

with the velocity defined by, $c^2(\mathbf{x}) = \frac{1}{\rho(\mathbf{x})\kappa(\mathbf{x})}$.

The expressions for the two-way Green's states of Equation A.2.4 through Equation A.2.7 can conveniently be written in terms of a matrix vector representation,

$$\partial_k\mathbf{G} - \mathbf{A}\mathbf{G} = \mathbf{S}, \tag{A.2.8}$$

	State A	State B
Field State	$\{p^A, v_k^A\}(\mathbf{x}, \omega)$	$\{p^B, v_k^B\}(\mathbf{x}, \omega)$
Material State	$\{\rho^A, \kappa_k^A\}(\mathbf{x})$	$\{\rho^B, \kappa_k^B\}(\mathbf{x})$
Source State	$\{q^A, f_k^A\}(\mathbf{x}, \omega)$	$\{q^B, f_k^B\}(\mathbf{x}, \omega)$

Table A.1: *The two states, used to define the two-way reciprocity relations.*

with,

$$\mathbf{G} = \begin{pmatrix} G^q & -G_l^f \\ -\Gamma_k^q & \Gamma_{k,l}^f \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 0 & -j\omega\rho \\ -j\omega\kappa & 0 \end{pmatrix}, \mathbf{S} = \begin{pmatrix} 0 & q \\ f_k & 0 \end{pmatrix} \delta(\mathbf{x} - \mathbf{x}^s). \tag{A.2.9}$$

with the elements of the Green's matrix $\mathbf{G}$ defined by Equation A.2.4 through Equation A.2.7.

■ A.2.2 Reciprocity relations for two-way wavefields

The reciprocity relations constitute the fundamentals of seismic wave-theory and can be used to describe the fundamental theory for near surface redatuming of wavefields. These relations describe the interaction between two states within a bounded medium D. The two states are described by their source characteristics, medium parameters and their wavefield distributions governed by the constitutional relations of Equation A.1.1 an Equation A.1.2, as shown by Table A.1. For an elaborate discussion on the Reciprocity theorems, the reader is referred to Fokkema and van den Berg (1993) and Rayleigh, J.W. Strutt, Baron (1877). Here, I briefly review the properties of the reciprocity theorems. To start with the derivation, first Gauss' integral theorem is considered, which relates a boundary integral to a volume integral,

$$\int_{\mathbf{x}\in\mathbb{D}} \partial_k U_k(\mathbf{x}) \mathrm{d}V = \int_{\mathbf{x}\in\partial\mathbb{D}} U_k(\mathbf{x}) \nu_k \mathrm{d}S. \tag{A.2.10}$$

From the constitutional equations of Equation A.1.1 and Equation A.1.2, we can define for the two interacting states,

$$\partial_k p^A(\mathbf{x}, \omega) + j\omega\rho^A(\mathbf{x}) v_k^A(\mathbf{x}, \omega) = f_k^A(\mathbf{x}, \omega), \tag{A.2.11a}$$
$$\partial_k v_k^A(\mathbf{x}, \omega) + j\omega\kappa^A(\mathbf{x}) p^A(\mathbf{x}, \omega) = q^A(\mathbf{x}, \omega), \tag{A.2.11b}$$

for state A, and

$$\partial_k p^B(\mathbf{x}, \omega) + j\omega\rho^B(\mathbf{x}) v_k^B(\mathbf{x}, \omega) = f_k^B(\mathbf{x}, \omega), \tag{A.2.12a}$$
$$\partial_k v_k^B(\mathbf{x}, \omega) + j\omega\kappa^B(\mathbf{x}) p^B(\mathbf{x}, \omega) = q^B(\mathbf{x}, \omega), \tag{A.2.12b}$$

for state B. For any surface discontinuity in the acoustic properties, Equation A.2.11a through Equation A.2.12a are supplemented by the boundary conditions as given in Section A.1. The field interaction analyzed by the reciprocity relations is chosen to be,

$$\begin{aligned}\partial_k \left(p^A(\mathbf{x})v_k^B(\mathbf{x}) - p^B(\mathbf{x})v_k^A(\mathbf{x})\right) = \\ v_k^B(\mathbf{x})\partial_k p^A(\mathbf{x}) + p^A(\mathbf{x})\partial_k v_k^B(\mathbf{x}) \\ -v_k^A(\mathbf{x})\partial_k p^B(\mathbf{x}) - p^B(\mathbf{x})\partial_k v_k^A(\mathbf{x}).\end{aligned} \tag{A.2.13}$$

If this relation is substituted into Gauss' theorem of Equation A.2.10, we arrive at

$$\begin{aligned}\int_{\partial D} \left(p^A(\mathbf{x})v_k^B(\mathbf{x}) - p^B(\mathbf{x})v_k^A(\mathbf{x})\right)\nu_k \mathrm{d}A = \\ \int_D \Big[v_k^B(\mathbf{x})\partial_k p^A(\mathbf{x}) + p^A(\mathbf{x})\partial_k v_k^B(\mathbf{x}) \\ -v_k^A(\mathbf{x})\partial_k p^B(\mathbf{x}) - p^B(\mathbf{x})\partial_k v_k^A(\mathbf{x})\Big]\mathrm{d}V.\end{aligned} \tag{A.2.14}$$

The right hand side of Equation A.2.14 can be rewritten by using the pertaining wave-equations of Equation A.2.11a through Equation A.2.12a, leading to,

$$\begin{aligned}\int_{\partial\mathbb{D}} \left[p^A(\mathbf{x})v^B(\mathbf{x}) - p^B(\mathbf{x})v^A(\mathbf{x})\right]\nu_k \mathrm{d}S = \\ \int_{\mathbb{D}} j\omega \left[\left(\rho^B - \rho^A\right)v_k^A(\mathbf{x})v_k^B(\mathbf{x}) - \left(\kappa^B - \kappa^A\right)p^A(\mathbf{x})p^B(\mathbf{x})\right]\mathrm{d}V \\ + \int_{\mathbb{D}} \left[f_k^A v_k^B + q^B pA - f_k^B v_k^A - q^A p^B\right]\mathrm{d}V,\end{aligned} \tag{A.2.15}$$

which is Rayleigh's reciprocity theorem in its global form. In case of equal material properties, the first term on the right-hand side vanishes. When the anti-causal wavefield of state A is taken, $\{p_A^*, v_{k,A}^*\}$ and the interaction quantity,

$$\begin{aligned}\partial_k \left(p^{*,A}(\mathbf{x})v_k^B(\mathbf{x}) + p^B(\mathbf{x})v_k^{*,A}(\mathbf{x})\right) = \\ v_k^B(\mathbf{x})\partial_k p^{*,A}(\mathbf{x}) + p^{*,A}(\mathbf{x})\partial_k v_k^B(\mathbf{x}) \\ - v_k^{*,A}(\mathbf{x})\partial_k p^B(\mathbf{x}) - p^B(\mathbf{x})\partial_k v_k^{*,A}(\mathbf{x}),\end{aligned} \tag{A.2.16}$$

we arrive at the power reciprocity theorem,

$$\begin{aligned}\int_{\partial\mathbb{D}} \left[p^{A,*}(\mathbf{x})v_k^B(\mathbf{x}) + p^B(\mathbf{x})v_k^{*,A}(\mathbf{x})\right]\nu_k \mathrm{d}S = \\ \int_{\mathbb{D}} -j\omega\left[\left(\rho^B - \rho^A\right)v_k^{*,A}(\mathbf{x})v_k^B(\mathbf{x}) + \left(\kappa^B - \kappa^A\right)p^{*,A}(\mathbf{x})p^B(\mathbf{x})\right]\mathrm{d}V \\ + \int_{\mathbb{D}}\Big[f_k^{*,A}(\mathbf{x})v_k^B(\mathbf{x}) + q^B(\mathbf{x})p^{*,A}(\mathbf{x}) + \\ f_k^B(\mathbf{x})v_k^{*,A}(\mathbf{x}) + q^{*,A}(\mathbf{x})p^B(\mathbf{x})\Big]\mathrm{d}V.\end{aligned} \tag{A.2.17}$$

Again, the first term on the right-hand side vanishes, when the material states are taken equal. These scalar-wave reciprocity theorems are used to derive wavefield representations of seismic data and the related forward model. They also prove a number of symmetry relations.

- By taking for state A field state $\{p^{q,A}, v_k^{q,A}\} = q^A\{G^q, \Gamma_k^q\}(\mathbf{x}, \mathbf{x}^A, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{q^A\delta(\mathbf{x}-\mathbf{x}^A), 0\}$ and for state B field state $q^B\{G^q, \Gamma_k^q\}(\mathbf{x}, \mathbf{x}^B, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{q^B\delta(\mathbf{x} - \mathbf{x}^B), 0\}$, the reciprocity theorem leads to,

$$q^B p^{q,A}(\mathbf{x}^B, \mathbf{x}^A) = q^A p^{q,B}(\mathbf{x}^A, \mathbf{x}^B), \tag{A.2.18}$$

$$G^q(\mathbf{x}^B, \mathbf{x}^A) = G^q(\mathbf{x}^A, \mathbf{x}^B). \tag{A.2.19}$$

- By taking for state A field state $\{p^{f,A}, v_k^{f,A}\} = f_l^A\{G_l^f, \Gamma_{k,l}^f\}(\mathbf{x}, \mathbf{x}^A, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{0, f_k^A\delta(\mathbf{x} - \mathbf{x}^A)\}$ and for state B field state $f_l^B\{G_l^f, \Gamma_{k,l}^f\}(\mathbf{x}, \mathbf{x}^B, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{0, f_k^B\delta(\mathbf{x} - \mathbf{x}^B)\}$, the reciprocity theorem leads to,

$$f_l^B v_l^{f,A}(\mathbf{x}^B, \mathbf{x}^A) = f_l^A v_l^{f,B}(\mathbf{x}^A, \mathbf{x}^B), \tag{A.2.20}$$

$$\Gamma_{k,l}^f(\mathbf{x}^B, \mathbf{x}^A) = \Gamma_{k,l}^f(\mathbf{x}^A, \mathbf{x}^B). \tag{A.2.21}$$

- By taking for state A field state $\{p^{q,A}, v_k^{q,A}\} = q^A\{G^q, \Gamma_k^q\}(\mathbf{x}, \mathbf{x}^A, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{q^A\delta(\mathbf{x} - \mathbf{x}^A), 0\}$ and for state B field state $f_l^B\{G_l^f, \Gamma_{k,l}^f\}(\mathbf{x}, \mathbf{x}^B, \omega)$, material state $\rho(\mathbf{x}), \kappa(\mathbf{x})$ and as source state $\{0, f_l^B\delta(\mathbf{x} - \mathbf{x}^B)\}$, the reciprocity theorem leads to,

$$f_l^B v_l^{q,A}(\mathbf{x}^B, \mathbf{x}^A) = q^A p^{f,B}(\mathbf{x}^A, \mathbf{x}^B), \tag{A.2.22}$$

$$\Gamma_l^q(\mathbf{x}^B, \mathbf{x}^A) = G_l^f(\mathbf{x}^A, \mathbf{x}^B). \tag{A.2.23}$$

A.3 One-way acoustic wave equation

As for instance imaging strategies, including redatuming, are based on back-propagation of primary one-way wavefields to their originating scattering contrast, it can be useful to use one-way reciprocity theorems. Moreover, by using two-way representations, the full Green's functions including, for instance, multiple energy have to computed. Slight errors in the background model can introduce severe errors as shown by Wapenaar and Berkhout (1989).

■ A.3.1 From two-way to one-way wavefields

For that purpose in this section, the one-way reciprocity theorem is derived, closely following Wapenaar and Grimbergen (1996). To derive the one-way reciprocity theorem, first a direction of interest has to be discriminated. Since in this thesis we are interested in redatuming of surface data, toward the target level, in the z-direction, it will be convenient to eliminate the two horizontal directions from the constitutive equations. The scalar wave-equations form a system of four equations,

$$\partial_1 p(\mathbf{x}) + j\omega\rho v_1(\mathbf{x}) = f_1(\mathbf{x}), \tag{A.3.1a}$$

$$\partial_2 p(\mathbf{x}) + j\omega\rho v_2(\mathbf{x}) = f_2(\mathbf{x}), \tag{A.3.1b}$$

$$\partial_3 p(\mathbf{x}) + j\omega\rho v_3(\mathbf{x}) = f_3(\mathbf{x}), \tag{A.3.1c}$$

$$\partial_1 v_1(\mathbf{x}) + \partial_2 v_2(\mathbf{x}) + \partial_3 v_3(\mathbf{x}) + j\omega\kappa p(\mathbf{x}) = q(\mathbf{x}), \tag{A.3.1d}$$

which can be rewritten as,

$$\partial_3 v_3(\mathbf{x}) = q - \partial_1 v_1(\mathbf{x}) - \partial_2 v_2(\mathbf{x}) - j\omega\kappa p(\mathbf{x}). \tag{A.3.2}$$

By substituting expressions for $v_1(\mathbf{x})$ and $v_2(\mathbf{x})$ taken from Equation A.3.1a and Equation A.3.1b, we arrive at,

$$\partial_3 v_3(\mathbf{x}) = \frac{1}{j\omega\rho}\mathcal{H}_2 p(\mathbf{x}) + q(\mathbf{x}) - \frac{1}{j\omega\rho}\left(\partial_1 f_1(\mathbf{x}) + \partial_2 f_2(\mathbf{x})\right), \tag{A.3.3}$$

in which we tacitly assumed the density ρ to be homogeneous, and the Helmholtz operator $\mathcal{H}_2$ being defined as,

$$\mathcal{H}_2 \cdot = \partial_1\partial_1 \cdot + \partial_2\partial_2 \cdot + k^2 \cdot . \tag{A.3.4}$$

The wavenumber k is defined by $k = \frac{\omega}{c}$, which is obtained by substituting $\kappa\rho c^2 = 1$. Thus, we can write the coupled wavefields $p(\mathbf{x})$ and $v_3(\mathbf{x})$ in matrix vector notation as,

$$\partial_3 \mathbf{Q} = \mathbf{A}\mathbf{Q} + \mathbf{D}, \tag{A.3.5}$$

in which $\mathbf{Q}$ contains the wavefield quantities,

$$\mathbf{Q} = \begin{pmatrix} p(\mathbf{x}) \\ v_3(\mathbf{x}) \end{pmatrix}, \tag{A.3.6}$$

matrix $\mathbf{A}$ the coupling operator,

$$\mathbf{A} = \begin{pmatrix} 0 & -j\omega\rho \\ \frac{-j}{\omega\rho}\mathcal{H}_2 & 0 \end{pmatrix}, \tag{A.3.7}$$

and $\mathbf{D}$ the source vector,

$$\mathbf{D} = \begin{pmatrix} f_3(\mathbf{x}) \\ q(\mathbf{x}) + \frac{j}{\omega\rho}(\partial_1 f_1(\mathbf{x}) + \partial_2 f_2(\mathbf{x})) \end{pmatrix}. \tag{A.3.8}$$

To decompose the wave-vector into up and down-going waves, the composition operator $\mathbf{L}$ is introduced. The decomposition is then carried out by the inverse, $\mathbf{L}^{-1}$. The coupling operator can be diagonalized through an Eigenvalue-decomposition,

$$\mathbf{A} = -j\omega\mathbf{L}\mathbf{\Lambda}\mathbf{L}^{-1}. \tag{A.3.9}$$

The factor $-j\omega$ is removed from the Eigen-value decomposition to assure equivalence with the differential equations for the two-way wave-equations. The diagonal matrix $\mathbf{\Lambda}$ contains the two Eigen-values,

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{1}{\omega}\mathcal{H}_1 & 0 \\ 0 & \frac{-1}{\omega}\mathcal{H}_1 \end{pmatrix}, \tag{A.3.10}$$

in which $\mathcal{H}_1$ is the pseudo-differential square-root operator, defined as,

$$\mathcal{H}_1\mathcal{H}_1 = \mathcal{H}_2. \tag{A.3.11}$$

Thus, the coupling operator consists of the Eigen-vectors and the decoupling operator is defined by the inverse, which through definition exists. Since an Eigen-vector is only one vector of a whole subspace of Eigen-vectors, the 2 by 2 coupling operator containing the Eigen-vectors is chosen consistently with the flux normalization in horizontally layered media as described in de Hoop (1992),

$$\mathbf{L} = \begin{pmatrix} L_1 & L_1 \\ L_2 & -L_2 \end{pmatrix}, \qquad \mathbf{L}^{-1} = \frac{1}{2}\begin{pmatrix} L_1^{-1} & L_2^{-1} \\ L_1^{-1} & -L_2^{-1} \end{pmatrix}, \tag{A.3.12}$$

in which

$$L_1 = \left(\frac{\omega\rho}{2\mathcal{H}_1}\right)^{\frac{1}{2}}, \qquad L_1^{-1} = \left(\frac{2\mathcal{H}_1}{\omega\rho}\right)^{\frac{1}{2}}, \tag{A.3.13}$$

$$L_2 = \left(\frac{\mathcal{H}_1}{2\omega\rho}\right)^{\frac{1}{2}}, \qquad L_2^{-1} = \left(\frac{2\omega\rho}{\mathcal{H}_1}\right)^{\frac{1}{2}}. \tag{A.3.14}$$

Substituting the Diagonal decomposition into Equation A.3.5, reads,

$$\partial_3\mathbf{Q} = -j\omega\mathbf{L}\mathbf{\Lambda}\mathbf{L}^{-1}\mathbf{Q} + \mathbf{D}. \tag{A.3.15}$$

The two-way wave vector $\mathbf{Q}$ can be expressed in terms of a one-way vector $\mathbf{P}$ by the aid of the composition matrix $\mathbf{L}$ through,

$$\mathbf{Q} = \mathbf{L}\mathbf{P}, \tag{A.3.16}$$

in which the one-way wave vector $\mathbf{P}$ consists of the up- and down-going waves, $\mathbf{P} = (p^+(\mathbf{x}), p^-(\mathbf{x}))^T$. In a similar way the source vector is composed from up and down-going sources, according to,

$$\mathbf{D} = \mathbf{L}\mathbf{S}, \tag{A.3.17}$$

If Equation A.3.16 and Equation A.3.17 are substituted in Equation A.3.15 the one-way system of wave equations is derived with a polarized direction along the z-axis,

$$\partial_3 \mathbf{L}\mathbf{P} = -j\omega \mathbf{L}\mathbf{\Lambda}\mathbf{L}^{-1}\mathbf{L}\mathbf{P} + \mathbf{L}\mathbf{S}, \tag{A.3.18}$$

$$\mathbf{P}\partial_3\mathbf{L} + \mathbf{L}\partial_3\mathbf{P} = -j\omega \mathbf{L}\mathbf{\Lambda}\mathbf{P} + \mathbf{L}\mathbf{S}, \tag{A.3.19}$$

$$\partial_3 \mathbf{P} = \mathbf{B}\mathbf{P} + \mathbf{S}, \tag{A.3.20}$$

$$\mathbf{B} = \mathbf{\Phi} - j\omega\mathbf{\Lambda}, \tag{A.3.21}$$

$$\mathbf{\Phi} = -\mathbf{L}^{-1}\partial_3\mathbf{L}. \tag{A.3.22}$$

The operator $\mathbf{B}$ contains two terms, first the pure diagonal part $-j\omega\Lambda$ which accounts for propagation of up- and down-going waves, and the coupling operator Φ which accounts for scattering due to vertical variation of medium parameters. Note that both terms account for scattering due to lateral variations. For a homogeneous earth $\mathbf{\Phi}$ vanishes and the one-way operator matrix $\mathbf{B}$ becomes a diagonal operator matrix and no longer couples the up and down-going waves, as expected. In case of inhomogeneous media, the up- and down-going wavefields are coupled through the coupling term $\mathbf{\Phi}$ which is formally written as,

$$\mathbf{\Phi} = \begin{pmatrix} \mathcal{T}^+ & \mathcal{R}^- \\ -\mathcal{R}^+ & -\mathcal{T}^- \end{pmatrix}, \tag{A.3.23}$$

in which T^+ and T^- account for transmission of up- and down-going waves respectively and R^+ and R^- account for reflection of up-going waves (into down-going waves) and down-going waves (into up-going waves) respectively .

Outside the volume V, the volume in which the one-way wave equations are analyzed, we choose as boundary condition for G a continuous, reflection free, fully absorbing media,

$$\kappa(x, y, z < z_0) = \kappa(x, y, z_0). \tag{A.3.24}$$

Thus, no down-going waves return from the upper half-space,

$$G = G^- \text{ at } z = z_0. \tag{A.3.25}$$

	State A	State B
Field State	$\{p^{+,A}, p^{-,A}\}(\mathbf{x}, \omega)$	$\{p^{+,B}, p^{-,B}\}(\mathbf{x}, \omega)$
Material State	$\mathbf{B}_A(\mathbf{x})$	$\mathbf{B}_B(\mathbf{x})$
Source State	$\{s^{+,A}, s^{+,A}\}(\mathbf{x}, \omega)$	$\{s^{-,B}, s^{+,B}\}(\mathbf{x}, \omega)$

Table A.2: *The two states, used to define the one-way reciprocity relations.*

■ A.3.2 One-way reciprocity theorem

Equivalent to the derivation of the two-way reciprocity theorem, again we would like to describe the interaction quantity of two field states as given in Table A.2. In contradiction to the two-way wavefield, the one-way wave-filed is described in terms of up and down-going wave, therefore, the interaction quantity to be investigated reads,

$$\partial_3 \left(p^{A,+}(\mathbf{x}) p^{B,-}(\mathbf{x}) - p^{A,-}(\mathbf{x}) p^{B,+}(\mathbf{x}) \right) = \partial_3 \mathbf{P}_A^T \bar{\mathbf{I}} \mathbf{P}_B, \tag{A.3.26}$$

in which $\bar{\mathbf{I}}$ is defined as,

$$\bar{\mathbf{I}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{A.3.27}$$

Then the interaction quantity of Equation A.3.26 leads to,

$$\partial_3 \mathbf{P}_A^T \bar{\mathbf{I}} \mathbf{P}_B = \mathbf{P}_B^T \bar{\mathbf{I}}^T \partial_3 \mathbf{P}_A + \mathbf{P}_A^T \bar{\mathbf{I}} \partial_3 \mathbf{P}_B. \tag{A.3.28}$$

If this relation is substituted into Gauss' theorem, we arrive at,

$$\int_{\partial \mathbb{D}} \mathbf{P}_A^T \bar{\mathbf{I}} \mathbf{P}_B \mu_3 \mathrm{d}A = \int_{\mathbb{D}} \mathbf{P}_B^T \bar{\mathbf{I}}^T \partial_3 \mathbf{P}_A + \mathbf{P}_A^T \bar{\mathbf{I}} \partial_3 \mathbf{P}_B \mathrm{d}V. \tag{A.3.29}$$

The right hand side of Equation A.3.29 can be written by using the pertaining one-way wave-equation of Equation A.3.20, leading to,

$$\int_{\partial \mathbb{D}} \mathbf{P}_A^T \bar{\mathbf{I}} \mathbf{P}_B \mu_3 \mathrm{d}A = \int_{\mathbb{D}} \mathbf{P}_A^T \bar{\mathbf{I}} \mathbf{S}_B + \mathbf{P}_B^T \bar{\mathbf{I}}^T \mathbf{S}_A \mathrm{d}V + \int_{\mathbb{D}} \mathbf{P}_A^T \left[\bar{\mathbf{I}} \mathbf{B}_B + \mathbf{B}_A^T \bar{\mathbf{I}} \right] \mathbf{P}_B \mathrm{d}V, \tag{A.3.30}$$

which is called the reciprocity relation of the convolution type for one-way wavefields. The term $\bar{\mathbf{I}} \mathbf{B}_B + \mathbf{B}_A^T \bar{\mathbf{I}}$ can be interpreted as a contrast operator. When considering the properties,

$$\bar{\mathbf{I}}^T \bar{\mathbf{I}} = \bar{\mathbf{I}} \bar{\mathbf{I}}^T = \mathbf{I}, \qquad \bar{\mathbf{I}}^{-1} = \bar{\mathbf{I}}^T, \tag{A.3.31}$$

the contrast operator can be rewritten as,

$$\psi = \left[\bar{\mathbf{I}} \mathbf{B}_B + \mathbf{B}_A^T \bar{\mathbf{I}} \right] = \bar{\mathbf{I}} \left[\mathbf{B}_B + \bar{\mathbf{I}}^T \mathbf{B}_A \bar{\mathbf{I}} \right]. \tag{A.3.32}$$

If also the symmetry-property of the medium operator $\mathbf{B}$ as derived by Wapenaar and Grimbergen (1996) is considered,

$$\mathbf{B}^H = -\bar{\mathbf{I}}\mathbf{B}^*\bar{\mathbf{I}}^{-1}, \tag{A.3.33}$$

the contrast operator ψ can be written as,

$$\psi = \mathbf{B}_B - \mathbf{B}_A. \tag{A.3.34}$$

Hence, when the two states are identical, the contrast term vanishes in the reciprocity relation of the convolution type for one-way wavefields. Furthermore, since the contrast function for one-way wavefields is composed of a propagation part and a scattering part, an independent choice of propagation and scattering is allowed to be analyzed.

To derive the reciprocity relation of the correlation type for one-way wavefields, the following interaction relation is considered,

$$\partial_3 \mathbf{P}_A^H \bar{\mathbf{I}} \mathbf{P}_\mathbf{B}. \tag{A.3.35}$$

Note that instead of the transpose of $\mathbf{P}_A$, as in Equation A.3.26 here the Hermitian transpose of $\mathbf{P}_A$ is taken. Then applying again Gauss' theorem leads in a similar way as in the previous section to,

$$\int_{\partial\mathbb{D}} \mathbf{P}_A^H \bar{\mathbf{I}} \mathbf{P}_B \mu_3 \mathrm{d}A \approx \int_{\mathbb{D}} \mathbf{P}_A^H \left(\mathbf{B}_B - \mathbf{B}_A\right) \mathbf{P}_B \mathrm{d}V + \int_{\mathbb{D}} \mathbf{P}_A^H \bar{\mathbf{I}} \mathbf{S}_B + \mathbf{P}_B^H \bar{\mathbf{I}}^T \mathbf{S}_A \mathrm{d}V. \tag{A.3.36}$$

The approximation sign in Equation A.3.36 refers to the use of an approximation in the derivation of a modified symmetry property of the medium operator $\mathbf{B}$; the evanescent waves are ignored (Wapenaar and Grimbergen, 1996). For further detailed comparison of the one-way reciprocity theorem and the two-way reciprocity theorem, the reader is referred to Wapenaar and Grimbergen (1996) and Hegge (2001).

Both the convolution type as well as the correlation type reciprocity relations for one-way wavefields are determined by the difference in medium operator $\mathbf{B}$ for the different states. In Equation A.3.21 the medium operator is explicitly defined in terms of a propagation part $-j\omega\Lambda$ and a scattering part Φ. Through the reciprocity relations we can chose to analyze those parts independently. This is major advantage over the two-way reciprocity theorems and will be used advantageously to set up the forward model, which is referred to as the WRW-model by Berkhout (1982).

■ A.3.3 One-way impulse responses

In analogy with the two-way Green's functions (Equation A.2.4 through Equation A.2.7), the one-way Green's functions $G^{+,+}$ and $G^{-,+}$ are defined as the up-going

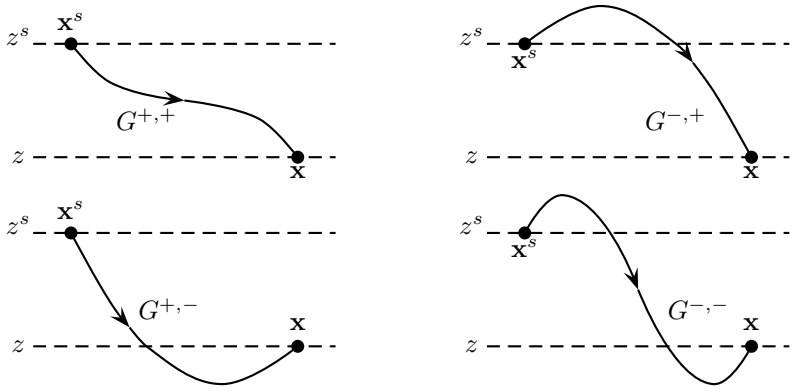

Figure A.2: *The elements of the one-way Greens function. The elements $G^{-,+}$ and $G^{+,-}$ are due to coupling effects, when scattering contrast are present.*

impulse responses of the one-way wave-equation and the Green's functions for the up-going wavefield and the coupled down-going wavefield. Similarly, the Green's functions $G^{-,-}$ and $G^{+,-}$ are defined as the down-going impulse response of the one-way wave equation for the down-going wavefield and the coupled up-going wavefield. The Green's state, satisfying the one-way wave equation can then be written in a matrix form as,

$$\partial_3 \mathbf{G} - \mathbf{B}\mathbf{G} = \mathbf{I}\delta(\mathbf{x} - \mathbf{x}^s), \tag{A.3.37}$$

in which the Green's matrix is defined as,

$$\mathbf{G}(\mathbf{x}, \mathbf{x}^s) = \begin{pmatrix} G^{+,+}(\mathbf{x}, \mathbf{x}^s) & G^{+,-}(\mathbf{x}, \mathbf{x}^s) \\ G^{-,+}(\mathbf{x}, \mathbf{x}^s) & G^{-,-}(\mathbf{x}, \mathbf{x}^s) \end{pmatrix}. \tag{A.3.38}$$

The elements of the Green's matrix are visualized in Figure A.2.

The choice of the medium operator $\mathbf{B}$ determines which elements of the Green's matrix are put under investigation. By taking $\mathbf{\Lambda}$ equal to the actual medium and $\mathbf{\Phi} = 0$, within the domain under investigation, coupling of wavefields is not taken into account. Thus only propagation of the different Green's components is taken into account,

$$\partial_3 \mathbf{G} + j \begin{pmatrix} \mathcal{H}_1 & 0 \\ 0 & -\mathcal{H}_1) \end{pmatrix} \mathbf{G} = \mathbf{I}\delta(\mathbf{x} - \mathbf{x}^s). \tag{A.3.39}$$

When only outgoing waves are considered in the domain under investigation, the diagonal elements of $\mathbf{G}$ vanish and we are left with the primary Green's function,

$$\mathbf{G}^p(\mathbf{x},\mathbf{x}^s) = \begin{pmatrix} \mathrm{h}(x_3 - x_3^s)W^+(\mathbf{x},\mathbf{x}^s) & 0 \\ 0 & -\mathrm{h}(x_3 - x_3^s)W^-(\mathbf{x},\mathbf{x}^s) \end{pmatrix}, \qquad \text{(A.3.40)}$$

in which h is the Heaviside function and W^+ and W^- are the primary propagators which describe downward and upward propagation. Note that due to the specific choices of the medium operator and the Green's function, the Green's state of Equation A.3.37 purely describes up and downward propagation of only up and downward going primary Green's functions. As coupling does not occur it allows for simple expressions for forward and inverse propagation of up and down-going wavefields. If use is made of a two-way formulation of forward and inverse propagation, the coupling effects as present in the Green's states need to included explicitly, which requires the computation of the full Green's function within the medium under consideration.

B

Bayes based parameter estimation

B.1 Introduction

In this thesis extensive use is made of inversion theory to obtain focusing operators, a near surface velocity model, operator amplitudes and optimally redatumed data. These steps comprise both linear as well as non-linear inverse theory. Within this appendix a thorough description is given of the different methods. The base of the inversion algorithms is given by Bayes' theorem, which will be discussed first. From this the so-called MAP estimator can be derived, which leads to a linear inverse formulation, which can be optimized by a proper choice of probability density functions. The derived system of normal equations can then be solved efficiently through Krylov subspace methods of which LSQR is one of the most efficient ones. For non-linear inversion, one can either linearize locally within the model space, such that a full non-linear inversion comprises of small linear inversion steps, for which the previously described methods hold. This works particularly well if we are not too far from the optimal solutions (not too non-linear, no multiple local optimal solutions). For very non-linear systems, of which we are not sure how far we are from the optimum solution, we need to restrict ourselves to global optimization methods like simulated genetic evolution. The appendix ends with an example of a genetic algorithm applied to a conventional static problem.

B.2 Bayes' rule

The Bayesian formalism Duijndam (1988a,b) tends to combine and describe the various sources of information available, such as various dependent and independent experiments as well as estimates of both measurement and model errors, in an elegant way.

Here, the various aspects of Bayesian estimation are reviewed. It gives clear insight how various experiments such as reflection and refraction experiments can be incorporated. Furthermore it will give a nice recipe for incorporation of various probability density functions in a clear sense.

The probability that a certain event $\tilde{\lambda}$ lies within the interval $[a, b]$ can be defined as,

$$P(a \leq \tilde{\lambda} \leq b) = \int_a^b p(\lambda) \mathrm{d}\lambda, \tag{B.2.1}$$

in which $p(\lambda)$ is the probability density function (pdf) of the n-dimensional random variable $\lambda \in L_\lambda^n$ with the property,

$$\int_{L_\lambda^n} p(\lambda) \mathrm{d}\lambda = 1. \tag{B.2.2}$$

In case of more than one random variable, let's say two variables λ and γ, we can define a joint pdf $p(\lambda, \gamma)$ from which the marginal pdf's can be derived through,

$$p(\gamma) = \int_{L_\lambda^n} p(\lambda, \gamma) \mathrm{d}\lambda. \tag{B.2.3}$$

In case of independent variables λ and γ, the joint pdf equals the product of the marginal pdf's. The conditional pdf of λ, given the observed value γ is defined as,

$$p(\gamma|\lambda) = \frac{p(\gamma, \lambda)}{p(\lambda)}. \tag{B.2.4}$$

With the conditional pdf $p(\lambda|\gamma)$ defined alike, Bayes' theorem can be derived, yielding

$$p(\lambda|\gamma) = \frac{p(\gamma|\lambda)p(\lambda)}{p(\gamma)}. \tag{B.2.5}$$

If we take γ to be our measured data (or one-way traveltimes from the subsequent subsurface Common Focal Points) and λ the model-parameters we would like to estimate, then Bayes' theorem states the probability density of the parameter space given our measured data and its accuracy and some sort of a priori information we have on the model parameters. In practical situations it appears almost impossible to compute $p(\gamma)$ but since we are interested in evaluating the model space, $p(\gamma)$ can be regarded as a constant in our inversion process.

$p(\lambda)$ resembles the a priori information we have on our model-parameters and describes possible first guesses and uncertainties or even hard constraints, like velocities can not be negative. The conditional pdf $p(\gamma|\lambda)$ is called the likelihood function. This function depends on the model which describes the relation between the random data variable γ and the unknown parameter vector λ. In our case, the measured

traveltimes $\gamma = \mathbf{d}$ depend on the subsurface model λ through the non-linear (theoretic) relation,

$$\gamma = \mathbf{m}(\lambda) + \mathbf{n}, \tag{B.2.6}$$

in which the noise term $\mathbf{n}$ resembles both theoretical errors as well as observational errors.
For a given measured data set γ the likelihood function is determined by the uncertainties and properties of the noise,

$$p(\gamma|\lambda) = p_n(\mathbf{n}) = p_n(\mathbf{d} - \mathbf{m}(\lambda)), \tag{B.2.7}$$

with p_n the pdf of the random noise distribution.

Bayes' theorem is merely a descriptive theorem, providing the basic blocks and insights of an inversion problem. A nice aspect is that Bayes' theorem is not restricted to describe probabilities of one set of measurements to one set of variables. Consider a new experiment with measurements ϕ, independent from the previous measurements γ then the joint probability reads,

$$p(\gamma, \phi) = p(\gamma)p(\phi), \tag{B.2.8}$$

and the conditional pdf's for the two experiments,

$$p(\gamma|\phi) = p(\gamma), \tag{B.2.9}$$
$$p(\phi|\gamma) = p(\phi), \tag{B.2.10}$$

then using the chain rule,

$$p(\phi, \gamma, \lambda) = p(\phi|\gamma, \lambda)p(\gamma, \lambda), \tag{B.2.11}$$
$$= p(\phi|\lambda)p(\gamma, \lambda), \tag{B.2.12}$$
$$= p(\phi|\lambda)p(\gamma|\lambda)p(\lambda). \tag{B.2.13}$$

Thus,

$$p(\lambda|\gamma, \phi)p(\gamma, \phi) = p(\lambda, \gamma, \phi), \tag{B.2.14}$$
$$p(\lambda|\gamma, \phi) = \frac{p(\lambda, \gamma, \phi)}{p(\gamma)p(\phi)}, \tag{B.2.15}$$
$$p(\lambda|\gamma, \phi) = \frac{p(\phi|\lambda)p(\gamma|\lambda)p(\gamma)}{p(\gamma)p(\phi)}. \tag{B.2.16}$$

From Equation B.2.16 it is concluded that independent measurements are easily incorporated by a chain rule. For m independent experiments, the a posteriori pdf can be written as

$$p(\lambda|\gamma_1, \cdots, \gamma_m) = p(\lambda)\frac{\prod_{i=1}^{m} p(\gamma_i|\lambda)}{\prod_{i=1}^{m} p(\gamma_i)}. \tag{B.2.17}$$

B.3 Characterization of Probability Density functions

The character of randomness of parameters can be described by certain properties of the probability density functions, such as the expectation, the median and the covariance of the distribution of events.

Expectation

The expectation or mean of the random variable λ with respect to its pdf $p(\lambda)$ is given by (Tarantola, 1987),

$$\bar{\lambda} = E(\lambda) = \int_{L_\lambda^n} \lambda p(\lambda) \mathrm{d}\lambda, \tag{B.3.18}$$

$$\rightarrow \text{MIN}\left[\int_{L_\lambda^n} \|\lambda - \bar{\lambda}\|_2^2, p(\lambda)\mathrm{d}\lambda\right], \tag{B.3.19}$$

and resembles a minimum L_2-norm.

Median

The median of the random variable λ with respect to its pdf $p(\lambda)$ is given by (Tarantola, 1987),

$$\hat{\lambda} = E(\lambda) = \int_{L_\lambda^n} \lambda p(\lambda) \mathrm{d}\lambda, \tag{B.3.20}$$

$$\rightarrow \text{MIN}\left[\int_{L_\lambda^n} \|\lambda - \bar{\lambda}\|_1 p(\lambda)\mathrm{d}\lambda\right], \tag{B.3.21}$$

and resembles a minimum L_1-norm.

Covariance

The covariance matrix $\mathbf{C}$ is defined by its elements (Tarantola, 1987),

$$C_{ij} = E(\lambda_i - \bar{\lambda}_j)(\lambda_i - \bar{\lambda}_j) = \int_{L_\lambda^n} (\lambda_i - \bar{\lambda}_j)^2 p(\lambda)\mathrm{d}\lambda. \tag{B.3.22}$$

The diagonal of $\mathbf{C}$ contains the variance σ^2, of which the square roots are defined as the standard deviations. It can easily be shown that median, mean, variance and standard deviations of one element λ_i are equal to those of the marginal pdf. The correlation coefficient ρ_{ij} of variables λ_i and λ_j are defined by,

$$\rho ij = \frac{C_{ij}}{\sigma_i \sigma_j}, \tag{B.3.23}$$

and define a measure to what extent certain variables are correlated or independent.

To formulate an inverse solution using Bayes' theorem, the user should define a probability density function, describing the probability a certain stochastic variable will have a specific value.

B.4 MAP estimator

When the data measured data can be described by,

$$\mathbf{d} = \mathbf{g}(\mathbf{m}) + \mathbf{n}, \tag{B.4.24}$$

in which $\mathbf{d}$ denotes the measurements, $\mathbf{g}(\mathbf{m})$ the forward modeling system acting on the model parameters $\mathbf{m}$ and $\mathbf{n}$ represents both measurement errors as well as theoretical errors, then using Bayes' theorem leads to the following a posteriori probability density function,

$$p(\mathbf{m}|\mathbf{d}) = p_n(\mathbf{d} - \mathbf{g}(\mathbf{m}))p(\mathbf{m}), \tag{B.4.25}$$

defining the a posteriori probability density function (pdf) $p(\mathbf{m}|\mathbf{d})$ of the model space $\mathbf{m}$ given the measurements $\mathbf{d}$. Properties of the noise are contained in the likelihood pdf p_n. A priori knowledge about the model parameters is contained in the marginal pdf, $p(\mathbf{m})$. Since we would like to find the most probable model, a practical solution to the inverse problem would be choosing the maximum of the a posteriori density function, or Maximum A Posteriori (MAP) estimator. When no a priori information on the model parameters is taken into account, setting $p(\mathbf{m}) = 1$, the MAP estimation is equivalent to Maximum Likelihood (ML) estimation; a uniform a priori distribution is chosen, taken as the state of null information. Properties on ML and MAP estimation are extensively discussed by Bard (1974). To define a quantitative solution to the MAP estimator, the probability density function on both noise, $p_n(\mathbf{d} - \mathbf{g}(\mathbf{m}))$ as well as model parameters, $p(\mathbf{m})$, has to be formulated.

■ B.4.1 Probability density function

Since Bayesian theorem is only provides a descriptive framework, an explicit choice of a proper pdf has to made in order to obtain useful expressions applicable for the inverse problems tackled in this thesis. In nature, random processes are often very well described by Gaussian distributions,

$$p(\mathbf{m}) = \frac{1}{(2\pi)^{n/2}\|\mathbf{C}\|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{m} - \bar{\mathbf{m}})^T\mathbf{C}^{-1}(\mathbf{m} - \bar{\mathbf{m}})\right], \tag{B.4.26}$$

which is completely described by the covariance matrix $\mathbf{C}$ and the mean value $\bar{\mathbf{m}}$ only, leading to simple mathematical expressions and containing the least information as determined by Shannon (1948).

As a consequence of the Gaussian description by variance and mean only, no knowledge of a scatter measure is used. This makes the Gaussian distribution non-robust. The need for a robust estimator comes to question when outliers are expected. In every day practice normal distributions are expected and a statistician would analyze the data and remove significant outliers on forehand. Robust estimators estimate models that are not sensitive to small changes in measurements. Non-robust estimators tend to leave residuals with glanally and locally low amplitudes, thus outliers have an unwanted effect on the outcome model. Outliers can be expected within the process of the determination of focusing operators, e.g. local pick errors due to absence of energy or frustration due to interference of several events are commonly present. Furthermore, in the process of amplitude determination boundary effects play a large role. Therefore, we will use when applicable, estimators derived from the class of type M-estimators.

In general, an estimator $\hat{m}$ should maximize the likelihood,

$$\hat{m} = \{m | \prod p(m) \geq \prod p(\gamma), \gamma \in \partial\mathbb{R}\}, \tag{B.4.27}$$

which states that the product of the probability of the components of the estimator is larger than the product of the probabilities of any other set of elements within the parameter space. Equivalently one can write, when for instance the Gaussian distribution of Equation B.4.26 is used,

$$\hat{m} = \{m | \sum -\ln p(m) \leq \sum -\ln p(\gamma), \gamma \in \partial\mathbb{R}\}, \tag{B.4.28}$$

transforming the problem into a minimization problem, called the Gaussian estimator. The Gaussian estimator is a member of the family of so called type M Estimators, which have the more general form,

$$\hat{m} = \{m | \sum f(\hat{m}) \leq \sum f(\gamma), \gamma \in \partial\mathbb{R}\}. \tag{B.4.29}$$

Note that Equation B.4.27 through Equation B.4.29 assume no interdependency between parameters. For the more general cases, including correlation between parameters, the reader is referred to Rey (1983) and Huber (1964). In this thesis we will consider exponential pdf's belonging to the group of M-estimators. Although the Gaussian estimator belongs to the group of M-estimators as well, the Gaussian is non robust whereas the exponential distribution has more robust properties.

With the statement of Equation B.2.27 that the estimator $\hat{m}$ fulfills the requirement that the product of the probability of the components of the estimator is larger than the product of the probabilities of any other set of elements within the parameter space, we can define the objective function as,

$$F(\mathbf{m}) = \sum f_i(m_i), \tag{B.4.30}$$

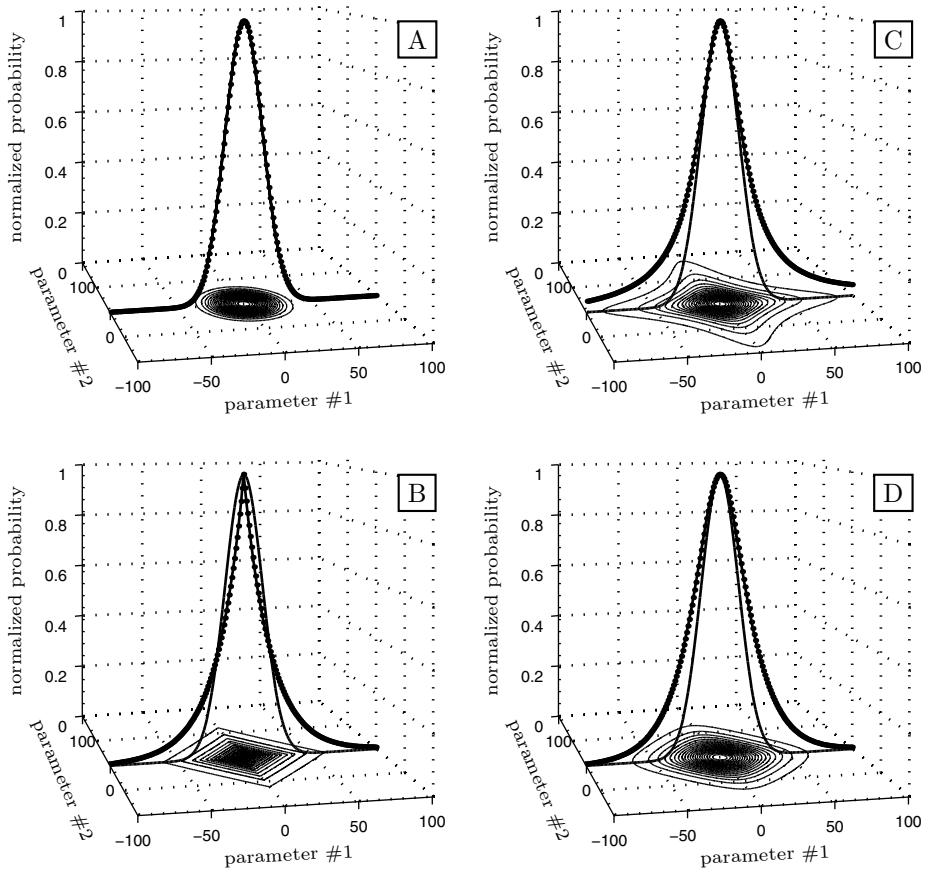

Figure B.1: *Different types of probability density functions for duo-variate (2 parameter) distributions. The horizontal axes denote the parameter distribution, the vertical axis the probability of a certain combination to occur. For all distributions the mean value is (0,0). Although for correct interpretation the integral of a pdf should be equal to one, the pdf's are scaled to their maximum for comparison reasons. A) Gaussian distribution. B) normal distribution compared to the Gaussian distribution. C) Cauchy distribution compared to the Gaussian distribution. D) Hybrid distribution compared to the Gaussian distribution.*

as defined by equation B.2.29, for which we have to find the minimum in order to define the most probable set of parameters. Following Huber (2003), Rey (1983) and Bube and Langan (1997), mathematical feasible objective functions of the form $F(\mathbf{m}) = \sum f_i(x_i)$ should obey the following restrictions,

- The objective should be minimum and equal to zero when the estimator equals zero,
$$F(\mathbf{0}) = 0. \tag{B.4.31}$$
- The minimum should be absolute,
$$F(\mathbf{m}) > 0, \text{ for } \mathbf{m} \neq \mathbf{0}. \tag{B.4.32}$$
- The objective function should be even,
$$F(-\mathbf{m}) = F(\mathbf{m}). \tag{B.4.33}$$
- $F(\mathbf{m}) \to \infty$ as $\mathbf{m} \to \infty$.
- F and $\frac{\partial F}{\partial \mathbf{m}}$ should exist and be continuous.
- $\frac{\partial^2 F}{\partial \mathbf{m}^2}$ should exist, be finite and positive definite, or in other words be strictly convex.

In literature various robust estimators are described, but do not obey to all requirements. A good estimator, which does obey the requirements, which has the convenient properties of the least squares estimator derived from a Gaussian distribution and which has the robustness of the exponential distribution, is the hybrid norm estimator as described in Bube and Langan (1997) and derived from the following distribution,

$$\begin{aligned} p(\mathbf{m}) &= \prod \exp(-f_i(m_i)), & \text{(B.4.34)} \\ &= \exp(-\sum f_i(m_i)), & \text{(B.4.35)} \\ f_i(m_i) &= \sqrt{1 + (\frac{m_i}{\sigma_i})^2} - 1. & \text{(B.4.36)} \end{aligned}$$

Then the estimator becomes,

$$F(\mathbf{m}) = \sum (\sqrt{1 + (\frac{m_i}{\sigma_i})^2} - 1), \tag{B.4.37}$$

which behaves like a Gaussian distribution for small $\mathbf{m}$ and as an exponential distribution for large $\mathbf{m}$,

$$f_i(m_i - \hat{m_i}) \approx \begin{cases} \frac{1}{2}\frac{(m_i - \hat{m_i})^2}{\sigma_i^2} & \text{for } \|(m_i - \hat{m_i})\| \text{ is small} \\ \frac{\|(m_i - \hat{m_i})\|}{\sigma_i} & \text{for } \|(m_i - \hat{m_i})\| \text{ is large} \end{cases}. \tag{B.4.38}$$

This type of objective functions effectively filter the data from bad, unexplained measurements. When applicable the hybrid norm probability density function will be used, denoted as

$$p_h = \text{constant} \times \exp(-\sum f_i(\mathbf{m}_i - \hat{\mathbf{m}}_i)), \tag{B.4.39}$$

or the Gaussian distribution, denoted as

$$p_g = \text{constant} \times \exp(-\sum \sigma_i^{-1}(\mathbf{m}_i - \hat{\mathbf{m}}_i)^2), \tag{B.4.40}$$

in which σ_i are the autocorrelation coefficients. Note that for both distributions, no correlation between measurements is assumed. One could choose to use the definition for the Gaussian of Equation B.4.26 which does include crosscorrelations. However, in practice very expensive noise tests need to be carried out to fill the covariance matrices completely and one would suffice only predicting the autocorrelation coefficients.

In Figure B.1 four possible possible probability density functions for duo-variate distributions are depicted for which the described objective functions can be defined. Figure B.1A displays the non-robust Gaussian pdf, which is displayed for comparison in the other three figures as well. Figure B.1B displays the normal distribution, which clearly shows a discontinuous behavior of the derivative around the expected value, which makes it a non-feasible objective function. In Figure B.1C the hybrid-norm distribution is displayed and in Figure B.1D the Cauchy distribution, which is another robust distribution found extensively in literature.

B.5 Linear inverse formulation

Now that the probability density functions are defined we can set up the inversion problem as given by Equation B.4.25. For now we assume a linear forward model, thus,

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}, \tag{B.5.41}$$

in which the linear forward problem, $\mathbf{g}(\mathbf{m})$ is written as a linear operator $\mathbf{G}$ acting on the unknown set of parameters $\mathbf{m}$. Non-linear optimization is subject to the next section. If for both noise as well model parameters, the Gaussian pdf is taken, then the a posteriori pdf would be,

$$p(\mathbf{m}|\mathbf{d}) = p_g(\mathbf{d} - \mathbf{G}\mathbf{m})p_g(\mathbf{m} - \hat{\mathbf{m}}), \tag{B.5.42}$$

with $\hat{\mathbf{m}}$ the a priori model values. Maximizing the a posteriori pdf is equivalent to minimizing the negative logarithm, taking into account Equation B.4.40,

$$\begin{aligned} F &= -\ln p_g(\mathbf{d} - \mathbf{G}\mathbf{m})p_g(\mathbf{m} - \hat{\mathbf{m}}) && \text{(B.5.43)} \\ &= (\mathbf{d} - \mathbf{G}\mathbf{m})^T\mathbf{C}_n^{-1}(\mathbf{d} - \mathbf{G}\mathbf{m}) + (\hat{\mathbf{m}} - \mathbf{m})^T\mathbf{C}_m^{-1}(\hat{\mathbf{m}} - \mathbf{m}), && \text{(B.5.44)} \end{aligned}$$

with the definition of Equation B.4.40, $\mathbf{C}_n$ and $\mathbf{C}_m$ are diagonal matrices, containing the inverse of the autocorrelation of the noise and a priori parameters respectively. Since the objective function F has a quadratic form it will contain one root only. Minimizing the objective function F can be done by finding its root,

$$\begin{aligned} \frac{\partial F}{\partial \mathbf{m}} &= \mathbf{0} && \text{(B.5.45)} \\ &= -2\mathbf{G}^T\mathbf{C}_n^{-1}\mathbf{d} + 2\mathbf{G}^T\mathbf{C}_n^{-1}\mathbf{G}\mathbf{m} - 2\mathbf{C}_m^{-1}\hat{\mathbf{m}} + 2\mathbf{C}_m^{-1}\mathbf{m}, && \text{(B.5.46)} \end{aligned}$$

and the parameter estimate under Gaussian assumptions is found to be,

$$\mathbf{m} = (\mathbf{G}^T\mathbf{C}_n^{-1}\mathbf{G} + \mathbf{C}_m^{-1})^{-1}(\mathbf{G}^T\mathbf{C}_n^{-1}\mathbf{d} + \mathbf{C}_m^{-1}\hat{\mathbf{m}}). \quad \text{(B.5.47)}$$

Instead of taking Gaussian distributions, the probability density function of Equation B.4.39 for the M-estimator of Equation B.4.37 is taken for the noise distribution and a Gaussian distribution for the parameter distribution,

$$p(\mathbf{m}|\mathbf{d}) = p_h(\mathbf{d} - \mathbf{G}\mathbf{m})p_g(\mathbf{m} - \hat{\mathbf{m}}). \quad \text{(B.5.48)}$$

Maximizing the a posteriori pdf is again equivalent to minimizing the negative logarithm, taking into account Equation B.4.39,

$$\begin{aligned} F &= -\ln p_h(\mathbf{d} - \mathbf{G}\mathbf{m})p_g(\mathbf{m} - \hat{\mathbf{m}}), && \text{(B.5.49)} \\ &= f(\mathbf{d} - \mathbf{G}\mathbf{m}) + (\hat{\mathbf{m}} - \mathbf{m})^T\mathbf{C}_m^{-1}(\hat{\mathbf{m}} - \mathbf{m}). && \text{(B.5.50)} \end{aligned}$$

Computing the gradient is not as straightforward as could be done when a Gaussian distribution is used. For this purpose the influence function is introduced,

$$g(r) = \frac{\partial f(r)}{\partial r}\frac{1}{r}. \quad \text{(B.5.51)}$$

Then with $\mathbf{r} = \mathbf{d} - \mathbf{G}\mathbf{m}$,

$$\begin{aligned} \frac{\partial f(r_i)}{\partial m_j} &= \frac{\partial f(r_i)}{\partial r_i}\frac{\partial r_i}{\partial m_j}, && \text{(B.5.52)} \\ &= G_{ij}g(r_i)r_i. && \text{(B.5.53)} \end{aligned}$$

Thus obtaining

$$\begin{aligned} \frac{\partial F}{\partial \mathbf{m}} &= 0, && \text{(B.5.54)} \\ &= \mathbf{G}^T\mathbf{W}(\mathbf{d} - \mathbf{G}\mathbf{m}) - 2\mathbf{C}_m^{-1}\hat{\mathbf{m}} + 2\mathbf{C}_m^{-1}\mathbf{m}, && \text{(B.5.55)} \end{aligned}$$

in which $\mathbf{W}$ is a diagonal matrix with $W_{ii} = g(r_i)$. When the estimator of Equation B.4.37 is adopted then the weights read,

$$\begin{aligned} W_{ii} &= \frac{1}{\sqrt{1 + \frac{r_i^2}{\sigma_i^2}}}, && \text{(B.5.56)} \\ \mathbf{r} &= \mathbf{d} - \mathbf{G}\mathbf{m}, && \text{(B.5.57)} \end{aligned}$$

and the estimated parameters are defined by,

$$\mathbf{m} = -(\mathbf{G}^T\mathbf{W}\mathbf{G} + \mathbf{C}_m^{-1})^{-1}(\mathbf{G}^T\mathbf{W}\mathbf{d} + \mathbf{C}_m^{-1}\hat{\mathbf{m}}), \tag{B.5.58}$$

with $\mathbf{W}$ as defined by Equation B.5.56. When comparing the latter with the Gaussian model estimate with the result of the M-estimator their results are comparable: the model dependent weighting matrix W of the M-estimator replaces the constant covariance matrix $\mathbf{C}$ of the Gaussian estimator. For the remainder we will define the objective function as,

$$F = \mathbf{e}^T\mathbf{e}, \tag{B.5.59}$$

$$\mathbf{e} = \begin{bmatrix} \mathbf{W}_1^{\frac{1}{2}}(\mathbf{d}_n - \mathbf{G}_1\mathbf{m}) \\ \vdots \\ \mathbf{W}_n^{\frac{1}{2}}(\mathbf{d}_n - \mathbf{G}_n\mathbf{m}) \end{bmatrix}, \tag{B.5.60}$$

in which several experiments are combined through the Bayes derived chain rule (see Section B.1). If Gaussian distributions are used, this expression leads directly through Equation B.5.44. If the hybrid distribution distribution is used, the formulation is derived by recognizing the similarity between Equation B.5.47 and Equation B.5.58, with the notion that the weights, $\mathbf{W}$, are depending on the parameters, $\mathbf{m}$. The objective function derived from the a posteriori pdf of Equation B.5.47 would lead for instance to,

$$F = \mathbf{e}^T\mathbf{e}, \tag{B.5.61}$$

$$\mathbf{e} = \begin{bmatrix} \mathbf{W}^{\frac{1}{2}}(\mathbf{d}_n - \mathbf{G}_1\mathbf{m}) \\ \mathbf{C}_m^{-\frac{1}{2}}(\hat{\mathbf{m}} - \mathbf{I}_n\mathbf{m}) \end{bmatrix}. \tag{B.5.62}$$

It should be noted that the objective function of Equation B.5.50 is non-linear since the weighting matrix $\mathbf{W}$ depends on the parameters. To solve this problem, iteratively reweighted least-squares are used: we start of with an initial weighting matrix, defined by $\mathbf{C}_n$ and compute the solution for the obtained least squares problem. Next, the weights are recomputed for the current parameter set and the parameters are recomputed with the newly defined least squares problem. This procedure is further iterated until a sufficiently small data residual is obtained. The actual solution to the least-squares step is discussed in the next section.

■ B.5.1 Comparison of performance between hybrid norm and Gaussian norm

To demonstrate the effect of reweighting the next simple experiment is considered. A random, linear forward operator $\mathbf{A}$ was created, mapping a parameter vector $\mathbf{m}$,

$$\mathbf{m} = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]^T, \tag{B.5.63}$$

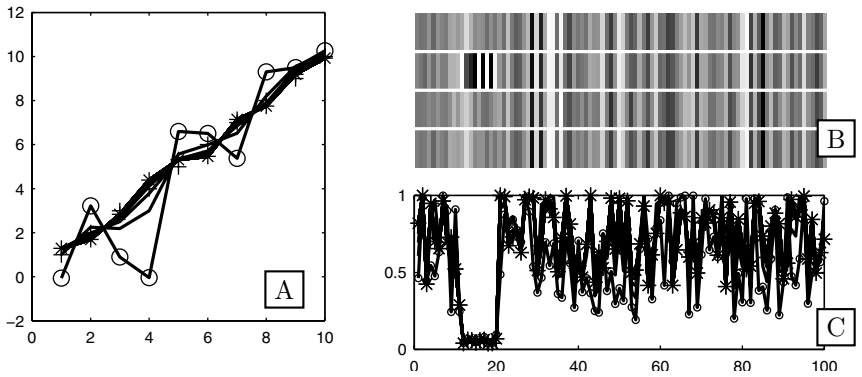

Figure B.2: *A random, linear forward operator* $\mathbf{A}$*, maps A) parameter vector* $\mathbf{m}$ *to B) a data-set. The 4 datasets in B) are realizations of respectively (from top to bottom) the exact parameters, the exact parameters with noise, parameters estimated by using a Gaussian norm and parameters estimated using a hybrid norm. C) The evolution of the weighting matrix as occurring in the hybrid norm formulation.*

into 100 data-realizations, contained in vector $\mathbf{d}$, of which the amplitudes are displayed in the top row of Figure B.2B. Next, 5 percent noise was added and an extra, extreme error on measurements 11 to 20, displayed in the second row of Figure B.2B. The least-squares solution to this problem, in finding the parameters, is depicted by circles in Figure B.2A. The parameters are severely deteriorated from the true solution. By remodeling the data from the least-squares solution, shown by the third row of Figure B.2B., the weights can be constructed for the first iteration of the reweighted scheme. Since least-squares assumes the data to have a normal distribution, the excessive errors of samples 11 through 20 are smeared over all data-realizations. Thus, the weights of the first iteration, denoted by the dots in Figure B.2C. already significantly suppress the mentioned data-points. Further iterating leads to the solution denoted by the stars in Figure B.2C. The remodeled data is shown by the fourth row of Figure B.2B. The data is almost fully reconstructed. Moreover, the estimated parameters denoted by the stars in Figure B.2A are close to the correct values, whereas the Gaussian solution renders large errors in the estimated parameters. This example clearly demonstrates the fact that the Gaussian estimator is sensitive to outliers whereas the hybrid estimator tends to take the outliers reasonably well into account. When the noise is assumed to be uncorrelated, the covariance matrix of the noise $\mathbf{C}$ can be replaced by the following diagonal matrix,

$$C_{ij} = \begin{cases} 0, & \text{for } i \neq j \\ \sigma_i^2 & \text{for } i = j \end{cases}. \tag{B.5.64}$$

■ B.5.2 Krylov subspace methods for linear system inversion

To invert the system of, for instance, Equation B.2.57,

$$\mathbf{m} = -(\mathbf{G}^T\mathbf{W}\mathbf{G} + \mathbf{C}_m)^{-1}(\mathbf{G}^T\mathbf{W}\mathbf{d} + \mathbf{C}_m\hat{\mathbf{m}}), \tag{B.5.65}$$

directly could be a costly procedure when we have to deal with large systems. Instead of computing the normal equations directly and invert the system of equations by means of for instant Gauss-Seidel iterations or LU-decomposition (Golub and van Loan, 1996), methods exist that try to compute the parameter vector in a least-squares manner implicitly. Such methods are called Krylov-subspace methods (Golub and van Loan, 1996; Strang, 1988). These iterative methods exploit the solution space efficiently such that each iteration updates the parameter vector in a direction perpendicular to the error vector of the current iteration. The majority of these methods are the so called Conjugate Gradient methods. LSQR, developed by Paige and Saunders (1982), is one of such methods as well. Within tomography problems, LSQR is widely accepted as the best appropriate method.

Krylov subspace methods do not compute the full normal equations, they only compute once or twice a vector product with the supplied system matrix. This is extremely efficient when one has to deal with sparse matrices. Furthermore, the number of iterations never exceeds the number of parameters (after N iterations in search of N parameters, N different orthogonal directions have been exploited, thus spanning the complete parameter space), and the first iterations tend to solve for the largest Eigenvalues first.

To solve the system $\mathbf{b} = \mathbf{A}\mathbf{x}$ the complete algorithm is defined by,

- Initialization of the system,

$$\begin{aligned}
\beta_1\mathbf{u}_1 &= \mathbf{b}, && \text{(B.5.66)}\\
\alpha_1\mathbf{v}_1 &= \mathbf{A}^T\mathbf{u}_1, && \text{(B.5.67)}\\
\mathbf{w}_1 &= \mathbf{v}_1, && \text{(B.5.68)}\\
\mathbf{x}_0 &= \mathbf{0}, && \text{(B.5.69)}\\
\bar{\phi}_1 &= \beta_1, && \text{(B.5.70)}\\
\bar{\rho}_1 &= \alpha_1. && \text{(B.5.71)}
\end{aligned}$$

- Continuation of a bidiagonalization procedure,

$$\begin{aligned}
\beta i + 1\mathbf{u}_{i+1} &= \mathbf{A}\mathbf{v}_i - \alpha_i\mathbf{u}_i, && \text{(B.5.72)}\\
\alpha_{i+1}\mathbf{v}_{i+1} &= \mathbf{A}^T\mathbf{u}_{i+1} - \beta_{i+1}\mathbf{v}_i. && \text{(B.5.73)}
\end{aligned}$$

- Construction and application of the next orthogonal transformation,

$$\begin{aligned}
\rho_i &= \sqrt{\hat{\rho}_i^2 + \beta_{i+1}^2}, && \text{(B.5.74)}\\
c_i &= \frac{\hat{\rho}_i}{\rho_i}, && \text{(B.5.75)}\\
s_i &= \frac{\beta_{i+1}}{\rho_i}, && \text{(B.5.76)}\\
\chi_{i+1} &= s_i \alpha_{i+1}, && \text{(B.5.77)}\\
\hat{\rho}_{i+1} &= -c_i \alpha_{i+1}, && \text{(B.5.78)}\\
\phi_i &= c_i \hat{\phi}_i, && \text{(B.5.79)}\\
\hat{\phi}_{i+1} &= s_i \hat{\phi}_i, && \text{(B.5.80)}
\end{aligned}$$

- updating of the parameter vectors,

$$\begin{aligned}
\mathbf{x}_i &= \mathbf{x}_{i-1} + \frac{\phi_i}{\rho_i}\mathbf{w}_i, && \text{(B.5.81)}\\
\mathbf{w}_{i+1} &= \mathbf{v}_{i+1} - \frac{\chi_{i+1}}{\rho_i}\mathbf{w}_i. && \text{(B.5.82)}
\end{aligned}$$

- Returning to the next bidiagonalization step (Equation B.2.70) and following steps, increasing the iteration number i.

After k iterations the algorithm has produced a bidiagonal matrix $\mathbf{S}$,

$$\mathbf{S} = \begin{bmatrix}
\alpha_1 & 0 & \cdots & & & \\
\beta_1 & \alpha_2 & \ddots & & & \\
0 & \beta_2 & \alpha_3 & \ddots & & \\
\vdots & \ddots & \ddots & \ddots & & \\
 & & & & \alpha_{k-1} & \\
 & & & & \beta_{k-1} & \alpha_k
\end{bmatrix}, \qquad \text{(B.5.83)}$$

which is close to the diagonal matrix containing the Ritz values of $\mathbf{A}$. Furthermore, two orthogonal matrices, $\mathbf{U}$ and $\mathbf{V}$ are obtained, for which the following holds,

$$\begin{aligned}
\mathbf{A} &= \mathbf{USV}, && \text{(B.5.84)}\\
&= \mathbf{U\acute{U}\Lambda\acute{V}V}, && \text{(B.5.85)}
\end{aligned}$$

in which $\mathbf{\Lambda}$ is a diagonal matrix containing the Ritz values and $\mathbf{U\acute{U}}$ containing the Ritz vectors, when $\mathbf{A}$ is square. Since the Ritz pairs are close to the k largest Eigenvalues, one could suffice to only compute a few iterations, thus enforcing numerical stabilization. Moreover, for the largest Ritz values, one could then compute the

resolution and covariance of the parameters. In this research we are not directly interested in those values apart from determining a posteriori properties of our system. However, Cox (2004) uses the intermediate properties during the iterative non-linear optimization to reparameterize the velocity model in a data driven manner. For a more elaborate discussion on the covariances and residuals of the data and parameters the reader is referred to Cox (2004).

B.6 Non-linear optimization

Within inversion procedures, the objective is to minimize[1] the objective function; a set of parameters, describing the modeling of our experiment, needs to be found which minimizes the weighted norm of the difference between the measured data and the modeled data, a priori information on the parameters and estimated parameters. Through the Bayesian approach statistical properties of measuring and modeling are incorporated.

Non-linear inversion is used in situations where a priori no information on the minima, curvature and gradient of the objective function is available. Various procedures exist in finding minima of non-linear objective functions. The most straight forward method would be evaluating the objective function at every point within the feasible region. In practice this method is not favorable. In Chapter 4 use was made of a genetic algorithm to find a set of focusing operators through the principle of data matching. The genetic algorithm is based on the simulation of the evolution of populations of possible models. Genetic algorithms require a large amount of model evaluations. If the forward modeling is expensive, genetic algorithms are in general not feasible. Instead of using genetic algorithms, one can also use local optimization algorithms. These type of algorithms iterate through the model space by linearizing the non-linear objective function around the current location in model space for each iteration. Linearized optimization methods find a minimum far more efficient but a risk remains to get trapped in a local minimum. Opposite to the problem of operator determination, traveltime inversion and amplitude determination, as described in Chapter 5 and Chapter 6, are computationally far more expensive, which makes the genetic algorithm unfeasible. Therefor we are forced to adopt locally linearized non-linear inversion, based on Bayes' theorem.

[1] In this description of non-linear optimization it is assumed that a minimum objective function value is sought. But the whole process works for a maximization problem as well.

■ B.6.1 Local linearization

In this section local optimization algorithms (Sorensen, 1980; Scales, 1985) will be discussed. In the previous section the objective function was defined as,

$$F(\mathbf{m}) = \frac{1}{2}\mathbf{e}^T\mathbf{e}, \tag{B.6.86}$$

$$= \sum_{l=1}^{N} e_l(\mathbf{m})e_l(\mathbf{m}), \tag{B.6.87}$$

in which, for the non-linear case, we take,

$$\mathbf{e} = \mathbf{L}\left(\mathbf{d} - \mathbf{g}(\mathbf{m})\right), \tag{B.6.88}$$

with $\mathbf{L}$ some sort of weighting function. If the objective function would be minimized in an iterative fashion, then at some iteration k, the update of the parameter vector will be defined as,

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha\mathbf{p}_k, \tag{B.6.89}$$

in which $\mathbf{p}_k$ is the update direction and α_k the step length[2]. Then, a Taylor expansion of the objective function reads,

$$F_{k+1} = F(\mathbf{m}_k - \alpha_k\mathbf{p}_k), \tag{B.6.90}$$

$$\approx F(\mathbf{m}_k) - \alpha_k\mathbf{g}_k^T(\mathbf{m}_k)\mathbf{p}_k + \alpha_k^2\mathbf{p}_k\mathbf{G}_k(\mathbf{m}_k)\mathbf{p}_k. \tag{B.6.91}$$

At a stationary point, the gradient of the objective function with respect to the model parameters should be zero. Hence, the following should hold,

$$\frac{\partial F(\mathbf{m}_k - \alpha_k\mathbf{p}_k)}{\partial\alpha_k p_{i,k}} = 0, \tag{B.6.92}$$

$$\mathbf{g}_k(\mathbf{m}_k) - \mathbf{G}_k(\mathbf{m}_k)\alpha_k\mathbf{p}_k = 0. \tag{B.6.93}$$

$$\tag{B.6.94}$$

and leads to the parameter update,

$$\alpha_k\mathbf{p}_k = \mathbf{G}_k^{-1}\mathbf{g}_k, \tag{B.6.95}$$

in which $\mathbf{g}_k$ is the gradient, or vector with first derivatives of F with respect to $\mathbf{m}_k$ and $\mathbf{G}_k$, the Hessian, or matrix with second derivatives of F with respect to $\mathbf{m}$. In the special situation of the weighted least squares objective function, the following

[2]In this and following section $\cdot_k$ refers to an iteration number and not to a subvector or subelement

holds,

$$
\begin{aligned}
g_{i,k} &= \frac{\partial F}{\partial m_{i,k}}, && \text{(B.6.96)}\\
&= \sum e_l \frac{\partial e_l}{\partial m_{i,k}}, && \text{(B.6.97)}\\
\mathbf{g}_k &= \mathbf{J}_k^T \mathbf{e}. && \text{(B.6.98)}\\
G_{ij,k} &= \frac{\partial^2 F}{\partial m_{i,k} \partial m_{j,k}}, && \text{(B.6.99)}\\
&= \sum \frac{\partial e_l}{\partial m_{i,k}} \frac{\partial e_l}{\partial m_{j,k}} + e_l \frac{\partial^2 e_l}{\partial m_{i,k} \partial m_{j,k}}, && \text{(B.6.100)}\\
\mathbf{G}_k &= \mathbf{J}_k^T \mathbf{J}_k + \mathcal{H}_e, && \text{(B.6.101)}
\end{aligned}
$$

in which $\mathbf{J}$ is the Jacobian and the term $\mathbf{J}^T\mathbf{J}$ will be called pseudo-Hessian. Hence the iterative procedure reduces to,

$$
\begin{aligned}
\mathbf{m}_{k+1} &= \mathbf{m}_k + \alpha_k \mathbf{p}_k, && \text{(B.6.102)}\\
\mathbf{p}_k &= \left(\mathbf{J}_k^T \mathbf{J}_k + \mathcal{H}_e\right)^{-1} \mathbf{J}_k^T \mathbf{e}_k. && \text{(B.6.103)}
\end{aligned}
$$

When derivative information is not directly available, the Jacobian and Hessian should be computed by finite differencing. Since this can become computationally expensive, often $\mathcal{H}_e$ is neglected when the initial model is not too far from the minimum of the objective function. Otherwise the full Hessian should be approximated in a smart way, which is subject to extensive research going beyond the scope of this thesis.

Note that in case of a linear set of equations,

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}, \qquad \text{(B.6.104)}$$

we could write for the gradient in case of a Gaussian estimator (Equation B.5.47),

$$\mathbf{g}_k = -2\mathbf{G}^T \mathbf{C}_n^{-1}(\mathbf{d} - \mathbf{G}\mathbf{m}) - 2\mathbf{C}_m^{-1}\hat{\mathbf{m}} + 2\mathbf{C}_m^{-1}\mathbf{m}, \qquad \text{(B.6.105)}$$

in which the error for the current iteration is defined as,

$$
\begin{aligned}
F &= \mathbf{e}^T \mathbf{e}, && \text{(B.6.106)}\\
\mathbf{e} &= \begin{bmatrix} \mathbf{C}_n^{-\frac{1}{2}}(\mathbf{d} - \mathbf{G}\mathbf{m}) \\ \mathbf{C}_m^{-\frac{1}{2}}(\hat{\mathbf{m}} - \mathbf{I}\mathbf{m}) \end{bmatrix}. && \text{(B.6.107)}
\end{aligned}
$$

For the Hessian we can write

$$\mathbf{G}_k = 2\mathbf{G}^T \mathbf{C}_n^{-1} \mathbf{G} + 2\mathbf{C}_m^{-1}. \qquad \text{(B.6.108)}$$

This would then lead to the parameter set according to Equation B.6.103,

$$\alpha_k \mathbf{p}_k = \mathbf{m}_k - \left(\mathbf{G}^T \mathbf{C}_n^{-1} \mathbf{G} + \mathbf{C}_m^{-1}\right)^{-1} \left(\mathbf{G}^T \mathbf{C}_n^{-1} \mathbf{d} + \mathbf{C}_m^{-1} \hat{\mathbf{m}}\right). \tag{B.6.109}$$

If this expression for the update is inserted in Equation B.2.46, it is concluded that a linear system would lead to a single iteration solution regardless of the chosen initial parameter vector. In case of an M-estimator, the gradient would read (Equation B.4.37)

$$\mathbf{g}_k = -2\mathbf{G}^T \mathbf{W}_k (\mathbf{d} - \mathbf{G}\mathbf{m}_k) - 2\mathbf{C}_m^{-1} \hat{\mathbf{m}} + 2\mathbf{C}_m^{-1} \mathbf{m}_k. \tag{B.6.110}$$

As the weighting matrix $\mathbf{W}$ depends on the current parameter distribution, the full Hessian is not straightforwardly computed and we would suffice by neglecting $\mathcal{H}$, which would lead to a similar parameter update,

$$\alpha_k \mathbf{p}_k = \mathbf{m}_k - \left(\mathbf{G}^T \mathbf{W}_k \mathbf{G} + \mathbf{C}_m^{-1}\right)^{-1} \left(\mathbf{G}^T \mathbf{W}_k \mathbf{d} + \mathbf{C}_m^{-1} \hat{\mathbf{m}}\right). \tag{B.6.111}$$

However, since W_k depends on the parameter set of the current iteration, this solution would indeed lead to just a parameter update. Recomputing the error would lead to an update of the weighting matrix W and we would again need to compute the updated gradient and Hessian, leading to a new parameter update. Note, though that only W needs to computed and the gradient and Hessian are computed by matrix products of W with a stationary system. What was called iteratively reweigthing within the previous section thus fully obeys to local linearization.

■ B.6.2 Error and resolution

The weight functions describe the uncertainties on the data/noise and the a priori uncertainties on the parameters. The objective of the inversion procedure is to estimate the parameters with a higher certainty than the a priori parameters. Thus an inversion procedure should be complemented with a posterior analysis of the uncertainties of the result.

When the forward equation can be linearized around the true parameter set, after sufficient iterations of the non-linear inversion scheme, we end up with,

$$\mathbf{g}(\mathbf{m}) = \mathbf{g}(\mathbf{m}_\infty) + \mathbf{H}_\infty (\mathbf{m} - \mathbf{m}_\infty), \tag{B.6.112}$$

in which $\mathbf{m}$ is the true (unknown) parameter vector, $\mathbf{m}_\infty$ the parameter vector after the final update and $\mathbf{H}_\infty$ the derivative operator after the final update,

$$H_{ij,\infty} = \frac{\partial g_i(\mathbf{m}_\infty)}{\partial m_{j,\infty}}. \tag{B.6.113}$$

Then the a posteriori pdf of the parameter vector reads

$$F = \text{constant} \times \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_\infty)\mathbf{C}_P^{-1}(\mathbf{m} - \mathbf{m}_\infty)\right), \tag{B.6.114}$$

in which the a posteriori weight operator $\mathbf{C}_P$ can be shown to read,

$$\mathbf{C}_P = \left[\mathbf{H}_\infty^T \mathbf{C}_n^{-1} \mathbf{H}_\infty + \mathbf{C}_m^{-1}\right]^{-1}. \tag{B.6.115}$$

The square roots of the diagonal elements of the posterior covariance operator define the posterior variances and can be read as error bars on the estimated parameters. The off-diagonal elements describe to what extent parameters are correlated. By rewriting the off-diagonal elements to

$$\rho_{ij} = \frac{C_{ij}}{\sqrt{(C_{ii})}\sqrt{(C_{jj})}}, \tag{B.6.116}$$

parameters m_i and m_j are said to be correlated when $|\rho_{ij}|$ is close to one. This means that the parameters are not independently solved, only combinations are. If so, the posterior covariance operator is singular and the inverse depends on on the a priori information given on the parameters (making the inverse non-singular). The full operator over the model-space defines the "ellipsoid of error" of which iso-density lines correspond to certain probabilities.

The inverse of the covariance operator to exist, is closely related to the resolution of the system. The resolution describes a linear relation between the true parameters and the estimated parameters, for which we can write by linearizing around $\hat{\mathbf{m}}$

$$\hat{\mathbf{m}} - \mathbf{m}_p = \mathbf{R}(\mathbf{m} - \mathbf{m}_p), \tag{B.6.117}$$

in which

$$\mathbf{R} = \left[\mathbf{H}_\infty^T \mathbf{C}_n^{-1} \mathbf{H}_\infty + \mathbf{C}_m^{-1}\right]^{-1} \mathbf{H}_\infty^T \mathbf{C}_n^{-1} \mathbf{H}_\infty, \tag{B.6.118}$$

which can be rewritten to

$$\mathbf{R} = \mathbf{I} - \mathbf{C}_P \mathbf{C}_m^{-1}. \tag{B.6.119}$$

When the resolution operator equals to the identity operator, the model is perfectly resolved. Gilbert (1968) state that the resolution operator acts as a filter to the true model resulting in the a posteriori model; we only see a filtered version of the real world. By taking the trace of the equation we can analyze to which extent the result is defined by the data and the a priori information, since,

- TRACE($\mathbf{I}$) = total number of parameters
- TRACE($\mathbf{R}$) = parameters determined by the data
- TRACE($\mathbf{C_p C_m^{-1}}$) = parameters determined by a priori information

■ B.6.3 Scaling

To solve the normal equations of equations, the inverse of $\mathbf{J}^T\mathbf{J}$, whether directly or through iterative schemes such as LSQR or Conjugate Gradients. For the inverse of

$\mathbf{J}^T\mathbf{J}$ to exist, $\mathbf{J}^T\mathbf{J}$ should be non-singular, or the ratio of the largest eigenvalue and the smallest eigenvalue should be close to one. For the system to be non-singular it has to be well defined and well conditioned.

To start, it is important to realize the meaning of an Eigenvalue decomposition of the system-matrix. As already mentioned by the Tarantola (1987), for covariance operators an Eigenvalue analysis is in principle meaningless. The decomposition is in mathematical sense only valid for automorphisms, operators mapping a space in itself. Covariance operators map dual spaces. If for instance the following simple Eigen value problem is considered

$$\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \phi = m\phi, \tag{B.6.120}$$

then the Eigenvalue has either the physical dimension of σ_1^2 or the physical dimension of σ_2^2 which is a discrepancy.

If we do not consider the physical dimensions, the results numerically obtained will not be intrinsic, they depend on the units chosen. According to Equation B.6.103 the parameter update within non-linear optimization is given by

$$\mathbf{H}\Delta\mathbf{x} = \mathbf{J}^T\Delta\mathbf{t}, \tag{B.6.121}$$

in which $\Delta\mathbf{x}$ resembles the parameter update, $\mathbf{H}$ the full Hessian, $\mathbf{J}$ the Jacobian and $\Delta\mathbf{t}$ the measurement error (including the a priori vector). The conditioning of the normal equations depends on the Eigenvalue decomposition of the Hessian $\mathbf{H}$,

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T, \tag{B.6.122}$$

for which $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. Thus, the Eigenvalue decomposition decomposes the operator in its principle axis, of which the Eigenvalues define the length of the principle axes. The posterior probability density function is given by Equation B.6.114, of which the curvature is defined by the inverse of the Hessian, $\mathbf{H}^{-1} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T$. For relative small Eigenvalues, the principle axis of the a posteriori pdf has an inverse proportional length. Practically speaking, this means that along these directions, the parameters are poorly resolved. When a parameter scaling (for instance meters to kilometers) is considered,

$$\Delta\tilde{\mathbf{x}} = \mathbf{D}\Delta\mathbf{x}, \tag{B.6.123}$$

the a posteriori pdf can be rewritten as

$$\exp \Delta\mathbf{x}\mathbf{D}\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T\mathbf{D}\Delta\mathbf{x}, \tag{B.6.124}$$

$$\exp \Delta\mathbf{x}\left[\mathbf{U}\mathbf{S}\mathbf{V}^T\mathbf{D}^\mathbf{T}\mathbf{D}\mathbf{V}\mathbf{S}\mathbf{U}^T\right]^{-1}\Delta\mathbf{x}, \tag{B.6.125}$$

in which $\mathbf{U}\mathbf{S}\mathbf{V}^T$ is the singular value decomposition of the Jacobian. A parameter scaling alters the direction of the principle axes and their lengths. In general numerical ill-conditioning can occur when large order differences exist in the used units for

the different parameters.

To overcome numerical instabilities scaling is necessary. This can be achieved by expressing the parameters in units that correspond to the range of interest, by taking a priori standard deviations as units or by statistically normalizing the parameters by $\mathbf{D} = \mathbf{C}_x^{\frac{1}{2}}$. In this thesis a very practical scaling is introduced, which effectively transforms the main diagonal of the Hessian to unity.

Since the eigenvalues are determined by the units of the parameters, one could define a scaling operator that assures the diagonal of $\mathbf{J}^T\mathbf{J}$ to become close to unity,

$$\mathbf{S} = \begin{bmatrix} \frac{1}{\sum_i^N \frac{\partial t_i}{\partial m_1}} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sum_i^N \frac{\partial t_i}{\partial m_j}} & & \\ & & & \ddots & \\ & & & & \frac{1}{\sum_i^N \frac{\partial t_i}{\partial m_M}} \end{bmatrix}. \tag{B.6.126}$$

The scaling matrix $\mathbf{S}$ is then used to scale the columns of $\mathbf{J}$ as follows,

$$\hat{\mathbf{J}} = \mathbf{JS}. \tag{B.6.127}$$

then elements of $\hat{\mathbf{J}}^T\hat{\mathbf{J}}$ can be written as follows,

$$\left[\hat{\mathbf{J}}^T\hat{\mathbf{J}}\right]_{kl} = \frac{\sum_i^N \frac{\partial t_i}{\partial m_k}\frac{\partial t_i}{\partial m_l}}{\sqrt{\sum_i^N \left(\frac{\partial t_i}{\partial m_k}\right)^2}\sqrt{\sum_i^N \left(\frac{\partial t_i}{\partial m_k}\right)^2}}, \tag{B.6.128}$$

$$= \frac{\mathbf{j}_k^T\mathbf{j}_l}{|\mathbf{j}_k||\mathbf{j}_l|}, \tag{B.6.129}$$

$$= \cos\phi, \qquad 0 \leq \phi \leq \frac{1}{2}\pi, \tag{B.6.130}$$

resulting in diagonal terms equal to one. The scaled system to be solved becomes,

$$\left(\hat{\mathbf{J}}^T\hat{\mathbf{J}}\right)\Delta\hat{\mathbf{m}} = \hat{\mathbf{J}}^T\mathbf{y}, \tag{B.6.131}$$

$$\left([\mathbf{JS}]^T[\mathbf{JS}]\right)\Delta\hat{\mathbf{m}} = [\mathbf{JS}]^T\mathbf{y}, \tag{B.6.132}$$

$$\left(\mathbf{S}^T\mathbf{J}^T\mathbf{JS}\right)\Delta\hat{\mathbf{m}} = \mathbf{S}^T\mathbf{J}^T\mathbf{y}, \tag{B.6.133}$$

$$\left(\mathbf{J}^T\mathbf{JS}\right)\Delta\hat{\mathbf{m}} = \mathbf{J}^T\mathbf{y}, \tag{B.6.134}$$

$$\mathbf{S}\Delta\hat{\mathbf{m}} = \Delta\mathbf{m}. \tag{B.6.135}$$

Secondly, it is important that under determined parameters are damped by stabilization of the operator. The damping of the system is defined by including a priori

information, such as defining parameter variances, or Tikhonov regularization. The damping should be consistent with the units of the parameters. To properly account for the parameter scaling, a damping factor is defined for each set of parameters with identical units. We state that a certain amount of parameters is under determined. When a parameter is well determined, it means that its equivalent column in the Jacobian has relatively large values compared to an under determined parameter, i.e. it has a significant contribution to the objective function. Therefore, for each set of parameters, the mean and standard deviation of the inverse scaling vector **S** are computed. Then the parameter sets will be stabilized by the value found at $mean - \alpha \times standarddeviation$.

■ B.6.4 Simulated evolution

Genetic algorithms or simulated evolution methods are a subset of so called Monte Carlo methods (Gallagher et al., 1991; Robert and Casella, 2005). Monte Carlo methods are based on random search techniques through parameter space. Within these random search techniques it is possible to optimize the search strategy and thus obtaining more efficient algorithms. Simulated evolution is one of these more sophisticated search methods and is based on evolutionary processes as observed in nature and was used by (Wilson et al., 1994) to determine residual statics. Another more sophisticated method is the Heat Bath Algorithm, which is based on cooling down processes and recrystallization of certain elements contained within the heated fluid, also referred to as simulated annealing (Cerny, 1985).

In this thesis the simulated evolution is adopted. The method is based on the evolution of populations. Within each iteration (in terms of evolution of populations, a generation of the population), during the population selection phase, population members are selected according to a certain fitness, to form a pair of parents. Such a pair of parents will produce a pair of children during the crossover phase (in terms of evolution, genes of both parents are mixed). Finally, at random some children or parents have a change to mutate in the mutation phase. This last phase keeps some randomness in the population pool and prevents the algorithm to converge prematurely. Next the different steps will be discussed individually an in more detail.

Start off

At the start of the algorithm a population will be created containing a fixed number N of members. Members of the population are characterized by a chromosome. This chromosome contains a set of genes. Each gene is a representation of the parameters we are looking for. For instance, in residual statics a chromosome contains a number of genes equal to the number of independent static components; number of source statics, receiver statics, residual move-out components, etc. Thus one specific gene equals a specific residual static component. Initially all the chromosomes are created randomly within the expected parameter range.

Population generation

In the next step a new population is generated from the initial population. For each member the fitness is calculated. The fitness is based on the function we would like to optimize (the objective function). For instance, we could choose to optimize the stack power, then the fitness of each member is equal to the stack power its genes will create. Several methods exist to create the new population set,

- Fitness proportionate selection
  From the fitnesses, $F_n(\mathbf{x}_n)$, a probability distribution is created,

$$P_n = \frac{F_n(\mathbf{x}_n)}{\sum_{i=1}^{N} F_i(\mathbf{x}_i)}, \tag{B.6.136}$$

  in which F_n is the fitness of member n characterized by the chromosome $\mathbf{x}_n$. Next, N random numbers are drawn from a uniform distribution. These numbers are used to select N members for the new population by using the created probability distribution. In this way members with a higher fitness are in more favor to be selected and duplicated than members with a low fit. As this is a random process, members with a low fitness still make a change to pass. Note that in this process some members will be selected more than once and some members will, thus, not be selected at all.

- Rank selection
  Instead of creating a distribution from the fitnesses, the fitnesses are used to rank the members of the population. The member with the highest fitness is ranked number one, the lowest is ranked number N. From these rankings again a probability distribution can be created,

$$P_n = \frac{f(R_n)}{\sum_{i=1}^{N} f(R_i)}, \tag{B.6.137}$$

  in which $f(R_n)$ is a function depending on the ranking R_n of member n. Creating a new population is again performed by drawing random numbers and using the distribution of Equation B.6.136. Rank selection depends on the functional $f(\cdot)$ and has the effect that the difference between high ranked members is exaggerated and the difference between the highest ranked and lowest ranked member is made less stronger. This means that relatively low ranked members all make an almost equal and larger chance of being selected.

- Tournament selection
  From the current population N pairs will be selected. The two members forming one pair are randomly drawn from the population. Each of the two members $\mathbf{x}_1$ and $\mathbf{x}_2$ is assigned a probability,

$$P_i = \frac{F_i(\mathbf{x}_i)}{F_1(\mathbf{x}_1) + F_2(\mathbf{x}_2)}, i \in \{1, 2\}. \tag{B.6.138}$$

By drawing a random number between 1 and 0, one of the two members is selected for the new population, according to its probability. This procedure has the effect that low ranked members have a smaller chance of being selected for the next generation.

Crossover

From the newly generated population, randomly $N/2$ pairs of parents are selected. Each pair will create two children through crossover. Again a random number is drawn to decide at which point the two parent chromosomes will be cut in two, thus creating two heads and to tails. Child 1 will be created from tail of parent 1 and head of parent 2 and child 2 will be created from the tail of parent 2 and the head of parent 1. Instead of single crossover, also multiple crossovers can be made. Then for each gene position randomly a 0 or 1 will be drawn. At 1, the genes of both parents will be interchanged, at 0 genes will stay intact. Both methods are graphically explained in Figure B.3. After crossover, the fitnesses of both children are computed. Some algorithms (Wilson et al., 1994) select the two members with the highest fitness, within the set of the two parents and the two children, to remain within the population, other algorithms replace the parents by their children (McCormack et al., 1999).

Mutation

Within the last step, all members are subject to mutation with a certain probability. If a member is selected for mutation (generally a chance between 1 and 5 percent), a few random genes are mutated with a value randomly taken from the set of feasible values within the defined parameter space. Again, algorithms do differ to some extend within literature. In Wilson et al. (1994) for instance, offspring is either created by crossover or by mutation. After crossover the two strongest members are passed through to the next generation out of the set of 4 (2 parents and 2 children), while when its decided to mutate the two parents, crossover will not take place and the mutated parents are passed through to the next generation.

After the mutation phase, the next generation is passed through to the population generation phase. For genetic algorithms, still no sound mathematic fundamentals exists which proves its convergence. Therefore, a lot of slightly differing algorithms exist based on intuitive improvements. One could for instance skip the phase of population generation and select parents in the crossover phase directly by using one of the selection criteria of the original population generation phase.

A more rigorous alternative was proposed by Coyne and Paton (1994). Instead of considering one population, several subpopulations are considered. Each population exploits part of model-space through the iteration scheme as described above. Every now and then randomly members of subpopulations are exchanged. Each group has

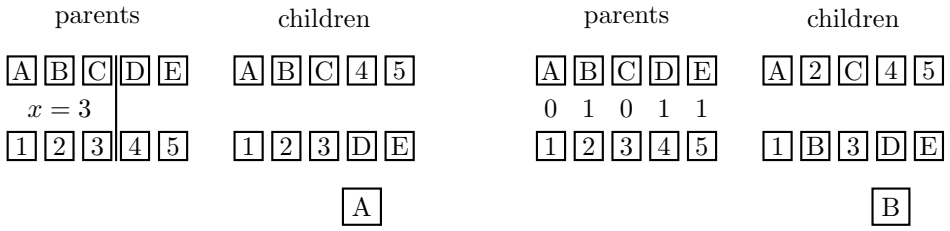

Figure B.3: *Children can be generated in two ways. 1) By choosing randomly a single crossover location. 2) Through multi crossover for each chromosome it is decided randomly whether it should crossover or not.*

a tendency toward a different (local) optimum and the algorithm is prevented from stalling due to the random interaction.

■ B.6.5 Residual statics example

To determine residual statics, Ronen and Claerbout (1985) proposed to optimize the stack power. Instead of finding optimized traveltime picks within consecutive CMP gathers, residual traveltime shifts are defined as those shifts that give the highest power within the stack, which they proved by considering the Cauchy Schwarz inequality. Although residual surface consistent statics are posed as a linear problem, the actual implementation by defining the optimal shift through picking of maxima within crosscorrelation panels turns the problem into a non-linear one; especially when one has to deal with large statics, cycle skips are easily introduced within the crosscorrelations, leading to local optima. To overcome the influence of cycle skips, Rothman (1986) used a simulating annealing algorithm to determine the optimal static shifts. Wilson et al. (1994) and Stork and Clayton (1991) further enhanced the use of non-linear global search algorithms to determine optimal residual statics. The described genetic algorithm of the previous section is used to demonstrate the search of residual statics, as implemented by Wilson et al. (1994). For this purpose the following objective function is optimized,

$$F(\mathbf{s},\mathbf{r},\mathbf{g}) = \sum_h \sum_t \left[\sum_i \hat{P}(x_i^S, x_{h-i}^R, t + s_i + r_{h-i} + g_h)\right] \left[\sum_i \hat{P}(x_i^S, x_{h+1-i}^R, t + s_i + r_{h+1-i})\right], \tag{B.6.139}$$

in which $\mathbf{s}$ is a vector containing the individual source statics s_i for surface location i, $\mathbf{r}$ is a vector containing the individual receiver statics r_j for surface location j and $\mathbf{g}$ is a vector containing the structural time shifts g_h for CMP location $h =$

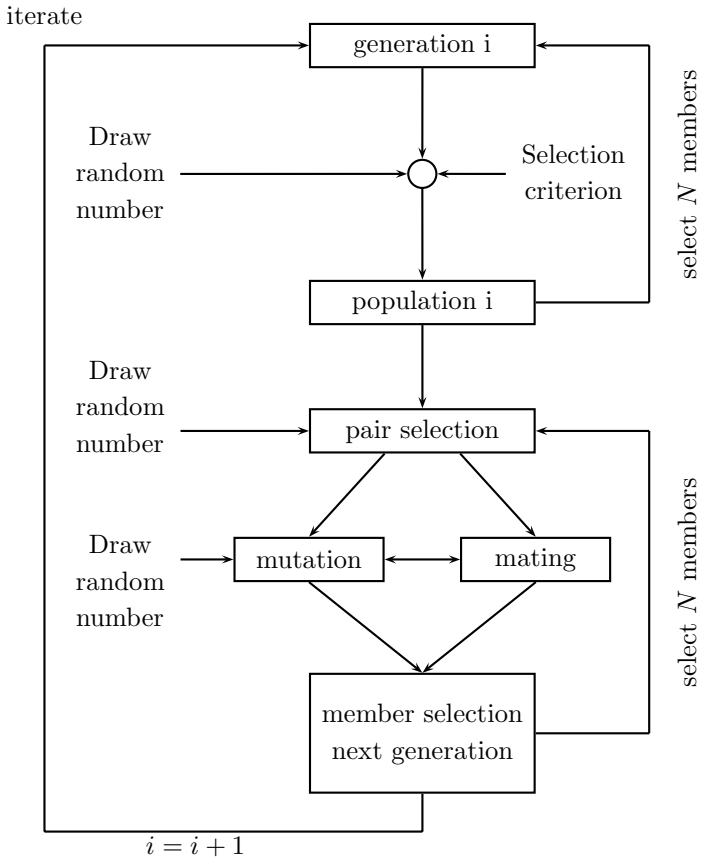

Figure B.4: *Flow chart of the genetic algorithm.*

$i + j$. The structural parameters $\mathbf{g}$ define time lags that optimize the coherence between two CMP gathers. To demonstrate the genetic algorithm, data was modeled with 20% noise, over a set a horizontal reflectors for which the optimal normal move-out was applied according to the known synthetic model. Surface consistent random source and receiver statics were added to the data. Since the reflectors are horizontally layered, the structural parameters can be set to zero and the objective can be rewritten in terms of taking the autocorrelation power of the CMP gathers instead of the crosscorrelation power,

$$F(\mathbf{s},\mathbf{r}) = \sum_h \sum_t \left[\sum_i \hat{P}(x_i^S, x_{h-i}^R, t + s_i + r_{h-i}) \right]^2 , \tag{B.6.140}$$

which reduces the problem into optimizing the stack power. In the example, 50 shot-experiments were used, for which the signal was recorded in the same set of 50 receivers. For each shot and receiver an individual static shift is present. Therefore, the total number of static parameters equals to 100. Within the optimization scheme a total of 200 members was used to evolve into a set of best fitting static time-shifts. Figure B.5 shows the stack results after different amounts of iterations during the optimization scheme. Although hardly any signal can be recognized, after termination the resulting stack images the reflectors perfectly. The performance of the scheme is displayed in Figure B.6, together with the original source and receiver statics and the final inversion result. Note, that a bulk shift within the source or receiver static does not have any influence on the resulting stack power. As was stated in Chapter 2, static corrections must be kept to a minimum and therefore, within the optimization scheme, after each iteration, the resulting parameter families are reduced by their average.

B.7 Sparse operators

When inverting measurements for a desired set of parameters, a linear forward system evaluation needs to be performed a large number of times. When using the LSQR algorithm, a forward system evaluation is carried out during the bidiagonalization phase, for each iteration as given by Equation B.2.70 and Equation B.2.70. In case of non-linear optimization, the linear operator is replaced by the Jacobian, as given by Equation B.3.101 when the non-linear problem is linearized. Otherwise, the computation of fitness requires a forward evaluation as outlined for instance by the statics example of Section B.4. Quite often the forward operator can become very large. For instance, a linear forward operation, involving 500 shot-locations with each each 500 receiver-locations, which are related to 1000 independent variables results in a system having $500 \times 500 \times 1000 = 250000000$ entries and at least the same number of floating point operations for each forward evaluations. When the number of iterations becomes large and when for instance multiple frequencies should be involved, the number of operations can even become at least an order of magnitude higher.

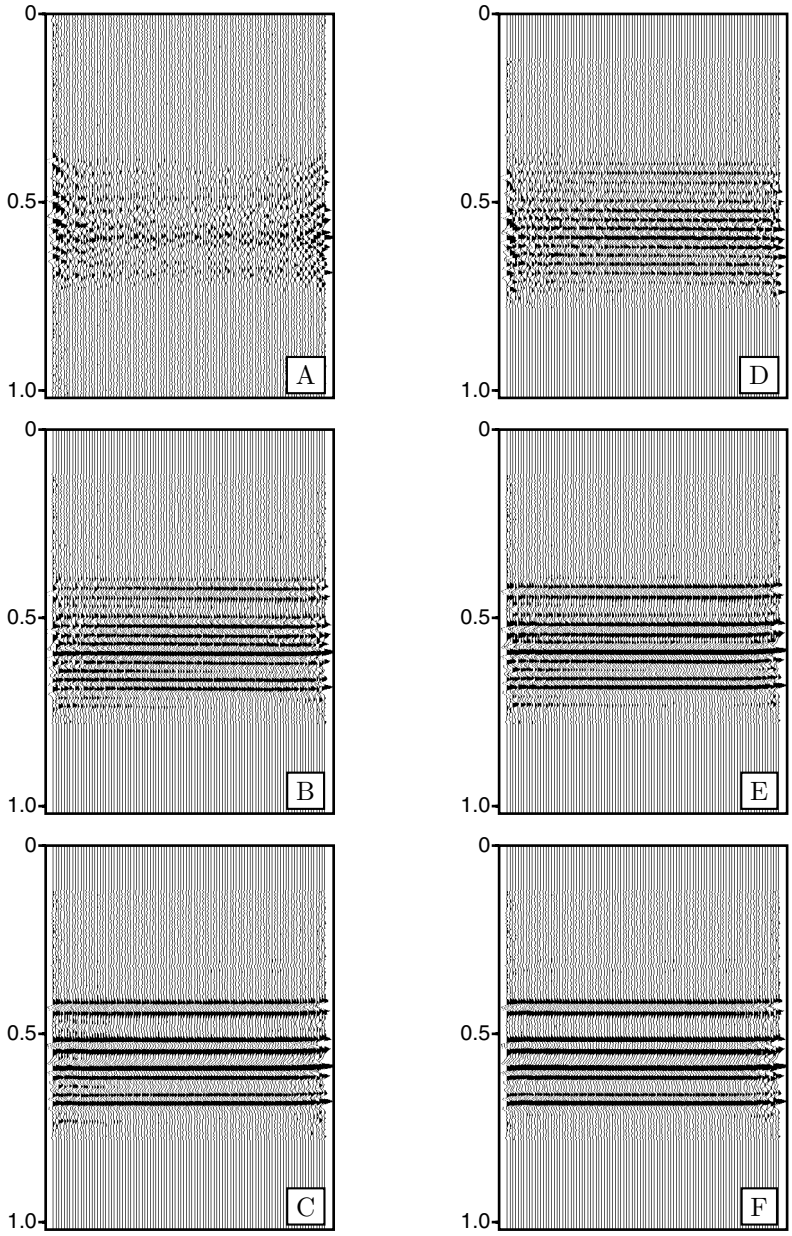

Figure B.5: *Optimization of the stack power through determination of source and receiver statics by a genetic algorithm. A) The original stack. B) through F) The result after respectively 100, 300, 500, 700 and 1000 iterations.*

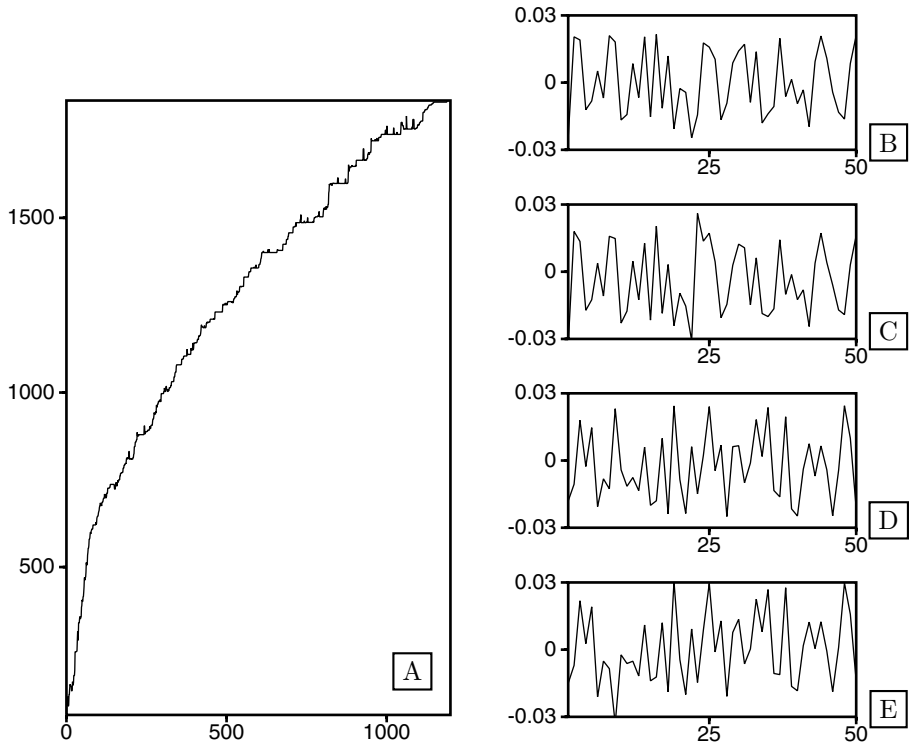

Figure B.6: *Performance of the genetic algorithm on static estimation through stack power optimization. A) shows the stack power as function of the number of iterations. B) and D) show the true static shifts of respectively the sources and receivers. C) and E) show the estimated static shifts.*

However, quite often for subsets of the measurements only a subset of the parameters is involved and the system becomes sparse, meaning a large number of the entries of the forward operator equals to zero and associated floating point operators do not need to be evaluated.

By storing the non-zero elements of the forward operator efficiently, a large advantage is created when evaluating the forward operations. Figure B.7. displays schematically the storage scheme of a sparse operator, of which the numerical implementation can be found in Press et al. (1992). The sparse operator is represented by two column-vectors, the first one containing all diagonal elements and nonzero elements and the second one containing all indices to the diagonal and nonzero elements. The first N elements of the data vector contain the diagonal elements of the

$M \times N$, operator. The next $M - N$ elements equal zero. The equivalent elements within the index-vector point to the first location of the first non-zero element for each row in the data-vector. As such it is easy to keep track of all elements and to evaluate each operation that resembles,

$$\mathbf{f} = \mathbf{A}\mathbf{b}, \tag{B.7.141}$$

in which $\mathbf{A}$ resembles the $M \times N$ operator, $\mathbf{b}$ the $N \times 1$ parameter-vector and $\mathbf{f}$ the $M \times 1$ data-vector.

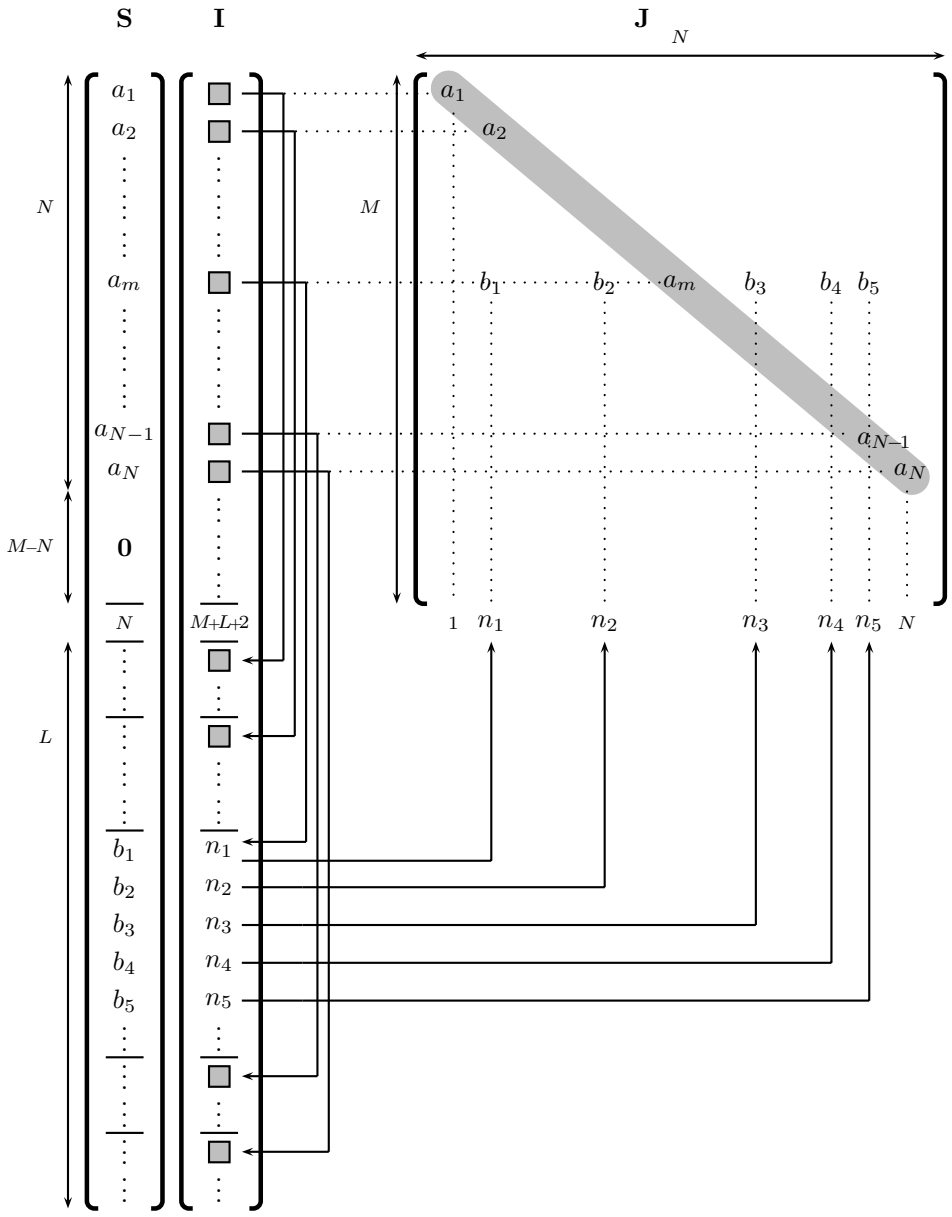

Figure B.7: *Representation of sparse operator* $\mathbf{J}$ *by means of an index vector* $\mathbf{I}$ *and a vector* $\mathbf{S}$ *containing all non-zero elements of* $\mathbf{J}$

Bibliography

Abramowitz, M., and Stegun, I. A., 1970, Handbook of mathematical function; with formulas, graphs, and mathematical tables: Dover Publications, Inc.

Aki, K., and Richards, P., 1980, Quantitative seismology, theory and methods, part 2: W.H. Freeman, San Francisco.

Al-Ali, M. N., 2007, Land seismic data acquisition and preprocessing, an operator solution to the near surface problem.: Ph.D. thesis, Delft University of Technology.

Al-Chalabi, M., 1997, Parameter nonuniqueness in velocity versus depth functions: Geophysics, **62**, no. 3, 970–979.

Bahorich, M., Coruh, C., Robinson, E., and Costain, J., 1982, Static corrections on the southeastern piedmont of the united states: Geophysics, **47**, 1540–1549.

Bakulin, A., and Calvert, R., 2004, Virtual source: new method for imagaing and 4d below complex overburden:, *in* Expanded Abstracts Soc. Expl. Geophys., 2477–2480.

Bard, Y., 1974, Nonlinear parameter estimation: Academic Press.

Berkhout, A., Ongkiehong, O., Volker, A., and Blacquiere, G., 2001, Comprehensive assessment of seismic acquisition geometries by focal beams, part i: Theoretical considerations: Geophysics, **66**, no. 3, 911–917.

Berkhout, A. J., 1982, Seismic migration, imaging of acoustic energy by wave field extrapolation, a: theoretical aspects: Elsevier.

Berkhout, A. J., 1997a, Pushing the limits of seismic imaging, part i: Prestack migration in terms of double dynamic focusing: Geophysics, **62**, no. 3, 937–953.

——— 1997b, Pushing the limits of seismic imaging, part ii: Integration of prestack migration, velocity estimation and avo analysis: Geophysics, **62**, no. 3, 954–969.

Berlioux, A., and Harlan, W., 1997, Iterative least-squares inversion for amplitude balancing: Stanford Exploration Report, **89**, no. www.sep.stanford.edu, 169–181.

Bleistein, N., and Handelsman, R., 1986, Assymptotic expansion of integrals: Dover Publications.

Blonk, B., and Herman, G., 1994, Inverse scattering of surface waves: A new look at surface consistency: Geophysics, **59**, no. 6, 963–972.

Bolte, J. F. B., Verschuur, D. J., and Hegge, R. F., 1999, CFP operator estimation and inversion demonstrated on a field dataset - part i: operator updating: 69th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1711–1714.

Bolte, J. F. B., 1997, Estimation of focusing operators using the common focal point method: Ph.D. thesis, Delft University of Technology.

Booker, A. H., Linville, A. F., and Wason, C., 1976, Long wavelength static estimation: Geophysics, **41**, no. 5, 939–959.

Bube, K. P., and Langan, R. T., 1997, Hybrid $l1/l2$ minimization with applications to tomography: Geophysics, **62**, no. 4, 1183–1195.

Cambois, G., and Stoffa, P., 1992, Surface-consistent deconvolution in the log/fourier domain: Geophysics, **57**, no. 6, 823–840.

Cambois, G., and Stoffa, P., 1993, Surface-consistent phase decomposition in the log/fourier domain: Geophysics, **58**, no. 8, 1099–1111.

Campman, X., Riyanti, D., and Herman, G., 2001, Statics and near-field scattering.: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 2029–2032.

Campman, X. H., 2005, Imaging and suppressing near-receiver scattered seismic waves data: Ph.D. thesis, Delft University of Technology.

Cerny, V., 1985, Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm: J. Opt. Theory Appl., **45**, 41–51.

Cerveny, V., 1985, The application of ray tracing to the propagation of shear waves in complex media., *in* Dohr, G., Ed., Seismic shear waves, part A: Theory: Geophysical Press, Handbook of Geophysical Exploration, 15A, 1–119.

Chemingui, N., Clapp, R., and Claerbout, J., 2001, 2d phase unwrapping: Stanford Exploration Report, **84**.

Cox, M., 1999, Static corrections for seismic reflection surveys: Society of Exploration Geophysicists.

Cox, B. E., 2004, Tomographic inversion of focusing operators: Ph.D. thesis, Delft University of Technology.

Coyne, J., and Paton, R., 1994, Genetic algorithms and directed adaptation, *in* Fogarty, T. C., Ed., Lecture Notes in Computer Science no. 865: Springer Verlag, 103–114.

de Hoop, M. V., 1992, Directional decomposition of transient acoustic wave fields: Ph.D. thesis, Delft University of Technology.

Denisov, M., Finikov, D., and Oberemchenko, D., 2001, A compact non-stationary wavelet parametrization for deconvolution and q estimation: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Extended Abstracts.

Dijkstra, E. W., 1959, A note on two problems in connection with graphs: Numerische Mathematik, **1**, 269–271.

Dix, C. H., 1955, seismic velocities from surface measurements: geophysics, **20**, no. 1, 68–68.

Docherty, P., 1992, Solving for the thickness and velocity of the weathering layer using 2-d refraction tomography: Geophysics, **57**, no. 10, 1307–1318.

Dubose, J., 1993, Practical steps toward realizing the potential of monte carlo automatic statics: Geophysics, **58**, no. 4, 399–407.

Duijndam, A. J., 1988a, Bayesian estimation in seismic inversion. part i: Principles.: Geophysical Prospecting, **36**, 878–898.

——— 1988b, Bayesian estimation in seismic inversion. part ii: Uncertainty analysis.: Geophysical Prospecting, **36**, 899–918.

Edge, A., and Laby, T., 1931, *in* The principles and practice of geophysical prospecting. Cambridge University Press, 339–341.

Epili, D., Criss, J., and Cunningham, D., 2001, Turning-ray tomography for statics solutions: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1910–1913.

Erdelyi, A., 1956, Asymptotic expansions: Dover Publications, Inc.

Ernst, F., 1999, Modelling, imaging and removal of guided waves: Ph.D. thesis, Delft University of Technology.

Fokkema, J. T., and van den Berg, P. M., 1993, Seismic applications of acoustic reciprocity: Elsevier Science Publ. Co., Inc.

Gallagher, K., Sambridge, M., and Drijkoningen, G., 1991, Genetic algorithms: an evolution from monte carlo methods for strongly non-linear geophysical optimization problems: Geophys. res. lett., **18**, 2177–2180.

Gilbert, G. B. F., 1968, The resolving power of gross earth data: Geophys. J. R. astron. Soc., **16**, 169–205.

Golub, G. H., and van Loan, C. F., 1996, Matrix computations: The John Hopkins University Press.

Gonzalez, A., Osypov, K., and Lau, A., 1999, Unconstrained near-surface velocity inversion of linear moveout velocities and its impact on structural mapping:, *in* 69th Ann. Internat. Mtg Soc. of Expl. Geophys., 520–523.

Guo, J., and Zhou, X., 2001, Surface consistent phase correction: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Extended Abstracts.

Hagedoorn, J., 1959, The plus-minus method of interpreting seismic refraction sections: Geophysical Prospecting, **7**, 158–182.

Hampson, D., and Russell, B., 1984, First-break interpretation using generalized linear inversion: Can. J. Expl. Geophys., **20**, no. 01, 40–54.

Harlan, W., 1990, Tomographic estimation of shear velocities from shallow crosswell seismic data: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 86–89.

Hatherly, P., Urosevic, M., Lambourne, A., and Evans, B., 1994, A simple approach to calculating refraction statics corrections: Geophysics, **59**, no. 1, 156–160.

Hegge, R., 2001, Seismic macromodel estimation by inversion of focusing operators: Ph.D. thesis, Delft University of Technology.

Hill, N. R., 1987, Downward continuation of refracted arrivals to determine shallow structure: Geophysics, **52**, no. 09, 1188–1198.

Hindriks, K., and Duijndam, A. J., 2000, Reconstruction of 3-d seismic signals irregularly sampled along two spatial coordinates: Geophysics, **65**, no. 1.

Huber, P., 1964, Robust estimation of a location parameter: Ann. Math. Statist., **35**, 73–101.

Huber, P., 2003, Robust statistics: Wiley, New York.

Kabir, M. M. N., and Verschuur, D. J., 1995, Restoration of missing offsets by parabolic radon transform: Geophys. Prosp., **43**, 347–368.

Kabir, M. N., and Verschuur, D., 2000, A constrained parametric inversion for velocity analysis based on cfp technology: Geophysics, **65**, 1210–1222.

Karcher, J. C., 1987, The reflection seismograph: its invention and use in the discovery of oil and gas fields: Geophysics, **42**, 10–19.

Kästner, U., and Buske, S., 1999, Computing geometrical spreading from traveltimes in 3-d: 69th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1739–1742.

Kirchheimer, F., and Ferber, R., 2001, Robust surface-consistent deconvolution with noise suppression: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts.

Larner, K., Gibson, B., Chambers, R., and Wiggins, R., 1979, Simultaneous estimation of residual statics and crossdip time corrections: Geophysics, **44**, no. 5, 1175–1192.

Lavielle, M., 1991, 2d bayesian deconvolution: Geophysics, **56**, no. 12, 2008–2018.

Lines, L., 1993, Ambiguity in analysis of velocity and depth: Geophysics, **58**, no. 4, 596–597.

Macrides, C., Kanasewich, E., and Bharatha, S., 1988, Multiborehole seismic imaging in steam injection heavy oil recovery projects: Geophysics, **53**, no. 1, 65–75.

Marsden, D., 1993a, Static corrections-a review, part i: The Leading Edge, **89**, no. 1, 43–48.

——— 1993b, Static corrections-a review, part ii: The Leading Edge, **89**, no. 2, 115–120.

——— 1993c, Static corrections-a review, part iii: The Leading Edge, **89**, no. 3, 210–216.

Mayne, W., 1956, Seismic surveying: United States Patent 2 732 906; Geophysics (abstract), **21**, 1856–1867.

McCormack, M. D., Stoisits, R. F., Macallister, D. J., and Crawford, K. D., 1999, Applications of genetic algorithms in exploration and production: The Leading Edge, **18**, no. 6, 716–718.

Meissner, R., 1961, Wave-front diagrams from uphole shooting: Geophysical Prospecting, **9**, 533–543.

Moser, J., 1991, Shortest path calculation of seismic rays: Geophysics, **56**, no. 1, 59–67.

Nolet, G., 1987, Seismic tomography: D. Reidel Publishing Company.

Osypov, K., 2000, Robust refraction tomography:, *in* 70th Ann. Internat. Mtg Soc. of Expl. Geophys., 2032–2035.

Osypov, K., 2001, Refraction tomography - a practical overview of emerging technologies: Recorder, **26**, no. 2, 5–8.

Paige, C. C., and Saunders, M. A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software, **8**, no. 1, 43–71.

Parry, D., and Lawton, D. C., 1994, Near-surface velocity structure from borehole and refraction seismic surveys: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 602–604.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1992, Numerical recipes in c: The art of scientific computing: Cambridge University Press.

Profeta, M., Moscoso, J., and Koremblit, M., 1995, Minimum field static corrections: The Leading Edge, **89**, no. 6, 684–687.

Rajasekaran, S., and McMechan, G., 1995, Prestack processing of land data with complex topography: Geophysics, **60**, no. 6, 1875–1886.

Rayleigh, J.W. Strutt, Baron, 1877, The theory of sound: Macmillan and co.

Reshef, M., 1991, Depth migration from irregular surfaces with depth extrapolation methods: Geophysics, **56**, no. 1, 119–122.

Rey, W. J., 1983, Introduction to robust and quasi-robust statistical methods: Springer Verlag, Berlin.

Robert, C. P., and Casella, G., 2005, Monte carlo statistical methods: Springer.

Rogers, A., 1981, Determination of static corrections, *in* Fitch, A., Ed., Developments in geophysical exploration methods, No.2: Applied Science Publ. Ltd., 1–36.

Ronen, J., and Claerbout, J., 1985, Surface-consistent residual statics estimation by stack-power maximization: Geophysics, **50**, no. 12, 2759–2767.

Rothman, D., 1986, Automatic estimation of large residual statics corrections: Geophysics, **51**, no. 2, 332–346.

Saleh, S. J., 1994, Sensitivity of refraction statics to choice of the picking range of offsets: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1525–1528.

Scales, L. E., 1985, Introduction to non-linear optimization: Macmillan.

Schneider, W. A., 1971, Developments in seismic processing and analysis (1968-1970): Geophysics, **36**, no. 6, 1043–1073.

Schoolmeesters, J. W. A. V., 2001, Three-dimensional processing of marine seismic data by spectral decomposition: Ph.D. thesis, Delft University of Technology.

Sen, M., and Stoffa, P., 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, **56**, 1624–1638.

Sethian, J., and Popovici, A., 1999, 3-d traveltime computation using the fast marching method: Geophysics, **64**, no. 2, 516–523.

Shannon, C., 1948, A mathematical theory of communication: Bell System Tech. J., **27**, 379–423.

Sherriff, R., 1991, Encyclopedic dictionary of exploration geophysics: Society of Exploration Geophysicists, Tulsa.

Sorensen, H. W., 1980, Parameter estimation, principles and problems: Marcel Dekker, New York.

Spagnolini, U., 1993, 2d phase unwrapping and phase abasing: Geophysics, **58**, no. 9, 1324–1334.

Stewart, R., 1991, Exploration seismic tomography: Fundamentals: Society of Exploration Geophysicists, Tulsa.

Stork, C., and Clayton, R., 1991, Linear aspects of tomographic velocity analysis: Geophysics DEZE IS NIET GOED, **56**, 483–495.

Strang, G., 1988, Linear algebra and its applications: Harcourt Brace Jovanovich, publishers.

Taner, M. T., and Berkhout, A. J., 1997, Statics, a physical view: 59th EAGE Conference and Technical Exhibition, EAGE, Extended Abstracts.

Taner, M. T., and Berkhout, A. J., 1998, Dynamics of statics: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1409–1412.

Taner, M. T., Koehler, F., and Alhilali, K., 1974, Estimation and correction of near-surface time anomalies: Geophysics, **39**, no. 4, 441–463.

Taner, M. T., Lu, L., and Baysal, E., 1988, Unified method for 2-d and 3-d refraction statics with first break picking by supervised learning: 58th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Session:S5.1.

Tarantola, A., 1987, Inverse problem theory: Elsevier.

Tegtmeier-Last, S., 2007, Redatuming of sparse 3d seismic data: Ph.D. thesis, Delft University of Technology.

Thorbecke, J., 199j, Common focus point technology: Ph.D. thesis, Delft University of Technology.

Tygel, M., Schleicher, J., and Hubral, P., 1992, Geometrical spreading corrections of offset reflections in a laterally inhomogeneous earth: Geophysics, **57**, 1054–1063.

van Dedem, E. J., and Verschuur, D. J., 2005, 3d surface-related multiple removal prediction: A sparse inversion approach: Geophysics, **70**, V31–V43.

van der Rijzen, M., Gisolf, A., and Verschuur, D., 2004, Infilling of sparse 3d data for 3d focusing operator estimation: Geophysical Prospecting, **52**, no. 6, 489–507.

van Wijngaarden, A. J., 1998, Imaging and characterization of angle-dependent seismic reflection data: Ph.D. thesis, Delft University of Technology.

Vanelle, C., and Gajewski, D., 2000, Traveltime-based true-amplitude migration: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 961–964.

Vasudeva, K., Wilson, W., and Laidlaw, W., 1991, Simulated annealing statics computation using an order-based energy function: Geophysics, **56**, no. 11, 1831–1839.

Verschuur, D., and Mahrfoul, B., 2005, Estimation of complex near surface focusing operators by global optimization: 67th Mtg. Eur. Assoc. Expl Geophys., EAGE, Houten, Extended Abstracts.

Vidale, J., 1988, Finite difference calculation of traveltimes: Bull. Seism. Soc. Am., **18**, 2062–2076.

Volker, A. W. F., 2002, Assessment of 3d seismic acquisition geometries by focal beam analysis: Ph.D. thesis, Delft University of Technology.

Wapenaar, C. P. A., and Berkhout, A. J., 1989, Elastic wave field extrapolation: redatuming of single- and mul ti-component seismic data: Elsevier Science Publ. Co., Inc.

Wapenaar, C. P. A., and Grimbergen, J. L. T., 1996, Reciprocity theorems for one-way wavefields: Geoph. J. Int., **127**, 169–177.

Weatherby, B. B., 1971, The history and development of seismic prospecting: Geophysics, **5**, 216–228.

Wilson, W., Laidlaw, W., and Vasudevan, K., 1994, Residual statics estimation using the genetic algorithm: Geophysics, **59**, no. 5, 766–774.

Winthaegen, P., and Verschuur, D., 2003, Local, tau-p-based interval velocity analysis applied at target level:, *in* 73rd Annual International Meeting Soc. of Expl. Geophys., 2096–2099.

Yilmaz, O., 1987, Seismic data processing: Society of Exploration Geophysicists.

Zhu, X., Sixta, D., and Angstman, B., 1992, Tomostatics: Turning-ray tomography and static corrections: The leading edge, **11**, no. 12, 15–23.

Summary

Estimation and removal of complex near surface effects in seismic measurements

Especially on land, near surface inhomogeneities tend to obscure the final quality of a seismic image. Almost all processes that are needed to obtain this image are affected by these near surface inhomogeneities. The commonly used procedures to account for these effects are known under the name 'static corrections', or in short 'statics'. Their aim is to "Determine the reflection arrival times which would have been observed if all measurements had been made on a (usually) flat datum with no weathering or low velocity material present." However, near surface anomalies do not only manifest themselves in time shifts but also in other effects, such as focusing and defocussing, amplitude anomalies etc. Therefore, in this thesis a method is developed to determine the recorded reflection energy which would have been observed if all measurements had been made on a datum without the influence of the weathering material present.

The commonly used family of 'statics' methods correct the data by applying corrections which solely depend on the surface locations of the sources and receivers, irrespective of offset and recording time. A detailed comparison shows that applying static corrections comprises further imaging by affecting migration velocities and introducing false structure, whereas full wavefield redatuming does not influence the migration velocities. However, to be able to use multi-trace wavefield redatuming, knowledge is required on the propagation characteristics of the near surface.

The determination of the propagation characteristics through the near surface can be accomplished through the Common Focus Point technology, which can be directly derived from the coupled wave equations. The two methodologies that arise from the Common Focus Point concept, operator updating based on the principle of equal traveltime and Fermat modeling, together form a recipe to derive operators containing these near surface characteristics. Using this recipe, operators are found which can be used to redatum the data. As the process is fully data-driven, resolution is not restricted by a constraining velocity-depth model. The recipe, however, does not guarantee uniqueness if each single focuspoint operator is updated independently from the other operators. Furthermore, no exact information is available

on the locations of the redatumed sources and receivers.

Moreover, strong near surface effects tend to degrade the operator determination procedure in such a way that the operators can not properly be recovered if the near surface effects are not been accounted for in some way. The source of this problem lies in the destructive interference caused by the fast variations in traveltime during the construction of the common focus point gathers. By using the so-called concept of wave-field healing in combination with Fermat modeling, we are able to constrain the operators with a sparse set of parameters that tend to include the near surface effects adequately. These parameters can be obtained through a global non-linear optimization process. Static time shifts can be used to initialize the parameters, for which two alternative procedures are given. Additional constraints can be included by measuring at downhole receivers physical near surface propagators in the field.

By using the determined operators as subject to a velocity inversion process, additional information is obtained about the redatuming locations. Using a raytracing algorithm as forward model within the optimization procedure turns out to be unstable. Therefore, grid-based methods are preferred. A drawback of using grid-based methods is the need of a large amount of additional computations to derive derivative information. Due to the large amount of computations involved, a locally linearized non-linear optimization scheme is employed. To diminish the influence of picking errors during the operator determination, use is made of a robust M-class estimator instead of a Gaussian estimator.

Besides traveltime information also amplitude information is needed to be able to properly redatum the data. This information can be obtained by parameterizing the amplitudes and optimizing the parameters by minimizing the difference between the measured amplitudes and the forward modeled amplitudes, using Fermat modeling. Again, use is made of a locally linearized non-linear optimization scheme, based on a robust M-class estimator. The amplitudes, found in a data-driven method, are then used in a least-squares redatuming process. If, however, the amplitudes are not optimally determined, one should refrain from a least-squares redatuming as it can severely degrade the final result.

The proposed method shows very good results for two synthetic datasets, generated in realistic subsurface models, containing both complex near surface inhomogeneities as well as topography. Optimal resolution is obtained as the data-driven operator estimation and near surface model estimation are implemented as decoupled processes. The proposed methodology turns out to be complimentary to methods that estimate and subtract multiplicative near surface scattering energy. A combination of the method outlined in this thesis and the methods that estimate near surface scattering energy is well capable of solving both the short wavelength as well as the longer wavelength near surface inhomogeneities.

As near-surface inhomogeneities pose a 3D problem, the proposed methodologies should be extended from a 2D implementation to a 3D implementation. Sparsely measured physical redatuming operators in the field, by placing downhole receivers below the near surface, could be used to guide the operator updating procedure, constrain the near surface model estimation process and constrain the amplitude determination and thus further enhance the proposed methodology. Although the presented methodology renders good results on specially designed synthetic data, they should be applied to real data in future research in combination with a proper statistical assessment of the results.

Kees Hindriks

Samenvatting

Afschatten en verwijderen van effecten in seismische metingen, geintroduceerd door complexe, ondiepe inhomogeniteiten

Voornamelijk op land, neigen inhomogeniteiten vlak onder het aardoppervlak het uiteindelijke seismische beeld te vertroebelen. Bijna alle processen die nodig zijn om dit uiteindelijke beeld te verkrijgen worden door deze inhomogeniteiten beinvloed. De gebruikelijke methoden om deze effecten te behandelen zijn bekend onder de naam 'static corrections', in het kort 'statics' genoemd, waarvan het doel is gedifinieerd als "Bepalen van de aankomsttijd van reflecties die geobserveerd zouden zijn als alle metingen gedaan zouden zijn aan een (gewoonlijk) vlak datum zonder de aanwezigheid van verweerd materiaal of materiaal met lage snelheid." Echter, ondiepe anomaliën manifesteren zich niet alleen in aankomsttijd maar ook in bijvoorbeeld focusering en defocusering, amplitude anomaliën, etcetera. Om die reden is in deze thesis een methode ontwikkeld welke de reflectie energie bepaalt die gemeten zou zijn als er geen invloed was geweest van verweerd materiaal.

De gebruikelijke toepassing van de familie van 'statics' methoden corrigeren de data door correcties toe te passen die slechts afhankelijk zijn van de locaties waar de bronnen en ontvangers zich bevinden, ongeacht de afstand tussen de bronnen en ontvangers alswel de gemeten aankomsttijden. Een gedetailleerde vergelijking laat zien dat het gebruik van statische correcties nadelig is voor verdere beeldvorming doordat het de beeldvormings snelheden beinvloedt en het nietbestaande structuren introduceert, terwijl het gebruik van volledige golfveld herdatering de beeldvormings snelheden niet beinvloed. Echter, om multi-trace golfveld herdatering toe te kunnen passen voor het corrigeren van effecten veroorzaakt door inhomogeniteiten in de ondiepe ondergrond is kennis vereist van de propagatie karakteristieken van de ondiepe ondergrond.

Het bepalen van de propagatie karakteristieken door de ondiepe ondergrond kan in principe bereikt worden met behulp van de gemeenschappelijke brandpuntanalyse die direct uit de gekoppelde golfvergelijkingen kan worden herleid. De twee methodologien die volgen uit het concept van de gemeenschappelijke brandpuntanalyse,operator verbetering gebaseerd op het principe van gelijke reistijd en Fer-

mat modelering, vormen gezamelijk een recept om operatoren te bepalen die de karakteristieken van propagatie door de ondiepe ondergrond in zich dragen. Deze operatoren kunnen dan gebruikt worden om de gemeten data te herdateren naar een locatie onder de ondiepe ondergrond, zodaning dat de inhomogeniteiten van de ondiepe ondergrond geen invloed meer hebben op de data. Aangezien dit proces volledig gebaseerd is op de gemeten data, wordt het oplossend vermogen niet aan banden gelegd door een beperkend snelheid-diepte model. Het recept garandeert echter niet een unieke oplossing wanneer elke operator horend bij één brandpunt onafhankelijk van alle andere opertoren wordt afgeschat. Ook is er geen informatie voorhanden over de exacte locatie van de geherdateerde bronnen en ontvangers.

Verder is het zo dat de effecten van de ondiepe ondergrond de procedure voor het bepalen van de operatoren neigen te degraderen op zo een wijze, dat de operatoren niet op een juiste manier herleid kunnen worden als deze effecten niet op enige manier worden meegenomen in de procedure. De kern van dit probleem bevindt zich in de destructieve interferentie die veroorzaakt wordt door snelle variaties in de reistijden tijdens het construeren van de gemeenschappelijk brandpunt verzamelingen. Door gebruik te maken van het concept van golffront-heling in combinatie met het principe van data constructie kunnen we de operatoren beschrijven met een schaarse verzameling aan parameters die de ondiepe ondergrond verschijnselen adequaat genoeg beschrijven. Deze parameters kunnen vervolgens verkregen worden met behulp van een globaal niet-lineair optimaliseringsproces. Extra beperkingen kunnen meegenomen worden door werkelijke propagatie verschijnselen in het veld te meten met behulp van ontvangers die onderin een boorgat geplaatst worden.

Door de uiteindelijk bepaalde opertoren te gebruiken in een inversie procedure om de golfsnelheden te bepalen, kan extra informatie verkregen worden over de daadwerkelijke herdatering locaties. Het blijkt dat het gebruiken van een algoritme dat golfstralen volgt als voorwaarts model binnen het optimaliserings proces leidt tot instabiliteiten. Om deze reden krijgen grid gebaseerde algoritmen de voorkeur. Een nadeel van het gebruiken van grid gebaseerde algoritmen is de noodzaak van een zeer groot aantal extra berekeningen om informatie over afgeleiden te verkrijgen. Omdat een groot aantal berekingen nodig is, is er gebruik gemaakt van een lokaal gelineariseerd niet-lineair optimalisatie schema. Om de invloed van pick-fouten tijdens het bepalen van de operatoren te verminderen is gebruik gemaakt van de robuuste klasse van M-schatters in plaats van een Gaussische schatter.

Naast reistijd informatie is ook informatie over de ampplitudes noodzakelijk om het mogelijk te maken de data op een juiste manier te herdateren. Deze informatie kan bepaald worden door de amplitudes te parameterizeren en vervolgens deze parameters te optimalizeren door het verschil tussen de gemeten amplitudes en de voorwaarts gemodelleerde amplitudes te minimalizeren met gebruikmaking van Fermat modelering. Wederom is gebruik gemaakt van een lokaal gelineariseerd niet-lineair optimalisatie schema gebaseerd op een robuuste klasse M-schatter. De

op een data-gedreven manier gevonden amplitudes worden vervolgens gebruikt in een kleinste kwadraten herdaterings proces. Echter, als de amplitudes niet optimaal gedefinieerd zijn dan dient men zich te weerhouden van het gebruiken van een kleinste kwadraten herdatering, gezien dit ernstig het uiteindelijke resultaat kan schaden.

De voorgestelde methode laat goede resultaten zien voor twee synthetische dataverzamelingen, die gemodelleerd zijn in realistische ondergrondse modellen welke beiden zowel complexe inhomogeniteiten als topologie bevatten. Een optimaal oplossend vermogen is bereikt doordat de datagedreven operatorbepaling en de modelschatting zijn geïmplementeerd als twee losgekoppelde processen. De voorgestelde methode blijkt complementair te zijn aan methoden die multiplicatieve energie, verstrooid door de ondiepe ondergrond, afschatten en aftrekken van de data. Een combinatie van de methode zoals uiteengezet in deze these en de methoden die door de ondiepe ondergrond verstrooide energy afschatten is zeer goed in staat om zowel te corrigeren voor ondiepe ondergrondse inhomogeniteiten met een korte golflengte als voor inhomogeniteiten met een langere golflengte.

Omdat inhomogeniteiten in de ondiepe ondergrond problemen veroorzaken in drie dimensies, zullen de voorgestelde methoden uitgebreid moeten worden van een tweedimensionale implementatie naar een drie-dimensionale implementatie. Hier en daar in het veld gemeten herdaterings-operatoren, door ontvangers te plaatsen in een boorgat onder de ondiepe ondergrond, kunnen gebruikt worden om het verbeteringsproces, dat gebruikt wordt om de operatoren te bepalen, te leiden, het schatten van het ondergrondse model te beperken en het schatten van de amplitudes te beperken. Op deze manier kunnen de veldgemeten operatoren de voorgestelde methodologie verder verbeteren. Hoewel de voorgestelde methodologie goede resultaten produceert op speciaal ontworpen synthetische data dient ze, binnen toekomstig onderzoek, toegepast te worden op werkelijk gemeten velddata, vergezeld van de juiste statische analyse van de resultaten.

Kees Hindriks

Dankwoord

Het schrijven van deze these is geen gemakkelijke taak geweest, en zeker de laatste periode waarin het schrijven niet langer meer een dagtaak kon zijn heeft vele opofferingen gevraagd niet alleen van mezelf maar ook van de mensen om mij heen. Ik wil hier een ieder danken die mij door de jaren gesteund heeft en zonder wie dit proefschrift niet tot een eind had kunnen komen. Hoewel de lijst met mensen die ik dank lang is, gebiedt de ruimte me te beperken tot het noemen van slechts enkelen bij naam.

Allereerst wil ik Eric Verschuur danken die behoorlijk wat geduld heeft moeten opbrengen met het continu verhappen van weer de zoveelste draft versie. Je pragmatische en inventieve manier van werken heeft op vele punten een grote bijdrage aan dit proefschrift geleverd.

Next, I would like to thank my promotor, Prof. Dr. ir. A.J. Berkhout for providing the concepts and hypotheses that forms the base of the presented thesis and the opportunity he has given me to conduct the research within the DELPHI consortium. The semi-yearly sponsor meetings have been inspiring events and an excellent opportunity for a Ph.D. student to discuss with and present his work to industry representatives. Therefore, I would also like to thank the members and sponsors of the consortium for their continuous support and fruitful input to the presented research. Especially, I would like to thank Dr. Panos Kelamis for his close involvement and giving me the opportunity to present and implement my work at Saudi-Aramco in Dahran.

Prof. Dr. ir. A. Gisolf wil ik bedanken voor het blijven tonen van interesse en blijven vrijhouden van een plek voor mij in de vakgroep tijdens de jaren dat ik niet meer werkzaam was aan de universiteit.

Dan wil ik mijn kamer- en ganggenoten in Delft bedanken, Paul, Barbara, John, Ewoud en Karin, voor al de wetenschappelijke discussies, niet wetenschappelijke beslommeringen en gezamelijke ergernissen. Arno, Frederiek, Remco en Gerrit wil

ik bedanken voor het wekelijkse rustpunt in de kelder, Edo, Henri en Paul K. voor hulp en inzet bij computerproblemen. Een speciaal bedankje ook voor onze vaste gast Xander, voor de inzichtelijke discussies op het Plein over de ondiepe ondergrond.

Mijn collega's, en met name Peter S., Flip, Cor, Christian, Najwa, John, Frans, Arjan en Sietse, toch bedankt voor het continue doorvragen naar het proefschrift maar ook voor het steunen in zoeken naar tijd, dat dan uiteindelijk toch geleid heeft tot dit boek.

Luc en Joris, alle beschouwingen en levensvragen tijdens de aandewindse rakken waren welkome momenten om de gedachten eens goed te verzetten. Met een zorg minder wordt de kans wel minder groot dat we ooit nog eens een aanmoedigingsprijs winnen.

Frank, Hessel, Roger, Marijn, Maarten, Valentine en Lysianne, ik ben er nog steeds niet uit of al het goede wining en dining het schrijfproces nu versneld of vertraagd heeft maar in ieder geval heeft het op vele momenten de noodzakelijke rust gegeven. Folkert, ik ben je weer een klein stapje voor.

Kees en Toos, bedankt voor alle jaren van steun en vertrouwen.

About the author

Cornelius Okko Henri (Kees) Hindriks was born in Haarlem, The Netherlands, on the fifth of April, 1970. From 1982 to 1988 he attended secondary school (VWO) at the "Zeldenrust College" in Terneuzen and "Ichthus College" in Driehuis, both in The Netherlands.

In September 1989, he started his studies in Mining and Petroleum Engineering at the Technical University of Delft in the currently known department of Applied Earth Sciences. In 1996 he graduated in the field of applied geophysics. His M.Sc. thesis on reconstruction of 3D irregularly sampled Wavefields was carried out under supervision of Prof. dr. ir. J.T. Fokkema and Dr. ir. A.J.W. Duijndam.

In 1997 he conducted further research on imaging of irregularly sampled VSP data at the section of applied Geophysics in cooperation with Norske Hydro S.A. In 1998 he started his Ph.D. research on removing of near surface effects through redatuming by focusing operators under the supervision of his promotor Prof. dr. ir. A.J. Berkhout and toegevoegd promotor dr. ir. D.J. Verschuur within the department of Applied Physics at Delft University of Technology. The results of this research are reflected within this thesis.

Since August 1, 2002, he is working for Shell International Exploration and Production B.V.