Tomographic inversion of focusing operators
Tomographic inversion of focusing operators

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 dinsdag 2 maart 2004 om 15:30 uur

door

Barbara Ellen COX
doctoranda in de geofysica geboren te Melick en Herkenbosch
Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. A.J. Berkhout
Prof. dr. ir. A. Gisolf

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

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. dr. ir. A.J. Berkhout, Technische Universiteit Delft, promotor
Prof. dr. ir. A. Gisolf, Technische Universiteit Delft, promotor
Dr. ir. D.J. Verschuur, Technische Universiteit Delft,
toegevoegd promotor
Prof. dr. W. Spakman, Universiteit Utrecht
Prof. dr. ir. P.M. van den Berg, Technische Universiteit Delft
Dr. A.P.E. ten Kroode, Shell International Exploration and Production
Dr. ir. P.M. van der Made, Fugro-Jason

ISBN 90-9017792-2

Copyright ©2004, by B.E.Cox, Laboratory of Aoustical Imaging and Sound Control, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author B.E. Cox, Faculty of Applied Sciences, Delft University of Technology, P.O. Box 5046, 2600 GA, Delft, The Netherlands.

SUPPORT
The research for this thesis was financially supported by the DELPHI consortium.

Typesetting system: LATEX.
Printed in The Netherlands by Fedoruk BV, Enschede.
to Roald
Contents

1 General introduction ........................................... 1
   1.1 A brief introduction to seismic exploration ................ 1
   1.1.1 The importance of seismic velocity models ............... 1
   1.1.2 The concept of seismic tomography ...................... 3
   1.1.3 The use of focusing operators in tomography ............ 5
   1.2 Statement of the problem ................................ 9
   1.3 Objectives of this research ................................ 9
   1.4 Outline of thesis ......................................... 10

2 Focusing operators ............................................ 13
   2.1 Introduction ................................................ 13
   2.2 The WRW model as a base for the CFP method ............. 14
   2.3 The CFP method as a base for focusing operators .......... 17
       2.3.1 Focusing of the receivers ......................... 17
       2.3.2 Focusing of the sources ............................ 19
       2.3.3 Applying both focusing steps ..................... 19
   2.4 Focusing operator updating ................................ 21
       2.4.1 Data-driven updating of focusing operators ........ 21
       2.4.2 Updating of focusing operators using sparse data ... 22
   2.5 Conclusions .............................................. 27

3 Tomographic inversion methodology .......................... 29
   3.1 Introduction .............................................. 29
   3.2 Fundamentals of tomographic inversion .................... 30
       3.2.1 The non-linearity of tomographic inversion ........ 30
       3.2.2 The procedure of tomographic inversion ............ 30
   3.3 Model parameterization .................................... 32
3.3.1 Overview of existing methods ........................................... 32
3.3.2 Parameterization used in tomographic inversion of focusing
operators ................................................................. 35
3.4 Forward modeling of traveltimes .......................................... 38
3.4.1 Overview of existing methods ........................................ 38
3.4.2 Analytical ray-tracing in Delaunay triangulation .............. 42
3.4.3 Determination of the optimization matrix ...................... 49
3.5 Optimization ...................................................................... 53
3.5.1 Overview of existing methods ........................................ 53
3.5.2 Optimization by the LSQR algorithm .......................... 55
3.6 Regularization of the inversion ............................................. 60
3.6.1 Overview of existing methods ........................................ 60
3.6.2 Regularization by resolution dependent parameterization .... 62
3.6.3 Calculation of resolution and covariance by LSQR ...... 64
3.7 Conclusions on the tomographic inversion strategy .......... 72
3.7.1 Tomographic inversion of focusing operators ................. 72
3.7.2 Data-driven procedure .............................................. 73

4 Additional modules to the tomographic inversion of focusing oper-
ators .................................................................................. 75
4.1 Introduction ................................................................. 75
4.2 Inversion of focusing operators using a priori information .... 76
4.2.1 Transferring a priori information to the model parameteriza-
tion by geostatistics ....................................................... 76
4.3 Joint inversion of P and S-wave focusing operators ........... 80
4.3.1 The WRW model for P and S-wave data ...................... 80
4.3.2 The CFP method for P and S-wave data ...................... 82
4.3.3 Focusing operator updating for P and S-wave data ......... 83
4.3.4 The aperture of the P and S-wave focusing operators ...... 83
4.3.5 Joint tomographic inversion of P and S-wave focusing operators 84
4.4 Evaluation of the inversion result using focal point clouds .... 87
4.4.1 Analyzing the focal point clouds to resolve the velocity-depth
ambiguity during updating ............................................... 89
4.4.2 Evaluating the focal point clouds to indicate anisotropy ..... 93
4.4.3 Evaluating the focal point clouds to discover 3D effects .... 96
4.4.4 Using the focal point clouds to influence the inversion .... 97
4.5 Conclusions ............................................................... 100

5 Evaluation of the methodology with synthetic data examples ... 101
5.1 Introduction ............................................................... 101
5.2 The 2D Turbidite model .................................................. 102
5.2.1 Description of the model .......................................... 102
5.2.2 Input data: the 2D focusing operators ....................... 102
5.2.3 Initial velocity model ............................................. 103
5.2.4 Optimization of the model parameters ................................ 104
5.2.5 Modification of the parameterization .................................. 106
5.2.6 Result ............................................................................... 109
5.2.7 Post-updating step .............................................................. 111
5.2.8 Including a priori information ............................................. 114
5.2.9 Joint inversion of P and S-wave focusing operators .......... 116
5.3 The 2D Marmousi model .......................................................... 120
5.3.1 Description of the model ..................................................... 120
5.3.2 Initial velocity model .......................................................... 121
5.3.3 Optimization of the model parameters ............................... 121
5.3.4 Modification of the parameterization ................................. 122
5.3.5 Result ............................................................................... 124
5.4 The 3D Anticline model ............................................................ 127
5.4.1 Description of the model ..................................................... 127
5.4.2 Input data; the 3D focusing operators ................................. 127
5.4.3 Initial velocity model .......................................................... 129
5.4.4 Optimization of the model parameters ............................... 130
5.4.5 Modification of the parameterization ................................. 131
5.4.6 Results ............................................................................... 133
5.4.7 Focal point clouds ............................................................. 136
5.5 The 3D SEG/EAGE Salt model ................................................ 137
5.5.1 Description of the model ..................................................... 137
5.5.2 Input data .......................................................................... 138
5.5.3 Initial velocity model .......................................................... 138
5.5.4 Optimization of the model parameters ............................... 138
5.5.5 Modification of the parameterization ................................. 140
5.5.6 Results ............................................................................... 141
5.5.7 Focal point clouds ............................................................. 144
5.5.8 2D versus 3D tomographic inversion ................................. 145
5.6 Conclusions ........................................................................... 147
5.6.1 Tomographic inversion of focusing operators ....................... 147
5.6.2 Data-driven procedure ........................................................ 147
5.6.3 Exploring the possibilities of the method: The additional mod-
ules ......................................................................................... 148
5.6.4 Additional remarks ............................................................ 148

6 Evaluation of the methodology with real data examples ........ 149
6.1 Introduction ............................................................................ 149
6.2 The 2D North Sea model ........................................................ 150
6.2.1 Description of the 2D field data set ..................................... 150
6.2.2 Input data: the 2D focusing operators ................................. 150
6.2.3 Initial velocity model .......................................................... 151
6.2.4 Optimization of the model parameters ............................... 152
6.2.5 Modification of the parameterization ................................. 152
6.2.6 Result of the P-wave focusing operator inversion .......... 155
6.2.7 Joint inversion of P and S-wave focusing operators .......... 159
6.3 Conclusions .................................................... 163
6.3.1 Tomographic inversion of focusing operators ............... 163
6.3.2 Data-driven procedure ...................................... 163
6.3.3 Exploring the possibilities of the method: The additional modules .................................................. 164

7 Conclusions and Recommendations 165
7.1 Tomographic inversion of focusing operators .................... 166
7.1.1 Conclusions .................................................... 166
7.1.2 Recommendations ........................................... 167
7.2 Data-driven procedure ........................................... 168
7.2.1 Conclusions .................................................... 168
7.2.2 Recommendation ............................................. 170
7.3 Exploring the advantages of the method: The additional modules .................................................. 170
7.3.1 Conclusions .................................................... 170
7.3.2 Recommendations ........................................... 171

A Parameterization by Delaunay triangulation 173
A.1 Delaunay triangles and tetrahedra ............................ 173
A.2 Delaunay triangulation .......................................... 173
A.2.1 Different methods ............................................ 173
A.2.2 Delaunay triangulation used in the tomographic inversion .. 175
A.2.3 Problems occurring in 3D Delaunay triangulation ........... 176
A.3 Initial Delaunay triangulation ................................... 177
A.4 Refining Delaunay triangulation ................................ 179

B Forward modeling by the ray-based method 181
B.1 Eikonal equation ................................................... 181
B.2 The ray-tracing system .......................................... 184

C The derivation of the partial derivatives with respect to quadratic slowness 189

D Notation, Symbols, and Abbreviations 193
D.1 Notation .......................................................... 193
D.2 Symbols .......................................................... 193
D.3 Abbreviations .................................................... 196

Bibliography 198
Summary 207
Samenvatting 211
CONTENTS

Curriculum vitae 215
Dankwoord 217
1

General introduction

1.1 A brief introduction to seismic exploration

Seismic exploration is the most important geophysical method in the upstream oil and gas industry. The popularity is due to the combination of high accuracy and large penetration. Moreover, with respect to borehole data, seismic measurements provide a large amount of lateral information. The widespread use of the seismic method is in exploration for new hydrocarbon reservoirs and for monitoring the production of these reservoirs. Seismic exploration is an offspring of earthquake seismology [Nolet, 1985]. When an earthquake occurs, the earth is fractured and the rocks on opposite sides of the fracture move relative to one another. Such a fault generates seismic waves that travel outward from the fracture and are recorded at various sites at the surface. These recorded data are used to get knowledge of the earth's interior, through which the earthquake waves traveled (Figure 1.1a). The seismic exploration method involves basically the same type of measurements as earthquake seismology. In seismic exploration an man-made source (mostly positioned at the surface) transmits a wavefield, which propagates into the subsurface, where it is reflected by inhomogeneities in the subsurface. The reflected wavefield is recorded at the surface and can be processed in order to obtain a reflection image of the geological structures in the subsurface (see Figure 1.1b). The goal of the geophysicist is to derive from this image an accurate geological model.

1.1.1 The importance of seismic velocity models

The quality of a seismic image is highly dependent on the accuracy of the acoustic propagation velocity model of the subsurface (Figure 1.2). Traditional imaging con-
Figure 1.1: a) Earthquake (star) generated by a fault. Seismic waves are recorded at various sites at the surface. b) Seismic measurement, in which an artificial source (star) transmits a wavefield, that is reflected by structures within the subsurface. The reflected wavefield is recorded by receivers (triangles) positioned at the surface. c) One-way seismic rays from a focus point (star) that transmits a wavefield in the subsurface, which is recorded at receivers (triangles) positioned the surface.

General introduction

...
1.1 A brief introduction to seismic exploration

an (initial) velocity model (see e.g. Al-Yahya [1989]; MacKay and Abma [1992]). A principally different approach is the use of seismic traveltimes. The traveltimes of the reflection events observed in the seismic data can be inverted to obtain a velocity model. This is performed by traveltime tomography (see for instance Bishop et al. [1985]).

![Seismic velocity model](image)

**Figure 1.2:** a) Seismic velocity model. b) Shot records from shots located at the surface, 1 km, 3 km, and 4 km away from the origin. c) The corresponding seismic image obtained by migration using an erroneous velocity model, and using d) the correct velocity model. The example shown here is the 2D Marmousi model [Versteeg, 1993], which will be discussed in more detail in section 5.3.

### 1.1.2 The concept of seismic tomography

Seismic tomography is defined as an imaging technique that transforms seismic waves, that probe the subsurface, to a cross-sectional velocity model (a tomogram) of the subsurface. Different forms of tomography exist: wave equation tomography [Tarantola, 1984a,b; Woodward, 1992] and traveltime tomography [Bishop et al., 1985]. In wave-equation tomography the complete seismic wavefield is taken into account, while in traveltime tomography only the arrival times of the seismic waves
are used. Those traveltimes are extracted from the reflection events, and inverted to obtain a velocity model of the subsurface. In the last decade several travelt ime tomography methods have been developed:

- **Migration Velocity Analysis (MVA) tomography**: The MVA tomography does not immediately use the pre-stack traveltimes, but iteratively combines velocity estimation and pre-stack depth migration. The migration is performed by an initial velocity model, and generates Common Image Gathers (CIG’s). In these gathers events that are not aligned indicate erroneous migration velocities. The picked events reveal depth errors that are converted into travelt ime errors. These time errors are tomographically inverted to obtain updated velocities and reflector depths (see e.g. Kosloff et al., [1996]).

- **Reflection tomography**: In this method two-way pre-stack traveltimes are used. An initial model describing the velocities and reflector depths is iteratively adjusted, in order to be consistent with the pre-stack traveltimes. Each iteration is driven by the difference between the observed two-way traveltimes and the traveltimes calculated in the model [Bishop et al., 1985].

- **Stereo-tomography**: This method is related to reflection tomography, as it also uses two-way pre-stack traveltimes and adjusts a velocity-depth model. However, in stead of using reflector depths directly, the model is parameterized using ray segment pairs. Such ray segment pairs are defined by a common point in depth, two incident angles, and two one-way traveltimes. In this way, in the estimated model the upward and downward going traveltimes are separated. As a consequence, a large number of parameters have to be estimated. As the two-way traveltimes are not sufficient to constrain this ray segment parameterization, the information of the local slope of the event at the surface is also used [Billette and Lambaré, 1998].

- **Cross-well tomography**: As the name already implies, this method does not use seismic reflection data, but data obtained by well-to-well measurements. As a consequence, the data are one-way traveltimes. Again, an initial model is adjusted by minimizing the difference between the modeled and observed one-way traveltimes. As only one-way traveltimes are used, no reflector depths have to be estimated, but only the velocities in between the wells [Michelen a and Harris, 1991].

- **Earthquake tomography**: Although the data of this method differ from the other methods (see Figure 1.1), the principles of tomography stay the same. The one-way traveltimes of recorded earthquakes are used to adjust an initial model describing the velocities of the earths interior and earthquake locations. This is performed by minimizing the difference between the traveltimes calculated in the initial model, and the traveltimes of the recorded earthquakes [Spakman, 1991; Iyer and Hirahara, 1993; Bijwaard, 1999].
Even though the MVA tomography is the only method that does not directly estimate a velocity model from the traveltimes, but requires an imaging step in the procedure, it is the most common method in the upstream oil and gas industry.

The purpose of all the methods described above is to translate the traveltimes of seismic data to a velocity model of the subsurface. A very important aspect of all these methods is the way the velocity model is represented. There are two main approaches to describe the velocity model: 1) The geology based description using layers and boundaries (global basis functions), and 2) the general description using a dense spatial grid (local basis functions). The advantage of the first method is that it reduces the number of parameters that describe the velocity model, but the unfavorable aspect is that it may introduce a serious bias in the solution. In the second method, the description using a spatial grid, this bias is avoided, but the number of parameters may increase significantly. In this thesis a solution is proposed that makes use of an irregular spatial grid, in which the number of parameters is determined by the data itself, the data being the one-way traveltimes of the focusing operators.

### 1.1.3 The use of focusing operators in tomography

In all the above methods, the velocity model is directly, or indirectly determined from the seismic data. The Common Focal Point (CFP) method [Berkhout, 1997a; Thorbecke, 1997] makes it possible to analyze and treat seismic data in a fundamentally different way. Starting with conventional seismic data, the CFP method is able to determine focusing operators by using 'the principle of equal traveltimes' [Morton and Thorbecke, 1996; Berkhout, 1997b; Thorbecke, 1997; Bolte, 2003]. These focusing operators can be seen as a response from a secondary source (called a focal point) in the subsurface towards receivers at the surface, and therefore they can also be considered as one-way Green’s functions of the subsurface (see Figure 1.1c). By these operators the two-way seismic data can be focused at the focal point in the subsurface. This is performed by applying the focusing operator to both the downward and the upward going wavefield. By doing this for all points in the subsurface, the data can be focused without using a velocity model [Berkhout, 1997a]. Although the focusing operators can focus the data, the position of the focused data in the subsurface is yet unknown. The positioning of the events in the subsurface requires a velocity model. In the CFP method, this velocity model is estimated directly from the focusing operators. Actually, by using the CFP method and the focusing operators, the focusing problem and positioning problem are separated (two-step approach). Or, in other words, first two-way reflection data are transformed into one-way data by estimating the focusing operators using the 'the principle of equal traveltimes' (step 1). Next, the focusing operators are used to estimate the velocity by using tomographic inversion (step 2).

An important difference between conventional seismic data and focusing operator
data is that the the seismic data obviously are in two-way traveltimes, and the focusing operators represent one-way traveltimes. The advantage of the one-way operators is that the presence of multi-valued arrivals will be reduced, as complicated reflections from structural inhomogeneities do not occur in the one-way traveltime domain (compare Figures 1.1b and c). Therefore, these less-complicated focusing operators are very suitable for the velocity model estimation, as it makes the inverse problem fundamentally simpler than the two-way inverse problem.

In the table 1.1 the focusing operator tomography is compared to the five conventional tomography methods as presented in section 1.1.2. The focusing operator tomography has something in common with all these methods, but is unique in its own way.

When it is compared to MVA tomography, the similarity can be found in the fact that they both operate in a transformed domain. However, the focusing operators are related to the single focused domain. In addition, no imaging step is involved. As a consequence, the focusing operator tomography is performed directly in the time domain. On the other hand, MVA tomography deals with the depth migrated domain, meaning the double focused domain together with the imaging step. Hence, errors in depth should be converted to errors in time by means of a velocity model. Therefore, we may conclude that the relation between observed traveltime differences and velocity updates are more straightforward in the CFP method.

The reflection tomography bears its similarity with the focusing operator tomography in the goal that should be achieved: minimizing the traveltime difference between observed and modeled data. However, the focusing operators represent one-way traveltimes while the reflected data obviously are in two-way traveltimes. This also results in a different parameterization, because for reflection tomography, reflectors must be defined in the model.

The same is valid for the stereo-tomography, that also uses two-way traveltimes. Like the CFP method, this method also tries to obtain one-way traveltimes. However, this is achieved by defining ray segment pairs, which is part of the parameterization of the model, while by using the focusing operators the one-way traveltimes are already in the data. This is also the reason that in stereo-tomography extra data, in the form of the slopes of the events, is required.

Between focusing operator tomography and cross-well tomography, more similarities appear. The cross-well data also contains one-way traveltimes and the traveltime difference between the observed and modeled data is minimized. However, in cross-well tomography the locations of the sources, which are situated in the well, are known. Therefore, only the velocities should be updated, while in focusing operator tomography the locations of the 'sources', the focal points, are unknown and should be inverted for as well. A drawback of the cross-well tomography method is that all the sources are located at one level only, while in focusing operator tomography the 'sources' can be located anywhere within the subsurface.
Finally, when focusing operator tomography is compared to earthquake tomography, the largest similarity emerges. Although the used data is different in the sense that earthquake recordings have a sparser and more irregular sampling, they are similar in the sense that both focusing operators and earthquakes recordings contain one-way travel times (Figure 1.1a,c) and the locations of the sources, the earthquakes and the focal points, are unknown.

It can be concluded that although the focusing operators are obtained from conventional two-way seismic traveltime data, they have the advantages of cross-well and earthquake data when they are used for the tomographic inversion. The reason for this is that they represent one-way travel times. Note, however, that there is an overwhelming amount of focusing operators compared to earthquake data, and that the 'sources' of the focusing operators are better distributed compared to the sources in the bore-hole of the cross-well data. In other words, focusing operator tomography can be seen as cross-well tomography with many wells, or earthquake tomography with many earthquakes.

Obviously, it is more correct to compare focusing operator tomography with tomography using two-way seismic data. The fundamental difference between, for example, reflection tomography and focusing operator tomography is in the procedure. The reflection tomography procedure only contains one step, in which two-way travel times are transformed to a velocity model. The procedure using the focusing operators is split into two steps. First, the two-way seismic data is translated to one-way focusing operators in a data-driven way. Second, these focusing operators are used to estimate the velocity by using tomographic inversion. This second step will be discussed in this thesis.
<table>
<thead>
<tr>
<th>method</th>
<th>data</th>
<th>model parameters</th>
<th>goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVA tomography</td>
<td>Common Image Gather convert depth error to traveltime error</td>
<td>velocities and reflector depths</td>
<td>obtain a flat CIG</td>
</tr>
<tr>
<td>Reflection tomography</td>
<td>pre stack two-way traveltimes</td>
<td>velocities and reflector depths</td>
<td>minimize traveltime difference</td>
</tr>
<tr>
<td>Stereotomography</td>
<td>pre stack two-way traveltimes and slope of the events</td>
<td>velocities and ray segment pairs</td>
<td>minimize traveltime difference</td>
</tr>
<tr>
<td>Cross-well tomography</td>
<td>one-way traveltimes</td>
<td>velocities</td>
<td>minimize traveltime difference</td>
</tr>
<tr>
<td>Earthquake tomography</td>
<td>one-way traveltimes</td>
<td>velocities and earthquake locations</td>
<td>minimize traveltime difference</td>
</tr>
<tr>
<td>Equal-traveltime tomography</td>
<td>Differential Time Shift panel</td>
<td>velocities and reflector depths</td>
<td>obtain a flat DTS panel</td>
</tr>
<tr>
<td>Focusing operator tomography</td>
<td>one-way traveltimes</td>
<td>velocities and focal point locations</td>
<td>minimize traveltime difference</td>
</tr>
</tbody>
</table>

**Table 1.1**: Comparison of the main features of five conventional tomography methods, the equal-traveltime tomography, and the focusing operator tomography.
1.2 Statement of the problem

In summary, it can be concluded that for an accurate seismic image of the subsurface, an accurate velocity model is required, and that this velocity model can be obtained by inversion of focusing operators. Velocity model estimation methods based on the focusing concept have already been applied [Kabir and Verschuur, 2000; Hegge and Bolte, 2000].

Kabir and Verschuur [2000] do not split the procedure in two steps, but estimate the focusing operators and the velocity model simultaneously. The operators are modeled in an initial model and applied to the two-way data. The method is based on 'the principle of equal traveltime' between the focusing operator and the corresponding response in the focused data. If the subsurface model is incorrect, there will be a traveltime difference between the response and the operator, which is analyzed in the Differential Time Shift (DTS) panel. By means of the traveltime differences in the DTS panel the velocities and reflector depths are updated until the differences are minimized and, as a result, the event in the DTS panel is flat at zero time. Hence, this method is called equal-traveltime tomography (table 1.1). Note that this tomography method is analogous to the MVA tomography, except for the fact that it is performed in the traveltime domain. It is important to realize that the focusing operators do not form the input of the equal-traveltime tomography, but that they are determined concurrently with the velocity model, i.e. in a velocity driven way.

Hegge and Bolte [2000] do use the focusing operators obtained in a data-driven way. They make use of the two-step procedure, in which the focusing operators are obtained from the seismic data first. Next, the focusing operators are inverted to a velocity model. The used inversion method has the following properties Hegge [2000]: 1) It uses a geology based parameterization and 2) It is developed for two spatial dimensions (2D) only. However, it is preferable that the method is not biased by the geology based parameterization and is also applicable to three-dimensional (3D) problems. These requirements can be satisfied by making use of 3D traveltime tomography for the inversion of focusing operators using a data-driven model parameterization, which will be the subject of this thesis.

1.3 Objectives of this research

This research has three main objectives:

[1] Develop a method for (3D) tomographic inversion of focusing operators.

This method should:

a) Generate unique results;

The translation from focusing operators into a velocity model should be per-
formed accurately. This is the first requirement of the algorithm. Related to this aspect, it is important to consider whether the method generates unique results, and does not end in a local minimum.

b) Generate accurate results;
   Tomographic inversion problems are generally not well determined, which means that minima are not well defined in one or more directions. Related to this aspect it is important to consider whether the method has well-defined minima, and generates accurate results.

c) Be efficient in memory;
   Tomography can be quite expensive when it comes to memory use. In particular for 3D problems. Therefore, it is necessary to use efficient algorithms.

d) Be efficient in speed;
   Tomography can also become expensive when it comes to calculation time, so the speed of the algorithms is important as well.

e) Be flexible;
   Any parameterization should be possible.

   In order to achieve this:
   a) The result should not depend on the user input;
   b) The method should require minimal user intervention;

   Find out how to:
   a) Include a priori information, such as water velocity and well-log velocities.
   b) Apply a joint inversion of P and S-wave operators.
   c) Evaluate the inversion result in terms of focal quality.

In this thesis, the methodology is presented that addresses these three objectives. Also, the performance of the method developed and the fulfillment of the criteria will be evaluated on synthetic as well as real seismic data in both 2D and 3D.

1.4 Outline of thesis

In this thesis a method of tomographic inversion of focusing operators is presented. The structure of the thesis is shown schematically in Figure 1.3. The current chapter, Chapter 1, is introductory and summarizes the concepts of seismic exploration, seismic tomography, and focusing operators. This is also the chapter that states the problem and the objectives of this thesis.
The next three chapters deal with the theoretical framework.

**Chapter 2** is a tutorial on the relatively new theory on focusing operators. Both the theory behind the focusing operators, and how these operators are obtained from seismic data in practice are discussed. The use of focusing operators in tomographic inversion is the first important new aspect of this research.

**Chapter 3** forms the main theoretical part of the thesis. All the aspects of tomographic inversion are discussed, i.e. model parameterization, forward modeling, optimization, and regularization. For each different aspect an overview of existing algorithms is presented. Next, the algorithm that is used for tomographic inversion of focusing operators is introduced. Actually, in Chapter 3 the methodology related to the *first objective* is presented. The methodology related to the *second objective* is also addressed in this chapter, as section 3.6 reveals how the method becomes data-driven. The data-driven method is the second important new aspect of this research.

**Chapter 4** addresses the methodology related to the *third objective* of this research. It presents additional modules to the inversion method as introduced in Chapter 3. The methodology related to the inclusion of a priori information is discussed. In this thesis a priori information is formed by hard measurements like well-log data and water velocities. Soft constraints are not considered here. Chapter 4 also discusses the joint inversion of P and S-wave operators. Also the focal point cloud concept is introduced, which can be used to evaluate the inversion result. Both the implementation of the joint P and S-wave operator inversion and the introduction of the focal point clouds are new aspects, which could only be realized thanks to the use of the focusing operators.

The subsequent chapters evaluate the methodology of tomographic inversion of focusing operators, and consider the application of the additional modules. **Chapter 5** presents both 2D and 3D synthetic data examples and **Chapter 6** describes the results on real data.

Finally, in **Chapter 7** the advantages and limitations of the methodology are discussed. Furthermore, recommendations for future research are given.

The appendices in this thesis are mainly related to the theory as presented in Chapter 3. **Appendix A** presents the Delaunay triangulation method that is used for the model parameterization. Not only existing methods are reviewed, but also new solutions to the problems occurring in 3D Delaunay triangulation are presented. In **Appendix B** the mathematical description of the eikonal equation and the ray tracing system is given, which is used in the forward modeling step. **Appendix C** derives the partial derivatives with respect to quadratic slowness, which are required in the optimization step. In **Appendix D** the notation, symbols, and abbreviations, which are used in this thesis are explained.
**Figure 1.3:** Schematic outline of thesis.
Focusing operators

2.1 Introduction

In this research, the input data for tomographic inversion is formed by focusing operators. These operators can be constructed by means of the Common Focus Point (CFP) method [Berkhout, 1997a,b; Thorbecke, 1997]. This method makes it possible to extract from conventional seismic data in two-way traveltimes, focusing operators that represent one-way traveltime. Analyzing seismic data through the CFP concept gives more insight into the seismic imaging process, as the problem of focusing the data and positioning the data in depth are decoupled. The concept of one-way traveltime focusing operators can be used in many phases of seismic processing, like internal multiple removal [Verschuur and Berkhout, 2001], redatuming [Hindriks and Verschuur, 2001; Tegtmeyer et al., 2003], elastic imaging [Berkhout and Verschuur, 2000], AVP analysis [Winther and Verschuur, 2001], and tomographic inversion. Using the one-way traveltime focusing operators generated by the CFP method as input for tomographic inversion has the advantage that the inverse problem becomes inherently simpler than the two-way inverse problem.

The concept of the CFP method, and as a result the concept of the focusing operators, is based on the WRW model [Berkhout, 1997a,b]. The WRW model makes it possible to unravel seismic data into an upward and a downward wavefield, which is explained in section 2.2. When the upward and downward wavefield are focused, this results in an image of the subsurface. This focusing can be achieved by applying focusing operators to the seismic data. The focusing procedure, and the concept of focusing operators is described by the CFP method, which is explained in section 2.3. Finally, section 2.4 shows how in practice focusing operators can be obtained from seismic data.
2.2 The WRW model as a base for the CFP method

The WRW model forms the framework for the CFP method and the focusing operators. It makes use of a matrix notation of seismic data [Berkhout, 1982]. Seismic wavefields can be represented as continuous multidimensional functions, or as discrete matrices. Since seismic measurements are uniformly sampled at discrete 'points' in space and time, a matrix notation is more convenient.

Seismic measurements can be represented with the help of a data matrix, in which source gathers are represented as columns and receiver gathers as rows (Figure 2.1(2D)). In other words, 2D seismic data can be represented in a rank 2 tensor. A rank 2 tensor can also be used for 3D data, by sorting the data set in such a way that one column of the matrix still represents a shot gather (Figure 2.1(3D)). However, without sorting the data, the 3D data set can also be represented as a rank 4 tensor, in which the locations of both the sources and the receivers are defined by 2 coordinates each.

Each point in the data matrix contains information about the frequency spectrum, the traveltime, and the attenuation factors of the signal from a particular source to a particular receiver. Following the model presented in Figure 2.2, the data matrix $P$ can be expressed in terms of source, receiver, propagation and reflection operators:

$$ P(z_0, z_0) = D^-(z_0) \sum_{m=1}^{M} \left[ W^-(z_0, z_m) R(z_m) W^+(z_m, z_0) \right] S^+(z_0). $$ (2.2.1)
2.2 The WRW model as a base for the CFP method

![Schematic representation of the WRW model](image)

**Figure 2.2:** Schematic representation of the WRW model. The data matrix $P$ can be expressed in terms of source $S^+$, downward propagation $W^-$, reflection $R$, upward propagation $W^+$, and receiver $D^-$ operators.

When $P(z_0, z_0)$ is considered to be a rank 2 tensor, containing the data recorded at interface $z_0$, the following meaning can be attached to these symbols:

- $P(z_0, z_0)$ is the data matrix, each column containing the measurements of one source record.
- $S^+(z_0)$ is the source matrix, each column containing the angle dependent transmission properties belonging to one source location along $z_0$. When single pressure sources are used $S^+$ is a diagonal matrix.
- $W^+(z_m, z_0)$ is the downward propagation matrix, each column defining the propagation operator from one position at the surface $z_0$, to many positions on the reflector at $z_m$.
- $R(z_m)$ is the reflection matrix, each column defining angle dependency of the reflectivity for one reflection point along $z_m$. When the reflection coefficient is angle independent $R$ is a diagonal matrix.
- $W^-(z_0, z_m)$ is the upward propagation matrix, each row defining the propagation operator from one point in the subsurface at $z_m$, to many points on the surface at $z_0$.
- $D^-(z_0)$ is the receiver matrix, each row containing the angle dependent detection properties belonging to one receiver at $z_0$. When single point receivers are used $D^-$ is a diagonal matrix.
- $z_0$ is the source and receiver interface.
- $z_m$ is the depth of the target reflector.
- $M$ is the total number of reflectors.
The superscripts '+' and '-' refer to downward and upward wavefields respectively. By means of this equation, the observed data can be described when the source, receiver, subsurface and reflector properties are known. The source and receiver matrices $S^+$ and $D^-$ are known or can be estimated, but the influences of propagating through the subsurface, expressed by the matrices $W^+$ and $W^-$, are unknown. However, when an estimated model of the subsurface is constructed, these parameters can be obtained by forward modeling. Finally, only the response of the reflector $R$ is unknown and this is the parameter that forms the target of seismic imaging. In other words, to obtain an image of the subsurface, the wave propagation matrices $W^+$ and $W^-$ and the acquisition effects $S^+$ and $D^-$ have to be removed from the data.

For the sake of simplicity, in the remainder of this section $S^+$ and $D^-$ are assumed to represent ideal single point sources and receivers that have the same response, so they are unity matrices. As in this research the seismic data is used for inversion of the kinematic effects only, and not for inversion of the reflection amplitudes, the reflection operator $R$ is assumed to be a constant diagonal matrix. For the response
from a single interface at \( z_m \) only, equation (2.2.1) becomes:

\[
P(z_0, z_0) = W^-(z_0, z_m) W^+(z_m, z_0).
\] (2.2.2)

In words, the kinematic information related to the target horizon at \( z_m \), contained in the seismic data, will only be described by a down-going wavefield operator \( W^+(z_m, z_0) \), and an up-going wavefield \( W^-(z_0, z_m) \) (Figure 2.3).

### 2.3 The CFP method as a base for focusing operators

The CFP method makes use of the WRW model to image the earth. This method has been discussed extensively by Berkhout [1997a,b] and Thorbecke [1997]. In this section an overview will be given. In the CFP method, imaging can be formulated as two consecutive focusing steps: focusing of the sources and focusing of the receivers. The sources and receivers are focused at a location in the subsurface to obtain the reflection image at that location. Focusing is performed by applying a focusing operator to the data matrix in equation (2.2.1).

#### 2.3.1 Focusing of the receivers

Focusing of the receivers to a position \( j \) at \( z_m \), occurs by applying a focusing operator \( F_j^{\dagger-} \) to the upward propagation side of the data matrix:

\[
P_j^{\dagger+}(z_m, z_0) = F_j^{\dagger-}(z_m, z_0) P(z_0, z_0) \\
= F_j^{\dagger-}(z_m, z_0) W^-(z_0, z_m) R(z_m) W^+(z_m, z_0). \] (2.3.3)

The focusing operator \( F_j^{\dagger-} \) is a row vector that is defined such that

\[
F_j^{\dagger-}(z_m, z_0) W^-(z_0, z_m) = I_j^{\dagger-}(z_m),
\]

with \( I_j^{\dagger-}(z_m) = (0, 0, ..., 1, 0, ..., 0) \). (2.3.4)

The vector \( F_j^{\dagger-}(z_m, z_0) \) defines an operator for the focusing of the receivers with its focal point on the interface \( z_m \) at position \( j \). In other words, the focusing operator is defined in such a way that it removes the upward propagation properties, so it has to be equal to the inverse of the upward propagation matrix. A stable approximation of the inverse of the upward propagation matrix is given by the complex conjugate of the downward propagation matrix, so \([W^+]^{-1} = [W^+]^* \) [Berkhout, 1982; Wapenaar and Berkhout, 1989]. The complex conjugate in the frequency domain is the same as the time reverse in the time domain. This approximation honors the kinematics of the propagation, but neglects the amplitudes of the focusing operator when it passes
through velocity and density variations. Using this approximation, the focusing operator can be defined by one row of \([W^+(z_m, z_0)]^*\):  

\[
\mathbf{F}^+_{j, j}(z_m, z_0) = \mathbf{I}^+_{j, j}(z_m)[W^+(z_m, z_0)]^*,
\]

(2.3.5)  

and as a result equation (2.3.3) becomes  

\[
\mathbf{P}^+_{j, j}(z_m, z_0) = \mathbf{R}^+_{j, j}(z_m)W^+(z_m, z_0)
\]

\[
\propto [\mathbf{F}^+_{j, j}(z_m, z_0)]^*.
\]

(2.3.6)  

The result \(\mathbf{P}^+_{j, j}(z_m, z_0)\) of this focusing step is called a Common Focus Point (CFP) gather, containing the reflection information of, and the downward propagation effects to, the target horizon at \(z_m\) (Figure 2.4). From equation (2.3.6) it can be observed that the CFP gather should coincide with its time reversed focusing operator. This feature is also known as 'the principle of equal travelttime' [Berkhout, 1997a]. Note that the CFP gather will contain reflection from other horizons than the one at \(z_m\). However, the principle of equal traveltime applies only to the target horizon to which the receivers were focused.
2.3 The CFP method as a base for focusing operators

2.3.2 Focusing of the sources

Focusing of the sources to a position \( j \) at \( z_m \) occurs by applying a focusing operator \( \mathbf{F}_j^+ \) to the downward propagation side of the data matrix:

\[
\mathbf{P}_j^-(z_0, z_m) = \mathbf{P}(z_0, z_0) \mathbf{F}_j^+(z_0, z_m) = \mathbf{W}^-(z_0, z_m) \mathbf{R}(z_m) \mathbf{W}^+(z_m, z_0) \mathbf{F}_j^+(z_0, z_m),
\]

(2.3.7)

where

\[
\mathbf{W}^+(z_m, z_0) \mathbf{F}_j^+(z_0, z_m) = \mathbf{I}_j^+(z_m) \quad \mathbf{I}_j^+(z_m) = (0, 0, ..., 0, 1, 0, ..., 0)^T.
\]

(2.3.8)

So, the column vector \( \mathbf{F}_j^+(z_0, z_m) \) defines an operator for the focusing of the sources with its focal point on the interface \( z_m \) at position \( j \). As a consequence, in this focusing step the downward propagation properties are removed. Therefore, the focusing operator has to equal the inverse of the downward propagation matrix. This inverse matrix can be approximated by the complex conjugate of the upward propagation matrix, so \( [\mathbf{W}^+]^{-1} = [\mathbf{W}^-]^* \). Using this approximation, the focusing operator can be defined by one column of \( [\mathbf{W}^-(z_0, z_m)]^* \):

\[
\mathbf{F}_j^+(z_0, z_m) = [\mathbf{W}^-(z_0, z_m)]^* \mathbf{I}_j^+(z_m),
\]

(2.3.9)

and as a result equation (2.3.7) becomes

\[
\mathbf{P}_j^-(z_0, z_m) = \mathbf{W}^-(z_0, z_m) \mathbf{R}_j(z_m) \alpha [\mathbf{F}_j^+(z_0, z_m)]^*.
\]

(2.3.10)

The result \( \mathbf{P}_j^-(z_0, z_m) \) of this focusing step is also a CFP gather, but in this case it contains the reflection and the upward propagation effects from the target to the surface (Figure 2.5). Again, the 'principle of equal traveltime' is valid, because equation (2.3.10) indicates that the CFP gather should coincide with its time reversed focusing operator.

2.3.3 Applying both focusing steps

By applying both focusing steps together, using a focal point at position \( j \) at the interface \( z_m \), both the sources and the receivers are focused at this location. The data matrix is reduced to one diagonal element of the reflection matrix which equals the reflection coefficient that is averaged over all angles of incidence:

\[
\mathbf{F}_j^{+-}(z_m, z_0) \mathbf{P}(z_0, z_0) \mathbf{F}_j^{+-}(z_0, z_m) = \mathbf{I}_j^{+-}(z_m) \mathbf{R}(z_m) \mathbf{I}_j^{+-}(z_m) = R_{jj}(z_m).
\]

(2.3.11)

This is the full stack image for one point in the subsurface. When this is repeated for each subsurface point of interest, a complete image of the subsurface is obtained.
Figure 2.5: Focusing of the sources by applying a focusing operator $F_j^+$ to the data matrix. This results in a CFP gather that contains the reflections $R$ and upward propagation properties $W^-$ of the seismic data.

One requirement for successful focusing of either the sources or the receivers is that the measured data should contain full reflection information from the defined focal point. This means that a full coverage of sources and receivers is required.

Although applying both focusing steps generates a final image of the subsurface, the outcome of applying one focusing step, the CFP gather, is a useful intermediate result. As shown in this section, focusing using a correct operator results in a traveltime response in the CFP gather that equals the traveltime of the focusing operator, referred to as the 'principle of equal traveltime' (Figure 2.6). A correlation in time of this (time reversed) focusing operator and its corresponding response in CFP gather, yields an aligned event at zero-time. This correlation means subtraction of the operator traveltime from the response in the CFP gather, which results in a Differential Time Shift (DTS) panel. This panel can be used to evaluate whether the used focusing operator is correct. When the DTS panel shows an aligned event at zero-time, a correct operator was used. When the event deviates from zero-time, this can be used to update the focusing operator.
2.4 Focusing operator updating

2.4.1 Data-driven updating of focusing operators

In the previous section the focusing operators resulted in accurate focusing of either the upward or downward propagation properties of the data. These operators were obtained from propagation matrices $W^+$ and $W^-$ that belong to an accurate subsurface model. However, usually the correct subsurface and as a consequence the correct propagation matrices and focusing operators are unknown. This section describes how the CFP method is used to obtain accurate focusing operators from the seismic data.

An initial estimate of the focusing operator can be made by means of the propagation matrices of an estimated subsurface model. However, when the model is erroneous, the estimated propagation matrices $W^+$ and $W^-$ used in constructing the focusing operators are not equal to the true propagation matrices $W^+$ and $W^-$. Let’s consider the focusing of sources where an error in the focusing operator is introduced.

Figure 2.6: a) The correct focusing time-reversed operator for focusing of the sources. b) The CFP gather after focusing of the sources. The response corresponding to the focal point equals the focusing operator. c) The DTS panel is aligned at zero-time.
by an error in the propagation matrix [Berkhout and Verschuur, 2001]:

\[
\begin{align*}
\hat{W}^{-}(z_0, z_m) &= W^{-}(z_0, z_m) \Delta W(z_m), \\
\hat{P}_j^+(z_0, z_m) &= [W^{-}(z_0, z_m)]^* [\Delta W(z_m)]^* \mathbf{I}_j^+(z_m), \\
&\approx [W^+(z_m, z_0)]^{-1} [\Delta W(z_m)]^* \mathbf{I}_j^+(z_m).
\end{align*}
\]

Note that in this illustration the error is located near the interface \(z_m\). When this erroneous focusing operator is applied to the data, this results in:

\[
\begin{align*}
\hat{P}_j^{-}(z_0, z_m) &= \mathbf{P}(z_0, z_0) \hat{P}_j^+(z_0, z_m), \\
&= W^{-}(z_0, z_m) R(z_m) W^+(z_m, z_0) \hat{P}_j^+(z_0, z_m), \\
&\approx W^{-}(z_0, z_m) R(z_m) [\Delta W(z_m)]^* \mathbf{I}_j(z_m).
\end{align*}
\]

When this CFP gather \(\hat{P}_j^{-}(z_0, z_m)\) is compared with the time reversed focusing operator used,

\[
[\hat{P}_j^+(z_0, z_m)]^* = W^{-}(z_0, z_m) \Delta W(z_m) \mathbf{I}_j(z_m),
\]

it is observed that an error in the operator gives the opposite error in the CFP gather \(\Delta W(z_m)\). This feature is a useful tool, because when both errors are opposite in sign, the correct traveltime is situated between the two erroneous traveltimes. It is on this feature that the focusing operator updating scheme is based.

The procedure for focusing operator updating is represented schematically in Figure 2.7. The DTS panel can be used to update an erroneous operator in a data-driven way [Berkhout, 1997a; Thorbecke, 1997; Bolte, 2003]. When the event in the DTS panel is not aligned at zero-time, this indicates that the focusing operator that was used for focusing is incorrect. However, the focusing operator can be updated in time by using the time difference in the DTS panel. Splitting the difference and adding it to the focusing operator generates a new focusing operator. This operator can be used to re-apply the focusing procedure. Finally, after some iterations, the alignment in the DTS panel is achieved and the correct focusing operators have been obtained. Note that by directly updating the operator from the observed traveltime differences, the underlying velocity model as well as the locations of the focal points remain unknown. Therefore, this is a fully data-driven operator estimation technique.

### 2.4.2 Updating of focusing operators using sparse data

The data-driven updating of focusing operators has been applied to 2D data [Bolte and Verschuur, 1998] and 3D data [Verschuur et al., 1999] with a full source and receiver coverage. For application of the CFP method to sparse (3D) data a problem arises, as there is no full areal coverage in the source and receiver domains which is required for application of the CFP method. Therefore, a focusing operator updating procedure that is applicable to sparse 3D data has been developed. The core of this
2.4 Focusing operator updating

Figure 2.7: Left: Flowchart for data-driven focusing operator updating. Middle: Gathers that correspond to the steps in the flowchart. Right: Schematic ray pictures that correspond to the gathers.

The procedure is still formed by the data-driven focusing operator updating. However, it will be assisted by additional steps, in which the missing data is filled in and the tomographic inversion of the operators is used during operator updating. In Figure 2.8 the complete procedure is depicted [Gisolf et al., 2002]. The procedure contains four steps:
[1] Calculation of focusing operators;
[2] In-fill of the data to a full 3D data set using the operators;
[3] Data-driven focusing operator updating;

These steps are explained more extensively in the remainder of this section.

Figure 2.8: Flowchart for updating of focusing operators for sparse 3D data. The numbers refer to the steps as explained in the text.

**Calculation of focusing operators (step 1)**

The procedure starts with the estimation of initial focusing operators and focal point locations. This can be performed by using an initial velocity model, and calculating the traveltimes of the focusing operators by ray-tracing. This does not differ from the conventional updating procedure as described in the previous section, where initial focusing operators are also required to start the updating.

The focusing operators can be used to define constant $p$ (ray-parameter)-contours at the surface, which are needed for the in-fill of the data. The ray-parameter values posted at the surface where rays emerge, are the $p$ values of those rays at
the reflection point. To obtain these \( p \)-values at the reflection point, the zero-offset times and time-dips of the target event, which can be obtained from the seismic data, are required. By means of the velocity model, the zero-offset time and time-dip can be translated to a focal point location and a dip of the reflector at this focal point. The zero-offset ray visualizes this translation from the surface to the subsurface properties. As the zero-offset ray is normal to the reflector dip at the focal point, the \( p \)-values are calculated with respect to this zero-offset ray during the ray-tracing (Figure 2.9a). Both the \( p \)-contours and the initial focusing operators are used in the in-fill of the data to a full coverage 3D data set.

![Figure 2.9: a) The constant ray-parameter contours (\( p \)-contours) with respect to the zero-offset ray (dashed arrow) are determined during the calculation of the focusing operator. b) An available source-receiver recording (filled star and triangle) is rotated along its corresponding \( p \)-contour (so around the zero-offset ray) and duplicated (open star and triangle) to fill in the data.](image)

**In-fill of the data to a full 3D data set (step 2)**

In order to apply focusing operator updating and imaging to sparse 3D data, it is necessary to fill in the gaps in areal coverage in the source and receiver domains. This in-fill procedure is based on the principle that for a given focal point location, the reflection from that location sampled by a source-receiver pair can be duplicated to other non-sampled source-receiver pairs [Gisolf et al., 2002; van de Rijzen et al., 2003]. The in-fill procedure is performed by rotation of the actual recorded ray around the zero-offset ray of the reflection point of this ray. In this rotation the ray-parameter \( p \) at the reflection point is the same for the sampled and non-sampled source-receiver pair, but the azimuth changes (Figure 2.9b). In practice this means that the sampled rays are rotated along the \( p \)-contours. The assumption is that each contour of each focal point contains at least one source-receiver combination that is actually acquired. If not, some form of interpolation is performed. In this way traces are duplicated by means of the ray-parameter contours. The correct two-way traveltimes for the target event of the duplicated traces are derived from the one-way
Focusing operators

traveltimes at the source and receiver locations before and after rotation. In this way the focusing operators and the p-contours are used to obtain the traveltimes and the reflections of non-sampled traces. When all contours for all focal points are filled, full 3D coverage of the target horizon is obtained. Therefore, this in-fill is only correct for the event related to the focal point. Note that the in-fill procedure is only correct if the operator is correct.

Data-driven focusing operator updating (step 3)

The focusing operators calculated in the first step and the in-filled data calculated in the second step are used in the data-driven updating of the focusing operators. This procedure has been presented in section 2.4.1 and can be carried out because the data is 'full 3D' now. By using the criterion of the flat DTS panels, this step results in new updated focusing operators. When these updated operators are sufficiently close to the operators of the previous update loop, the procedure stops and the updated operators represent the true focusing operators of the data set. When the difference between the updated and the previous operators is not small enough, we have to go through another iteration.

Tomographic inversion of the focusing operators (step 4)

The updated focusing operators are tomographically inverted to obtain a new velocity model. As a result, this velocity model will match the current focusing operators as well as possible. From this model a new physically consistent set of operators is determined by raytracing. First, the focusing operators will determine the model, then the model will make the operators physically consistent. There is no more complexity in the model than needed to explain the operators. This new velocity model is replacing the previous model and is used to:

- convert the zero-offset time and time-dip to the location of the focal point and the dip of the reflector at the focal point position;

- generate new focusing operators and p-contours.

Moreover, it ensures lateral and vertical consistency of the focusing operators. This is particularly important in combination with the data-driven focusing operator updating, as this will not necessarily lead to a physically realizable set of operators.

The four steps of the focusing operator updating procedure (Figure 2.8) are repeated till no more significant updates can be made. This is the point that the final accurate focusing operators are obtained.
2.5 Conclusions

In this chapter the concept of focusing operators is discussed, and it is shown how these operators can be extracted from the data by using the WRW-model and the CFP method. The data-driven estimation of these operators for both ’full coverage’ and ’sparse’ seismic data is presented. The operators contain the kinematic effects of the subsurface, which can be seen as the one-way traveltimes from a point in the subsurface to locations at the surface (Figure 2.10). These focusing operators can also be considered as the traveltimes of the Green’s functions that describe the propagation effects of the subsurface. This will form the input for tomographic inversion of focusing operators.

Figure 2.10: Illustration of the traveltimes of a focusing operator in a 2D and a 3D situation.
Tomographic inversion methodology

3.1 Introduction

In the previous chapter the concept of focusing operators is explained. The focusing operators are very suitable for velocity model estimation, because they represent one-way traveltimes, which are more suitable for inversion to a velocity-depth model. For the inversion of focusing operators we want to develop a data-driven velocity estimation method, that is applicable to 3D problems. This is done by 3D traveltime tomography in which the kinematic information, the traveltimes of the focusing operators, are inverted to a velocity model.

In this chapter, the methodology of tomographic inversion applied to focusing operators is discussed, which is the basis for the first objective of this research (section 1.3). In section 3.2, the fundamentals of the method are described. The non-linearity of the problem and the procedure that is used to solve the tomographic problem are addressed. The procedure contains four main steps. The first three steps: parameterization, forward modeling, and optimization are discussed more thoroughly in the subsequent sections 3.3, 3.4, and 3.5 respectively. The fourth step, which is a very important part of tomographic inversion of focusing operators, is the regularization of the inversion. This is the part that makes the method data-driven, as will be presented in section 3.6, which is the basis for the second objective of this research (section 1.3).
3.2 Fundamentals of tomographic inversion

3.2.1 The non-linearity of tomographic inversion

In traveltime tomography the velocity model parameters are estimated from the traveltimes of seismic data. If the relation between the model parameters and the data would be linear, a result could easily obtained. However, the tomographic problem is in principle non-linear, because the data depends on the model parameters and the ray-path geometry of the data, and the latter itself depends again on the model parameters. In order to be able to invert for the model parameters, the tomographic problem needs to be linearized around an approximate ray-path.

Approximate ray-paths can be defined by the introduction of an initial model of the subsurface. In this initial model, the approximate ray-paths of the focusing operators can be calculated. The difference between the traveltimes of the modeled ray-paths and the observed traveltimes becomes the data to be inverted. As a consequence, the difference between the initial model and the actual subsurface (the update of the model parameters) becomes the output of the inversion. This relation is assumed to be linear, and forms the base of many tomography studies [Tarantola, 1986; Nolet, 1987; Scales, 1987; Phillips and Fehler, 1991]. The relation is represented by:

$$\Delta d = A \Delta m,$$  \hspace{1cm} (3.2.1)

in which $\Delta d$ is the vector that contains the traveltime differences for all measurement locations, $\Delta m$ is the vector that contains the model parameter updates, and $A$ is the matrix that describes a linear relation between the differences in traveltimes and the model parameter updates. In this way, the tomographic problem is linearized, and the inversion of the traveltime differences leads to an update of the initial model parameters. Note that this linear relation is only an approximation of the true relation between $\Delta d$ and $\Delta m$. In the updated model, new traveltime differences can be calculated and these differences can be inverted again, and so on. This iterative process is commonly denoted as non-linear tomography.

3.2.2 The procedure of tomographic inversion

In order to set up the non-linear tomographic system for the focusing operators, the following steps should be made:

[1] Parameterization;
   an (initial) velocity model of the subsurface is defined.

[2] Forward modeling;
   in the velocity model focusing operators are modeled. The traveltime differences between the observed and modeled focusing operators is calculated.
[3] Optimization;
the differences between the traveltimes of the observed and the modeled focusing operators is minimized, in order to optimize the model parameters. The updates of the model parameters are added to the initial model to obtain the updated velocity model.

Once the velocity model is optimized in the third step, one can return to the second step, to get further optimization of the velocity model. If the traveltime differences are small enough an accurate velocity model is obtained. The strategy for focusing operator inversion is presented in Figure 3.1. The parameterization, the forward modeling, and the optimization are described more thoroughly in the next subsections.

![Figure 3.1: Procedure flowchart of tomographic inversion of focusing operators.](image-url)
3.3 Model parameterization

In this section the parameterization of the velocity model is discussed, which is the first main step in the tomographic inversion of focusing operators. In this tomographic inversion, the model is described by parameters that influence the travel-times: the velocities and the location of the focal points in the subsurface. The velocity parameterization has to describe a complete area or volume, and the focal point parameters describe the position of each focal point.

The parameterization of the focal point is simple, as its location is described by a vector containing the coordinates. However, the number of focal points that are used can be critical, as described later in this section.

The velocity parameterization can be more complex. The key to the success in tomographic inversion is the definition of an optimum number of parameters that accurately describe the subsurface velocity structure. Too many parameters result in an under-determined system during the inversion process, while too few parameters will be unable to describe the subsurface accurately. So, for the velocity parameterization it is important to find a method that optimally describes the subsurface.

3.3.1 Overview of existing methods

There are several methods for parameterization of the subsurface model. The three categories that will be discussed are:

- layers defined by splines (global basis functions);
- regular grid (local basis functions);
- irregular grid (local basis functions).

Layers

Parameterization by means of layers (Figure 3.2a) seems straightforward, as in general, the geology consists of a layering of depositional units having variable thicknesses. A way to describe the subsurface is by defining layers and velocities within the layers by splines (i.e. by global basis functions) [van der Made, 1988; Hegge, 2000]. The depths of the boundaries of the layers are chosen at the reflectors and the velocities within the layers can be taken either homogeneous, vertically varying, or laterally varying. The advantage of the parameterization by layers is that the number of parameters to be inverted for is limited. However, the change in velocity does not always coincide with the reflector on which the boundary of the layer is defined. On the other hand, a reflector that is not clearly visible can also cause a considerable velocity change, while there is no layer defined at that position. So, the
result will depend very strongly on the location and the number of layers that are
chosen.

\[ \text{Figure 3.2: Different methods of parameterization presented in 2D. a) Layers: the boundaries are defined by splines through a set of definition points and the velocities are defined by functions based on a few velocity points within the layers. b) Regular grid: the velocities are defined within the cells or at the grid-points. c) Irregular grid connected by Delaunay triangulation: the velocities are defined within the triangles or at the grid-points.} \]

**Regular grid**

Parameterization by means of a grid has a maximum flexibility in modeling the subsurface. In a regular grid parameterization the subsurface is defined by regular grid-points (Figure 3.2b). These points are connected which results in rectangular cells (local basis functions). The velocity of the subsurface can be defined at the grid-points or within the cells. As a result, every point/cell can adopt an optimum velocity which results in a method that can describe any kind of subsurface. In this way, the velocity of the model is updated, and the velocities are not tied to predefined layers. A drawback of the regular grid method is that it generates a very large number of parameters, which can result in an inaccurate inversion.

**Irregular grid**

In parameterization by means of an irregular grid method (Figure 3.2c) the velocity is also defined by a grid, but the number and the position of the grid-points is flexible. For example, the density of the grid-points can be chosen higher at regions with high heterogeneity (model dependent [Vesnaver et al., 1995]) or at regions where more data information is available (data dependent [Spakman and Bijwaard, 2001]). The points can be connected by polygons, which results in a grid. A large number of polygon definitions and gridding methods exist, but only two important methods, the Voronoi polygon and the Delaunay triangulation, will be discussed in the next part. In Figure3.2c a Delaunay triangulation is used. The velocity of the subsurface
can be defined at the grid-points or within the cells (polygons). An irregular grid can
describe any kind of subsurface and uses the parameters efficiently. This can result
in a reduction of parameters with respect to the regular grid method. Although the
irregular grid method is very flexible, it requires more bookkeeping than the regular
grid method [Sambridge et al., 1995; Böhm and Vesenaver, 1996].

**Voronoi polygons and the Delaunay triangulation**

Although there exist a large number of polygon definitions and gridding methods,
only two of them are discussed here: the Voronoi polygon and the Delaunay trian-
gulation [Delaunay, 1924; Virieux, 1908; Watson, 1981; Sambridge et al., 1995].
The reason for this is that these methods have attractive spatial properties and are
closely related to each other. The Voronoi polygon of an irregular set of grid-points
divides the plane into a set of regions, one for each grid-point, such that any point
in a particular region is closer to that region’s grid-point than to any other grid-
point (Figure 3.3a). The Delaunay triangulation is formed by simply connecting the
grid-points whose Voronoi polygons have common boundaries (Figure 3.3b). The
Delaunay triangulation results in the set of least ‘long and thin’ triangles that can be
generated among the many triangulations that are possible with the irregularly dis-
tributed grid-points. In a 3D situation Delaunay triangulation generates tetrahe-
dra, and the Voronoi polygons become Voronoi polyhedra. Because of the spatial prop-
ties of the Voronoi polygons and the Delaunay triangles, their surface will reflect the
local grid density. The Delaunay triangles are generally more convenient for build-
ing parameterizations containing particular interfaces, because the grid-points can
be distributed on the interfaces. The Voronoi polygons are more useful when areal
or volumetric coverage is of concern [Sambridge and Gudmundsson, 1998; Böhm and
Vesenaver, 1999].

![Figure 3.3: Two ways of defining grids based on a distribution of grid-points by a) Voronoi polygons and b) Delaunay triangulation.](image)
3.3.2 Parameterization used in tomographic inversion of focusing operators

The parameterization that is used for tomographic inversion of focusing operators is presented here for both the 2D and the 3D situations.

Parameterization of the velocity

For the parameterization of the velocity in the subsurface, an irregular parameterization is used. This method is chosen, because it does not introduce a bias in the result and allows to avoid over-parameterization of the subsurface. It was already discussed that the gridding can be done by both Delaunay triangulation and Voronoi polygons. Despite the fact that Voronoi polygons are better for volumetric coverage [Böhm and Vesnáver, 1999], the Delaunay triangulation has been chosen for the gridding. The motivation for this choice is that it is desirable to generate a smooth velocity model without abrupt transitions between the cells. Continuity across the faces of the cells can be achieved by specifying the velocity at the faces of the cells. This can only be achieved by defining the velocities at the grid-points using a Delaunay triangulation.

To get continuity across the interfaces in the velocity model, it is not only necessary to use Delaunay triangulation, but it is also required to define the velocities at the grid-points. Besides generating continuity, defining velocities at grid-points also reduces the number of parameters. When velocities are defined at grid-points and not within cells, the number of parameters can be reduced by a factor of two in 2D models. This is due to the fact that each triangle is defined by three grid-points, but each grid-point can define more than three triangles (see Figure 3.3b). In 3D models this can even be up to a factor of five or six. The velocity within the Delaunay triangulation is calculated by a linear interpolation of the velocities at the grid-points defining each triangle.

The Delaunay triangulation forms the triangles and tetrahedra that are used in this research, which are presented in Figure 3.4. In Figure 3.5a,b it is shown how the Delaunay triangulation is used to define a velocity model. How Delaunay triangulation is carried out and how points can be added to an existing triangulation will be described extensively in appendix A.

Parameterization of the focal points

Besides the velocity parameters there are also the other model parameters that influence the traveltime: the locations of the focal points. These can simply be described by a vector containing the coordinates. The focal point locations are related to positions where reflection energy is available, which can be anywhere in
the subsurface. The focal points are defined independently of the velocity points (white dots in Figure 3.5c). By parameterizing them independently of the velocity grid-points, the velocity changes are not dependent on, or restrained to the reflectors. This generates more freedom than the method of layer parameterization, where velocity changes are connected to reflectors.

Each focusing operator is connected to one focal point and, as a consequence, more focal points means more data. This suggests that an endless number of focal points would be optimal. However, despite the fact that focusing operators originate from different focal points, they may still contain similar information. This is, for example, the case when focal points are too closely positioned to each other. A good distribution of focal points within the subsurface is therefore very important.
3.3 Model parameterization

Figure 3.5: Parameterization of 2D and 3D velocity models. a) Irregular parameterization: a velocity is defined at the black grid-points (for sake of simplicity only 9 points are shown in the 3D example. b) The irregular grid is connected by the Delaunay method resulting in triangles (2D) or tetrahedra (3D). The velocity within the Delaunay triangulation (shown for one triangle and one tetrahedron) is calculated by a linear interpolation between the velocity values at the grid-points defining each triangle/tetrahedron. c) The focal point locations (open dots) are defined independently of the velocity points (black dots).
3.4 Forward modeling of traveltimes

In this section, the forward modeling of the focusing operators is discussed, which is the second main step in the tomographic inversion of focusing operators. Forward modeling is done to obtain the traveltimes of focusing operators in the current velocity model. The differences between the traveltimes of these operators and the observed focusing operators are used to optimize the parameters of the velocity model. The forward modeling is also needed to define the relation between the traveltime data and the model parameters.

3.4.1 Overview of existing methods

Forward modeling of the traveltimes requires a model of the subsurface that contains parameters that influence these traveltimes. The quality of the tomographic inversion is not only dependent on the quality of the model, as discussed in the previous section, but also on the quality of the forward modeling. Therefore, it is important to consider the main features of the forward modeling algorithms, such as the accuracy of the method, the capacity to model geophysical events, and, last but certainly not least, the cost of the method in terms of both speed and memory. There are several ways to forward model traveltimes in the subsurface [Audebert et al., 1995; Hegge, 2000]. They can roughly be divided into three groups:

[1] Ray-based methods;

Ray-based methods

The first forward modeling approach that is considered is the ray-based method. In this method, the ray geometry and the corresponding traveltimes for a seismic ray between two points (e.g. source and receiver) are calculated (Figure 3.6). This is called geometric ray-tracing. The calculation is done by a high-frequency ray approximation of the scalar wave-equation; the eikonal equation. The fundamentals of the method are described by Červený [2001]. The method does not only provide the traveltimes, but also provides the corresponding ray-paths. This is an important advantage, because these ray-paths directly describe the relationship between traveltimes and model parameters that is needed in tomographic inversion. Moreover, the ray-based method can handle complex model structures, multi-valued arrivals, and computes traveltimes accurately (Figure 3.8a). Once the geometric ray-tracing is extended to dynamic ray-tracing [Červený, 2001], it can find amplitudes and phase...
3.4 Forward modeling of traveltimes

\[ T \]

Figure 3.6: Ray-based method: Calculation of the ray geometries and the corresponding traveltimes for seismic rays. Note that shadow-zones can be present due to the complex velocity distribution in the model.

Grid-based methods

In the past decade, several methods have been introduced to calculate the traveltimes directly on a regular grid [Vidale, 1988]. Again, the calculation is done by a high-frequency ray approximation. The eikonal equation is solved by finite-differences on a rectangular grid. The incremental traveltimes are calculated on an ever-expanding rectangular front (Figure 3.7). There is no relaxation and no updating back into the

shifts. A major advantage, particularly in 3D problems, is that the method requires relatively little memory.

However, the method has some drawbacks. Although it requires little memory, it is expensive in calculation time compared to the grid-based methods, which will be described in the next subsection. Another disadvantage is that the method makes use of a high-frequency ray approximation. As a consequence, the solution of the eikonal equation is not necessarily close to the solution of the wave equation with a band-limited source. This is particularly the case if the velocity model contains small structures compared to the wavelengths of the seismic signal. This problem can be dealt with by smoothing the velocity model. Lomax [1994] solves this problem by ray-path dependent smoothing of the model. Spetzler and Snieder [2001] used the variations in the vicinity of the ray to obtain a band-limited solution. However, these procedures will increase the costs of the method. Another drawback of the ray-based method is that it produces shadow-zones, where the ray theory breaks down. This problem can be dealt with by extending the method with Gaussian beam modeling [Cervený and Pšencík, 1983], which generates a good result in the transition zone between an illuminated region and a shadow-zone. Another way to circumvent the shadow-zone problem is by including wavefront reconstruction, but also these methods will increase the costs [Vinje et al., 1993].
inside of the computation front (Figure 3.8b).

This finite-difference method is computationally quick and complicated velocity structures can be handled. Another important advantage is that shadow-zones are avoided as traveltimes are calculated throughout the complete grid. By considering the computational front, the first arrival is automatically followed. This is also a drawback of the method, as the first arrival is not necessarily the most important arrival. Moreover, amplitudes and phases are not calculated by the finite-difference method. Podvin and Lecomte [1991] made an improvement by introducing the Huygens' principle. In this way, the influences from inside of the computation front are taken into account. However, this makes the method more expensive. Another drawback is that the grid-based method assumes a square front. This can be improved by defining the expanding front as a semi-circle [van Trier and Symes, 1991]. The disadvantage that remains, for the use of tomographic inversion, is that ray-paths are not calculated explicitly, so the relationship between the traveltimes and the model parameters does not result from forward modeling automatically. However, this relation can be obtained by the perturbation method [Hegge, 2000]. As this should be done for each model parameter separately, it causes a serious decrease in the efficiency of the algorithm.

**Limited bandwidth method**

This method does not use the high-frequency approximation. It is a reduced version of a full wavefield modeling [Nichols, 1996]. For a decimated set of frequencies the wavefield is modeled, so the wavefield is correctly modeled only in a small time

---

**Figure 3.7**: Grid-based method: Grid used in the finite-difference algorithm for wave-front modeling. a) The source grid-point A and the eight points in the ring surrounding point A are used in the first step. b) 2D grid as the numerical calculation of traveltimes is progressing. The ring of the points shown as filled circles are about to be timed. The hollow circles indicate points that have had their traveltimes calculated. The double circle in the middle shows the source point. The small dots are not timed yet [after Vidale, 1988].
3.4 Forward modeling of traveltimes

<table>
<thead>
<tr>
<th></th>
<th>ray-based</th>
<th>grid-based</th>
<th>limited bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>geophysical events</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>amplitudes</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>multi-valued arrivals</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>accuracy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>traveltime accuracy</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>shadow-zones</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>limited bandwidth</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td><strong>costs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>speed efficiency</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>memory efficiency</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>derivation of</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>partial derivatives</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Table 3.1: Comparison of the main features of three different forward modeling methods.*

window around the maximum energy arrival (Figure 3.8c). By keeping track of this time window, it is possible to pick either first or most energetic arrivals. In this way, at a cost lower than full wavefield modeling, a reliable traveltime is produced. The corresponding phases and amplitudes are also estimated. This reduced output is similar to the output of the ray-based method, but it is expected to be more reliable.

The main advantage of the method is that it is a traveltime computation method, but it takes into account the frequency-dependence of propagation. Moreover, no shadow-zones are produced as a solution is found at every point in the subsurface. However, the method is expensive, which is the main drawback. Besides, no explicit ray-paths are calculated, which forces the user to develop a costly algorithm to obtain the relationship between traveltime and model parameters.

The main features of the methods described above are presented in Table 3.1. For this comparison only the core of the methods is considered. The possible extensions that were described above are excluded. This table indicates that each method has its specific advantages. The limited-bandwidth methods have the highest accuracy, the grid-based methods have the highest speed, and the ray-based methods have the highest efficiency when it comes to memory usage in the application of tomographic problems.
3.4.2 Analytical ray-tracing in Delaunay triangulation

Despite the fact that limited bandwidth methods have the highest accuracy and grid-based methods are computationally quick, the ray-based method has been chosen for the forward modeling of focusing operators. There are two reasons for this choice:

[1] The ray-based method has the highest memory efficiency, which is important for large 3D problems.

Figure 3.8: Comparison of the different methods for forward modeling of traveltime maps of the Marmousi model (see section 5.3) generated by a) the ray-based method, b) the grid-based method, and c) the limited bandwidth method [after Nichols, 1996].
3.4 Forward modeling of traveltimes

[2] The relation between traveltime and model parameters are automatically generated during ray tracing, as the ray-path is calculated. This makes the method quite applicable to tomographic problems.

Even though the calculation of amplitudes and multi-valued arrivals is also an advantage of this method, they will not be used in this research. Only the traveltimes of the first arrivals will be inverted.

In the ray-based method, an imaginary shot (focal point) is positioned in the subsurface model, and rays are traced towards the surface. Together, the traveltimes of these rays form the traveltimes of the modeled focusing operator for this particular focal point (Figure 3.9).

The forward modeling of focusing operators should be carried out in a model that is parameterized in the way presented in the previous section. As a ray-based method is used, this can be done analytically. The procedure of analytical ray-tracing within Delaunay triangulation is explained in this section. This is done for 3D problems, but can easily be redefined for problems in 2D.

![Figure 3.9](image)

Figure 3.9: An imaginary shot (focal point) is positioned in the subsurface and rays are traced towards the surface

The theory of the ray-based method is described by Červený [2001]. It makes use of the eikonal equation, which is a high-frequency result derived from the wave equation (see appendix B.1):

$$\nabla \tau (x) \cdot \nabla \tau (x) = s^2 (x),$$

where $\tau (x)$ is the eikonal function, $s(x)$ is the slowness, and the location in the Cartesian coordinate system is defined by $x$ ($x_i$, with $i = 1, 2, 3$). The equation $\tau (x) = t$ represents a wavefront of the wavefield at time $t$. In other words, the eikonal equation relates wavefronts to a seismic velocity distribution, as the constant traveltime surfaces in the traveltimes field correspond to the wavefronts of a propagating wave.
The eikonal equation is usually solved in terms of ray-paths. A ray-path in a wave-field is defined by the curve for which $\nabla \tau$ is tangent in each point. Hence, the rays are perpendicular to the wavefront and parallel to the slowness vector $p_i$ (with $i = 1, 2, 3$), so $\nabla \tau(x) = p_i$ is the slowness vector. The ray-paths can be described by an independent ray-trace parameter along the path. There are several ways to define this parameter, e.g., the arc-length can be used. The traveltime from point $A$ to point $B$ along a ray-path with arc-length $l$ can be written as:

$$
\tau_{AB} = \int_{A}^{B} s \, dl.
$$

(3.4.3)

However, the simplest ray-tracing system is obtained using the 'natural variable along the ray'. This variable is denoted as $\sigma$ and its dimension is $[m^2/s]$. This parameter is related to the arc-length $l$ and traveltime $t$ along the ray by:

$$
d\sigma = \frac{1}{q} \, dl = \frac{1}{q} \, dt,
$$

(3.4.4)

where $q$ is the quadratic slowness ($s^2$), which is directly connected with $\sigma$ by the traveltime. Thus, to derive a simple form of the ray-tracing system in which the 'natural variable along the ray' can be used, the velocity model is described by the quadratic slowness. For sake of simplicity, the quadratic slowness is denoted by $q$ in the remainder of this thesis. Using equation (3.4.4), the traveltime from point $A$ to point $B$ along a ray-path with ray-trace parameter $\sigma$ can be written as:

$$
\tau_{AB} = \int_{A}^{B} q \, d\sigma.
$$

(3.4.5)

As a result, the ray-trace parameter $\sigma$ defines the ray-tracing system that describes the ray-path:

$$
\frac{dx_i(\sigma)}{d\sigma} = p_i(\sigma), \quad \frac{dp_i(\sigma)}{d\sigma} = \frac{1}{2} \nabla q, \quad \frac{dt(\sigma)}{d\sigma} = q.
$$

(3.4.6)

A derivation of this system is given in appendix B.2. By means of this system the variation of $x$, $p$, and $t$ along the ray can be calculated. In order to solve the ray tracing system, initial conditions should be defined. The initial conditions for a single ray are defined in the coordinates of the initial point $(\sigma[0])$ and can be written as:

$$
x_i[0] = x_i(\sigma[0]), \quad p_i[0] = p_i(\sigma[0]), \quad t[0] = t(\sigma[0]) = 0,
$$

(3.4.7)

for $i = 1, 2, 3$.

Assume that the quadratic slowness $q$ is a linear function of the coordinates $x$:

$$
q = q[0] + \sum_i (\nabla q)_i (x_i - x_i[0]),
$$

(3.4.8)
where \((\nabla q)_i\) is the quadratic slowness gradient in direction \(x_i\). The ray tracing system (3.4.6) can be solved by means of (3.4.7) and (3.4.8):

\[
x_i(\sigma) = x_i^{[0]} + p_i^{[0]} \Delta \sigma + \frac{1}{4}(\nabla q)_i \Delta \sigma^2,
\]

\[
p_i(\sigma) = p_i^{[0]} + \frac{1}{2}(\nabla q)_i \Delta \sigma,
\]

\[
t(\sigma) = t^{[0]} + q^{[0]} \Delta \sigma + \sum_i \left\{ (\nabla q)_i x_i^{[0]} \Delta \sigma + \frac{1}{2}(\nabla q)_i p_i^{[0]} \Delta \sigma^2 + \frac{1}{12}(\nabla q)_i (\nabla q)_i \Delta \sigma^3 \right\},
\]

for \(i = 1, 2, 3\)

where \(\Delta \sigma = \sigma - \sigma^{[0]}\).

This is an exact analytic solution representing the ray-path, the ray-parameter, and the traveltime. How this solution can be used for a parameterization by Delaunay triangulation is explained in the next part.

**Ray-tracing within a tetrahedron**

In the forward modeling of focusing operators a focal point is positioned in the subsurface model and rays are traced towards the surface. Assume that the initial focal point of a ray is situated within a tetrahedron. The ray will travel through the tetrahedron and end at a facet of the tetrahedron (see Figure 3.10). The equation of the facet is:

\[
\sum_i n_i (x_i - x_i^{[0]}) = 0,
\]

where \(x\) is the point where the ray intersects the facet, \(x^{[0]}\) is an arbitrary point on the facet, and \(n\) is the normal to the facet. By substituting \(x_i\), which is defined in (3.4.9a), in (3.4.10), the location of incidence of the ray at the facet \(x_i\) can be calculated:

\[
\sum_i n_i \left( x_i^{[0]} + p_i^{[0]} \Delta \sigma + \frac{1}{4}(\nabla q)_i \Delta \sigma^2 - x_i^{[0]} \right) = 0.
\]

This is a quadratic equation in which \(\Delta \sigma\) can be computed. Once \(\Delta \sigma\) is calculated, it can be substituted in (3.4.9) to obtain the location of incidence of the ray at the facet \(x(\sigma)\), and the ray parameter \(p(\sigma)\), and traveltime \(t(\sigma)\) at this location.

**Ray-tracing across a facet**

Once the ray has been calculated through the tetrahedron, it will arrive at one of the facets of the tetrahedron. How the ray behaves at the interface can be determined
by the laws of reflection and transmission. Using Snell’s law it can be derived that (see appendix B.2):

\[
p^{[2]}_i = p^{[1]}_i - \left\{ (p^{[1]}_i \cdot \mathbf{n}) \pm \text{sign}(p^{[1]}_i \cdot \mathbf{n}) \sqrt{q^{[2]} - q^{[1]} + (p^{[1]}_i \cdot \mathbf{n})^2} \right\} n_i,
\]

for \(i = 1, 2, 3\),

(3.4.12)

where \(p^{\text{new}}\) and \(q^{\text{new}}\) are the ray parameter and the quadratic slowness on the other side of the facet. If \(\text{sign}(\mathbf{p} \cdot \mathbf{n})\) equals 1 the incident wave makes an acute angle with \(\mathbf{n}\), and if it equals -1 the angle is obtuse. The upper sign applies to reflected waves, the lower sign to the transmitted waves. In the modeling of focusing operators only transmission is used. The new ray parameter \((p^{\text{new}})\) is needed to start the ray-tracing in the tetrahedron at the other side of the facet. The location \(\mathbf{x}\) at the facet and the traveltine \(t\) stay the same. The ray arrives in the new tetrahedron and the ray can again be calculated within this tetrahedron. The two steps of ray-tracing within the tetrahedron and across a facet are repeated until the outer boundary of the model is reached.

**Reaching the receivers**

The ray-based method that is presented here, sprays rays from a buried shot (focal point) into the model. Therefore, it is also referred to as a 'shooting method'. The method does not define where the rays end, because this depends on the path through
3.4 Forward modeling of traveltimes

the quadratic slowness model (Figure 3.11). As it is desired to know the ray-path from the source to a predefined receiver, it is necessary to develop an interpolation algorithm. Vinje et al. [1993] described a method to calculate the wavefront that corresponds to the calculated ray-paths. In this way the traveltimes are known at any location at the surface, consequently also at the receiver locations. Nevertheless, the ray-path geometries to the receivers, which are useful for the tomographic inversion, are still unknown.

![2D](image1.png) ![3D](image2.png)

**Figure 3.11:** a) A ray-fan with arbitrary arrivals at the surface. b) Rays that arrive at receivers (triangles) at the surface.

One way to obtain the ray-paths is by re-shooting of rays until the receiver is hit. This seems to be an inefficient method, however, when it is done in a smart way the efficiency of the method can be increased. Figures 3.11 and 3.12 explain the procedure of re-shooting. First rays are calculated for increasing takeoff angles, resulting in an initial ray-fan. The initial rays that have similar takeoff angles are connected, resulting in an arrival grid represented by lines in 2D, or triangles in 3D. Next, the grid is used to find the takeoff angle for the ray that hits the receiver. The rays that enclose the receiver are used to calculate the new takeoff angle. In 2D the distances between the two initial rays and the receiver, $d_1$ and $d_2$ (Figure 3.12b),
are used for the new takeoff angle $\theta_{rec}$:

$$
\theta_{rec} = \frac{d_1 \theta_1 + d_2 \theta_2}{d_1 + d_2},
$$

(3.4.13)

where $\theta_1$ and $\theta_2$ are the takeoff angles of the initial rays. In 3D, the areas formed by the three surrounding arrivals and the receiver, $a_1$, $a_2$, and $a_3$, are used for the new takeoff angle $\theta_{rec}$:

$$
\theta_{rec} = \frac{a_1 \theta_1 + a_2 \theta_2 + a_3 \theta_3}{a_1 + a_2 + a_3},
$$

(3.4.14)

and for the new takeoff azimuth $\phi_{rec}$:

$$
\phi_{rec} = \frac{a_1 \phi_1 + a_2 \phi_2 + a_3 \phi_3}{a_1 + a_2 + a_3}.
$$

(3.4.15)
In Figure 3.12b the definition of the areas \( a_1, a_2, \) and \( a_3 \) are shown. These new takeoff angles (and azimuths) are used to calculate the new ray, which can approach the receiver or even hit the receiver. When the receiver is only approached but is not exactly hit, the new ray can be used for further interpolation, replacing one of the initial rays. Once the receiver is hit, our goal is reached: both the traveltime at the receiver and the ray-path from the focal point to the receiver are known. Note that it is possible that there are more ray-paths that connect one source to one receiver. In this case, only the first arrivals are considered in this research.

### 3.4.3 Determination of the optimization matrix

Now that the ray-paths of the modeled focusing operator are known, they can be used to describe the relation between the traveltime and the model parameters. This is needed to update the model parameters during the optimization step. In tomographic inversion of focusing operators both the updating of the quadratic slowness and the focal point locations should be incorporated in the optimization algorithm. The matrix equation that should be solved to obtain the model parameter updates was introduced in equation (3.2.1). The complete matrix equation is:

\[
\begin{bmatrix}
\Delta t^{[1]} \\
\Delta t^{[2]} \\
\Delta t^{[3]} \\
\vdots \\
\Delta t^{[N]}
\end{bmatrix} = \\
\begin{bmatrix}
\frac{\partial g^{[1]}}{\partial x_1} & \cdots & \frac{\partial g^{[1]}}{\partial x_2} & \cdots & \frac{\partial g^{[1]}}{\partial x_3} \\
\frac{\partial g^{[2]}}{\partial x_1} & \cdots & \frac{\partial g^{[2]}}{\partial x_2} & \cdots & \frac{\partial g^{[2]}}{\partial x_3} \\
\vdots & \cdots & \vdots & \cdots & \vdots \\
\frac{\partial g^{[N]}}{\partial x_1} & \cdots & \frac{\partial g^{[N]}}{\partial x_2} & \cdots & \frac{\partial g^{[N]}}{\partial x_3}
\end{bmatrix}
\begin{bmatrix}
\Delta q^{[1]} \\
\Delta q^{[2]} \\
\Delta q^{[3]} \\
\vdots \\
\Delta q^{[F]}
\end{bmatrix}.
\]

\( \Delta t^{[n]} \) is the traveltime difference for ray \( n \), \( \Delta q^{[i]} \) is the quadratic slowness update of vertex \( i \), and \( x_i^{[f]} \) with \( i = 1, 2, 3 \) is the location update of focal point \( f \). If there are \( N \) data points, \( L \) quadratic slowness parameters, and \( F \) focal points, then the number of model parameters is \( M = L + 3F \) in 3D problem, and \( M = L + 2F \) in 2D. The \( N \) data points and the \( M \) model parameters are related by a \( N \times M \) optimization matrix \( A \), that describes the relation between the traveltimes and the model parameters. This matrix thus contains the partial derivatives of the \( N \) traveltimes with respect to the \( M \) model parameters.
Partial derivatives with respect to quadratic slowness

As the forward modeling can be done analytically, the equations needed for the inverse problem can also be solved in an analytical way [Trinks, 2001]. The inverse problem requires that the partial derivatives of the traveltime with respect to the model parameters are known. The traveltime within a tetrahedron is described by (3.4.9c):

\[
 t(\sigma) = t[0] + q[0] \Delta \sigma + \sum_i \left\{ (\nabla q)_i x_i[0] \Delta \sigma + \frac{1}{2} (\nabla q)_i p_i[0] \Delta \sigma^2 + \frac{1}{12} (\nabla q)_i (\nabla q)_i(\nabla q)_i \Delta \sigma^3 \right\}. 
\]

The model parameters are defined at the vertices of the tetrahedra. One tetrahedron is described by four model parameters, so the partial derivatives are calculated with respect to these four vertices. The partial derivative of the traveltime with respect to the quadratic slowness of the \( k^{th} \) vertex \( (k = 1, 2, 3, 4) \) of a tetrahedron is:

\[
 \frac{\partial t(\sigma)}{\partial q[k]} = \frac{\partial t[0]}{\partial q[k]} + \frac{\partial q[0]}{\partial q[k]} \Delta \sigma + \left\{ \sum_i \frac{\partial (\nabla q)_i}{\partial q[k]} x_i[0] \Delta \sigma + \sum_i \frac{1}{2} \frac{\partial (\nabla q)_i}{\partial q[k]} p_i[0] \Delta \sigma^2 + \sum_i \frac{1}{6} \frac{\partial (\nabla q)_i}{\partial q[k]} (\nabla q)_i(\nabla q)_i \Delta \sigma^3 \right\}. 
\]

(3.4.17)

The initial conditions \( x_i[0], p_i[0], \) and \( t[0] \) are obtained during ray-tracing, so only \( \frac{\partial (\nabla q)_i}{\partial q[k]} \) and \( \frac{\partial q[0]}{\partial q[k]} \) have to be calculated to obtain the partial derivative with respect to quadratic slowness. To obtain \( \frac{\partial (\nabla q)_i}{\partial q[k]} \), \( (\nabla q)_i \) is required, which can be calculated by means of:

\[
 \begin{bmatrix}
 x_i[2] - x_i[1] \\
 x_i[3] - x_i[1] \\
 x_i[4] - x_i[1]
\end{bmatrix}
\begin{bmatrix}
 x_i[2] - x_i[1] \\
 x_i[3] - x_i[1] \\
 x_i[4] - x_i[1]
\end{bmatrix}^{-1} 
(\nabla q)_i = 
\begin{bmatrix}
\end{bmatrix}, 
\]

(3.4.18)

in which the \( x_i[k] \) is the location of vertex \( k \) in direction \( i \), and \( q[k] \) is the quadratic slowness of this vertex. By rewriting the system, the linear equations for \( (\nabla q)_i \) can be obtained. In these equations, the partial derivatives of \( (\nabla q)_i \), with respect to the quadratic slowness of the vertices \( q[k] \) \( (k = 1, 2, 3, 4) \) can be calculated easily (see appendix C).

To obtain \( \frac{\partial q[0]}{\partial q[i]} \), \( q[0] \) is required, which is defined by (3.4.8):

\[
 q[0] = q[k] - \sum_i (\nabla q)_i (x_i[k] - x_i[0]). 
\]

(3.4.19)

So, the partial derivatives of \( q[0] \) with respect to the quadratic slowness of the vertices
3.4 Forward modeling of traveltimes

$q^{[k]}$ ($k = 1, 2, 3, 4$) can be calculated by means of the partial derivatives \( \frac{\partial (\nabla q)}{\partial q^{[k]}} \):

\[
\frac{\partial q^{[0]}}{\partial q^{[k]}} = 1 - \sum_{i} \frac{\partial (\nabla q)_{i}}{\partial q^{[k]}} (x_{i}^{[k]} - x_{i}^{[0]}).
\]  
(3.4.20)

Now that \( \frac{\partial (\nabla q)}{\partial q^{[k]}} \) and \( \frac{\partial q^{[0]}}{\partial q^{[k]}} \) are calculated for vertex \( k \), they can be inserted in equation (3.4.17), together with \( \Delta \sigma \) and the initial conditions. In this way, the partial derivative of travel time with respect to the quadratic slowness of vertex \( k \) is obtained analytically. This can be done for all vertices of the tetrahedron and for all tetrahedra the ray is crossing. As a result, the complete relation between the traveltime of the ray and the quadratic slownesses in the model is described.

**Partial derivatives with respect to focal point location**

Matrix \( A \) also contains partial derivatives of the traveltimes with respect to the location of the focal points. These partial derivatives are related to the propagation direction at the focal points. This can easily be shown using equation (3.4.5). From this equation we can derive that:

\[
dt(\sigma) = \int_{\sigma_{0}}^{\sigma_{1}} q d\sigma, \quad dt(\sigma) = \sum_{i=1}^{3} \int_{x_{i}}^{x_{f}} p_{i}(\sigma) dx_{i}(\sigma),
\]  
(3.4.21)

where \( x_{i} \) is the position along the ray. From this equation it follows that \( p_{i} \) gives the partial derivative \( \frac{\partial t}{\partial x_{i}} \). Since \( p_{i} \) is calculated during ray-tracing, these partial derivatives are directly available. In 2D situations they are written as:

\[
\frac{\partial t}{\partial x_{1}} = s \sin \theta, \quad \frac{\partial t}{\partial x_{2}} = s \cos \theta,
\]  
(3.4.22)

where \( \theta \) is the takeoff angle. In 3D problems the takeoff azimuth \( \phi \) is also relevant:

\[
\frac{\partial t}{\partial x_{1}} = s \sin \theta \sin \phi, \quad \frac{\partial t}{\partial x_{2}} = s \sin \theta \cos \phi, \quad \frac{\partial t}{\partial x_{3}} = s \cos \theta.
\]  
(3.4.23)

This can also be illustrated (in a 2D example) by means of Figure 3.13. When an imaginary perturbation \( \Delta x_{1} \) is considered, the imaginary perturbed ray can be ‘projected’ on the initial ray. This results in a right angled triangle defined by the takeoff angle \( \theta \) of the ray. In this triangle \( \Delta x_{1} \) forms the hypotenuse, and the difference between the rays is related to the opposite side. To be more precise, the length of the opposite side can be defined by \( \frac{\Delta t}{s} \), in which \( s \) is the local slowness. Using this, the partial derivatives of traveltime with respect to focal point can be calculated by:

\[
\frac{\partial t}{\partial x_{1}} = \frac{\Delta t}{\Delta x_{1}} = s \sin \theta, \quad \frac{\partial t}{\partial x_{2}} = \frac{\Delta t}{\Delta x_{2}} = s \cos \theta,
\]
which corresponds to (3.4.22).

Thus, the forward modeling stage consists of determining the ray-paths and traveltimes from each focal point to all receiver positions, and determining the partial derivatives of the traveltime with respect to the quadratic slowness and the focal point coordinates. Next, the updates of the model parameters can be calculated.
3.5 Optimization

The optimization is the third main step in tomographic inversion of focusing operators (see Figure 3.1). Generally, inversion and optimization are not always clearly distinguished. The inversion theory is the combination of statistical and physical theories that may lead to a function that has to be minimized. This is obtained by the complete scheme as presented in Figure 3.1. Optimization is the mathematical or numerical technique of the actual determination of that minimum, which is only one step in the complete inversion scheme.

In this optimization step, the updates of the model parameters are calculated by minimizing the difference between the travel times of the forward modeled and those of the observed focusing operators. The updates of the model parameters are added to the initial model, resulting in an improved model of the subsurface.

3.5.1 Overview of existing methods

Local versus Global inversion methods

The solution space of an optimization problem may have both local and global minima. The global minimum is the desired solution of the problem. There are inversion methods that look for the nearest (local) minimum, and methods that look for the absolute (global) minimum. Methods that will find a local minimum are usually much more efficient. Nevertheless, as they will find a local minimum, a good initial model is necessary to obtain a correct solution (i.e. the local minimum that is found equals the desired global minimum). Methods that will find a global minimum are not very sensitive to the choice of the initial model, because they are able to escape from local minima. However, they are computationally expensive [Rothman, 1985; Mosegaard and Vestergaard, 1991; Sambridge and Drijkoningen, 1992]. Besides, experience shows that local minima are no major problem in traveltime inversion [van der Made, 1988]. Therefore, a local inversion method will be used for the tomographic inversion of focusing operators. An attempt will be made to reduce the risk of ending up in a possible local minimum by making use of a dynamic parameterization, as will be explained in section 3.6.

Local inversion methods

Local inversion methods look for a local minimum in the solution space. Generally, the gradient information at the initial position in the solution space defines the direction in which the minimum can be found. Note that taking the gradient of the objective function with respect to the parameters as a direction of the solution is inherent to the linearization of the problem (section 3.2). So, in the local inversion methods the relation between the traveltime data and the model parameters is
assumed to be linear (equation (3.2.1)):  

\[ \Delta d = A \Delta m, \]  

(3.5.24)

where \( \Delta d \) contains the difference between the traveltimes of the observed and the modeled focusing operator, \( \Delta m \) contains the update in model parameters, and \( A \) describes the linear relation between these traveltimes and the model parameters (see section 3.4.3). In practice, it appears to be very difficult to solve for the model vector \( \Delta m \) in terms of the data vector \( \Delta d \), because matrix \( A \) generally is not invertible. The problem is that \( \Delta d \) has size \( N \), while \( \Delta m \) has size \( M \), which makes \( A \) a non-square \( N \times M \) matrix, that does not have an inverse. However, this problem can be solved by using the generalized inverse, which is denoted by a dagger:

\[ \Delta m = A^\dagger \Delta d. \]  

(3.5.25)

The generalized inverse can be obtained by explicit matrix calculation methods.

**Explicit optimization methods**

For small scale problems, generalized inverses can be calculated by explicit matrix calculation methods like:

- Least squares inversion;
- Singular-Value Decomposition (SVD).

Typically, a tomographic system will be over-determined, as the size of the data vector \( N \) is much larger than the size of the model vector \( M \). When the least squares method is used in such an over-determined system, an update of \( \Delta m \) is calculated by \( \Delta m = (A^T A)^{-1} A^T \Delta d \) [Jackson, 1972]. So, in the least squares method the generalized inverse is represented by:

\[ A^\dagger = (A^T A)^{-1} A^T. \]  

(3.5.26)

The size of \( A \) is \( N \times M \), so the size of the generalized inverse \( A^\dagger \) is \( M \times N \). Note that the size of matrix to be inverted \( A^T A \) is only \( M \times M \).

In the case of applying SVD, the matrix \( A \) will be decomposed into \( A = U \Lambda V^T \), where \( U \) and \( V \) are the \( (N \times N) \) and \( (M \times M) \) orthonormal matrices respectively. \( U \) contains the eigenvectors of the data space, \( V \) contains the eigenvectors of the model space, and \( \Lambda \) is a \( (N \times M) \) matrix containing the singular-values in decreasing order in the matrix elements \( \Lambda_{ii} \). The generalized inverse can be constructed by using only the \( k \) largest singular-values:

\[ A^\dagger = V^{[k]} \Lambda^{[k-1]} U^{[k]} \]  

(3.5.27)
where $\Lambda^{[k]}$ is a square diagonal matrix, containing only the $k$ largest singular-values and $U^{[k]}$ and $V^{[k]}$ contain the corresponding columns of $U$ and $V$.

The relation between the least squared method and the SVD is discussed by Menke [1984]. Both methods are very well known and are feasible for small systems. However, for larger problems an explicit matrix calculation in least squares inversion or in SVD becomes impractical. Several iterative optimization methods for approximation of the generalized inverse have been developed to circumvent this problem.

### Iterative optimization methods

Iterative optimization methods are used when the number of parameters is large. These methods approximate the generalized inverse iteratively, which results in a minimal CPU memory requirement. In addition, they are very suitable for cases where matrix $\mathbf{A}$ is sparse. This is generally the case in seismic inverse problems, since each seismic ray passes through only part of the total model parameters, which causes most elements in $\mathbf{A}$ to be zero.

Iterative optimization methods have been realized with techniques such as Conjugate Gradients (CG) [Hestenes and Stiefel, 1952; Scales, 1987] and Large Sparse QR factorization (LSQR) [Paige and Saunders, 1982]. These are projection methods in which a model update is calculated, which minimizes the squared residual error of the data. To find the exact minimum, it is required that the number of steps equals the number of model parameters. This is similar to an explicit optimization method, and therefore, in large scale problems, the computational cost makes an exact solution infeasible. Nevertheless, iterative methods allow to get a reasonable estimate of the solution in a reasonable limited amount of iterations. Berryman [2001a] provides a complete overview of the iterative methods and an analysis of their capacities.

#### 3.5.2 Optimization by the LSQR algorithm

In tomographic inversion of focusing operators an iterative optimization method is used, because it reduces the memory storage and it is suitable for sparse matrices, which is typically the case in tomographic problems. The LSQR algorithm is algebraically equivalent to applying the CG method to the equation $\Delta \mathbf{m} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \Delta \mathbf{d}$, but it has better numerical properties, especially if $\mathbf{A}$ is ill-conditioned. Therefore, the LSQR algorithm is used.

The input of the LSQR algorithm are vector $\Delta \mathbf{d}$ and matrix $\mathbf{A}$. The steps applied in the algorithm are as follows:
Initialize:

\[
\hat{u}^{[1]} = \Delta d; \quad \beta = ||\hat{u}^{[1]}||; \quad \hat{u}^{[1]} = \frac{\hat{u}^{[1]}}{\beta}
\]

\[
\hat{v}^{[1]} = A^T \hat{u}^{[1]}; \quad \alpha = ||\hat{v}^{[1]}||; \quad \hat{v}^{[1]} = \frac{\hat{v}^{[1]}}{\alpha}
\]

Iterate for \(i=1,k\):

\[
\hat{u}^{[i+1]} = A \hat{v}^{[i]} - \alpha \hat{u}^{[i]}; \quad \beta = ||\hat{u}^{[i+1]}||; \quad \hat{u}^{[i+1]} = \frac{\hat{u}^{[i+1]}}{\beta}
\]

\[
\hat{v}^{[i+1]} = A^T \hat{u}^{[i+1]} - \beta \hat{v}^{[i]}; \quad \alpha = ||\hat{v}^{[i+1]}||; \quad \hat{v}^{[i+1]} = \frac{\hat{v}^{[i+1]}}{\alpha}
\]

In \(k\) iterations, this algorithm produces \(k\) vectors \(\hat{u}\) and \(\hat{v}\), and \(k\) values of \(\alpha\) and \(\beta\). The \(\hat{u}\) and \(\hat{v}\) vectors are called basis-vectors. Note that the subspace described by the \(k\) vectors is also called a Krylov subspace [Berryman, 2001b]. The basis-vectors can be used to construct matrices \(\hat{U}^{[k]}\) and \(\hat{V}^{[k]}\) by:

\[
\hat{U}^{[k]} = (\hat{u}^{[1]}, \hat{u}^{[2]}, ..., \hat{u}^{[k]})
\]

\[
\hat{V}^{[k]} = (\hat{v}^{[1]}, \hat{v}^{[2]}, ..., \hat{v}^{[k]})
\]

These matrices correspond to the eigenvector matrices \(U^{[k]}\) and \(V^{[k]}\) obtained by SVD. The values \(\alpha\) and \(\beta\) obtained by LSQR can be organized in a matrix structure (Figure 3.14). The obtained matrix plays a similar role as the singular-value matrix \(A^{[k]}\) in SVD:

\[
B^{[k]} = \begin{bmatrix}
\alpha^{[1]} & 0 & \cdots & 0 \\
\beta^{[1]} & \alpha^{[2]} & \cdots & 0 \\
0 & \beta^{[2]} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \alpha^{[k]} \\
0 & 0 & 0 & \beta^{[k]}
\end{bmatrix}
\]

It should be noted that this is a bi-diagonal matrix and in SVD the singular-value matrix is a diagonal matrix. The exact relation between the two matrices will be discussed in the next section.
Figure 3.14: Matrix $A$ of an over-determined system; 88 data points ($N$) and 31 model parameters ($M$). Applying the LSQR algorithm to matrix $A$ results in the matrix system $U^{[k]}B^{[k]}V^{[k]T}$. Matrix $B^{[k]}$ containing the bi-diagonal values after 31 LSQR iterations. Applying SVD to matrix $B^{[k]}$ results in $U_B\Lambda_BV_B^T$, where $\Lambda_B$ contains the singular-values of matrix $A$ when $M = 31$ iterations are performed.

Similarities to SVD

The LSQR algorithm bears some very useful similarities to the generalized inverse method based on SVD [Zhang and McMechan, 1995; Yao et al., 1999]. This will be demonstrated by means of an optimization problem of a model defined by $M = 31$ model parameters and $N = 88$ data. The optimization matrix $A$ of this system is shown in Figure 3.14. In the LSQR algorithm the matrices $U^{[k]}$, $V^{[k]}$, and $B^{[k]}$ are obtained after $k$ iterations. These can be used, in the same way as they are used in SVD to calculate the approximate generalized inverse $\tilde{A}^\dagger$:

$$\tilde{A}^\dagger = \hat{V}^{[k]}B^{[k-1]}\hat{U}^{[k]T}. \quad (3.5.28)$$
As mentioned before, the matrix $\mathbf{B}^{[k]}$ is a bi-diagonal matrix. To make a better comparison between the singular-values obtained by SVD and the bi-diagonal values obtained by LSQR, matrix $\mathbf{B}^{[k]}$ can be diagonalized (Figure 3.14). This is performed by applying SVD to matrix $\mathbf{B}^{[k]}$:

$$
\mathbf{B}^{[k]} = \mathbf{U}_B \mathbf{A}_B \mathbf{V}_B^T.
$$

(3.5.29)

$\mathbf{U}_B$ is a $(k+1 \times k+1)$ and $\mathbf{V}_B$ is a $(k \times k)$ matrix, both containing the eigenvectors belonging to bi-diagonal value matrix $\mathbf{B}^{[k]}$. $\mathbf{A}_B$ is the diagonal matrix containing the real singular-values of this bi-diagonal matrix. $\mathbf{U}_B \mathbf{U}_B^T$ and $\mathbf{V}_B \mathbf{V}_B^T$ will always equal the identity matrix, as matrix $\mathbf{B}^{[k]}$ is of full rank. Generally, the SVD procedure applied to matrix $\mathbf{B}^{[k]}$ is feasible, because it is only a $(k+1 \times k)$ matrix in which $k$ is the number of LSQR iterations. This procedure makes a qualitative comparison between the SVD and the LSQR algorithm possible. When we substitute (3.5.29) into (3.5.28) we obtain:

$$
\hat{\mathbf{A}}^\dagger = \left(\hat{\mathbf{V}} \mathbf{V}_B\right) \mathbf{A}_B^{-1} \left(\hat{\mathbf{U}} \mathbf{U}_B\right)^T,
$$

(3.5.30)

which can be directly compared to result of direct SVD (equation (3.5.27)):

$$
\mathbf{A}^\dagger = \mathbf{V} \mathbf{A}^{-1} \mathbf{U}^T.
$$

When the number of basis-vectors (i.e. the number of iterations $k$) equals the number of model parameters, the result of the LSQR algorithm is equivalent to the result of SVD. In this case, the singular-values in matrix $\mathbf{A}_B$ equal the singular-values in $\mathbf{A}$. However, when the number of basis-vectors is smaller than the number of model parameters, the vectors obtained do not completely span the range of the model space [Deal and Nolet, 1996]. In this case, the singular-values in $\mathbf{A}_B$ are called the ‘pseudo’ singular-values (also known as Ritz values (Chapter 3 in Nolet [1987])). In the first $k$ iterations, only the $k$ largest basis-vectors are resolved, and only $k$ ‘pseudo’ singular-values are obtained (Figure 3.15). When the number of iterations increases, the ‘pseudo’ singular-values converge to the complete set of singular-values obtained by SVD. The ‘pseudo’ singular-values, obtained by LSQR, can be used quantitatively to evaluate how well the system is determined.

**Re-orthogonalization**

In theory, the LSQR algorithm produces a sequence of orthogonal basis-vectors, but in practice the orthogonality may break down after several iterations when the calculations are performed at a finite precision. Normally, after a sufficient number of iterations, ‘pseudo’ singular-values should converge to singular-values. However, as the number of iterations becomes large the basis-vectors are no longer orthogonal, and some duplicate ‘pseudo’ singular-values and basis-vectors are generated.
Therefore, it may be necessary to perform re-orthogonalization. In each iteration basis-vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ can be re-orthogonalized by:

\[
\text{for } i = 1, k : \quad \hat{\mathbf{u}}^r = \hat{\mathbf{u}} - (\hat{\mathbf{U}}^T_i \cdot \hat{\mathbf{u}}) \hat{\mathbf{U}}_i \\
\hat{\mathbf{v}}^r = \hat{\mathbf{v}} - (\hat{\mathbf{V}}^T_i \cdot \hat{\mathbf{v}}) \hat{\mathbf{V}}_i
\]

\[\text{(3.5.31)}\]

Of course this procedure makes the algorithm more expensive. However, it has been shown by Berryman [2001b] that re-orthogonalization during the first iterations is most effective. As re-orthogonalization is not very expensive during the first iterations, it is worthwhile applying it.

**Updates versus Iterations**

In the tomographic inversion theory as presented in this chapter, both the word 'update' and the word 'iteration' are used. These words are related to the 'inversion' and 'optimization' respectively. An update is one entire loop in the inversion scheme as presented in Figure 3.1. An iteration is one calculation step in the iterative optimization method. Although the literal meaning of both words is quite similar, the usage in this thesis indicates two different procedures.
3.6 Regularization of the inversion

Now that the first three main steps, parameterization, forward modeling, and optimization have been discussed, we can continue with the final fourth step which is a very important part of the tomographic inversion of focusing operator method. Regularization is required because, typically, geophysical inverse problems are mixed-determined. This means that, even thought the number of model parameters is less than the available data, a lot of parameters are under-determined. The mixed-determined nature of the inversion should be faced in order to obtain stable and accurate results. This can be performed by regularization the inverse problem. Regularization is a very important part of tomographic inversion of focusing operators because this step will not only attempt to generate stable and accurate results, but also will make the method data-driven as will be shown in section 3.6.2.

3.6.1 Overview of existing methods

In tomographic inversion the model parameters are determined by the seismic rays that pass through the subsurface, which is defined by these parameters. When a lot of rays sense a parameter, this results in many non-zero elements in matrix $A$, and as a consequence this parameter is well-determined. On the other hand, when only a few rays sense a parameter, it will be under-determined. In the least squares inversion method, this parameter will cause large elements in matrix $(A^T A)^{-1}$ (see section 3.5.1), which may have dramatic effects on $\Delta m$. In case of using the SVD method, this parameter will result in large elements in $\Lambda^{-1}$, which has the same effect. Therefore, mixed-determined inverse problems should be regularized. Regularization methods can be divided into two groups:

[1] Regularization of the optimization;


Both methods will be highlighted in this subsection.

Regularization of the optimization

Regularization of the optimization can be performed by several methods. Three of these methods are discussed here: 1) damping of the least squares method, 2) truncation of the SVD, 3) regularization of the model parameters.

Damping of the least squares method can be performed by extending the system $A \Delta m = \Delta d$ (equation (3.5.24)) by another set of equations that force the update in the model parameters $\Delta m$ to be as small as possible. As a result, the following
3.6 Regularization of the inversion

The equation should be solved:

\[
\begin{bmatrix}
A \\
\lambda I
\end{bmatrix} \Delta m = \begin{bmatrix}
\Delta d \\
0
\end{bmatrix}.
\]  

(3.6.32)

This results in damping of the inversion, in which \( \lambda \) is called the damping factor. The model parameter updates are obtained by the generalized inverse \((A^T A + \lambda I)^{-1} A^T\), in which the elements of \((A^T A + \lambda I)^{-1}\) are always smaller than \( \frac{1}{\lambda} \). This circumvents the problem of large values in the inversion system that make the inversion unstable. Note that applying SVD to this system results in the same damping. In the generalized inverse \(V[k](\Lambda[k])^{-1}U[k]^T\) the elements in \( \Lambda[k] \) are always smaller than \( \frac{1}{\lambda} \). Another way of regularizing the optimization is by truncation of the SVD. This is performed by considering the singular-values in matrix \( \Lambda \). Only the \( k \) singular-values above a certain threshold should be incorporated in the calculation of the model parameter updates. Again, \((\Lambda[k])^{-1}\) cannot contain large values, which results in a stable inversion. Note that also a limited number of LSQR iterations regularize the optimization automatically, because the LSQR algorithm finds the largest basis-vectors first. Regularizing the optimization by damping and truncation of the SVD makes the inversion more stable. However, the accuracy of the result is not ensured, as the minimum might not be well-shaped, which results in an interaction between parameters. Regularization of the model parameters can also be performed in a more intelligent way in which an attempt is made to obtain a well-shaped minimum. An attractive approach is to add an additional model regularization term to the objective function. In theory, this regularization term should be the inverse model covariance [Tarantola, 1987] obtained from a-priori information sources. Another way of model regularization can be achieved by adding a term encouraging the velocity gradient to follow the reflector position [Käipio et al., 1999; Clapp, 2001]. Although regularization of the optimization is powerful and generally necessary, it has a drawback: the problem of over-parameterization is still not solved. This can cause problems when the inverse problem is regularized; the model parameters need different levels of regularization, as certain regions in the model are more under-determined than other regions. In other words, a well-shaped minimum might still not be accomplished.

Modification of the parameterization

By modifying the parameterization an attempt is made to avoid over parameterization. The parameterization can be modified by the user. This can be achieved by using a global parameterization in which layers are defined [van der Made, 1988; Kosloff et al., 1996; Hegge, 2000] (see section 3.3.1). However, as discussed before, this user-defined parameterization may introduce a bias in the solution. The parameterization can also be modified by using varying sized grid cells [Michelena and Harris, 1991], but also this parameterization is prone to error when the wrong
cells are chosen. Modification of the parameterization can also be performed in a model dependent way. For example, cells can be clustered on basis of similarity of the current velocity updates [Hyndman and Harris, 1996; Eppstein and Dougherty, 1998]. In this way, the structure of the dominant velocity populations are estimated automatically and is not restricted by the user-defined parameterization. By using these kinds of parameterization an attempt is made to reduce the number of parameters without introducing a bias in the solution. However, it is not necessarily the case that the remaining parameters are well-determined and as a consequence the problem may remain mixed-determined. In addition, the minimum of the solution space might still not be well-shaped. Spakman and Bijwaard [2001] aim at avoiding over-parameterization and the mixed-determined nature of earthquake tomography problems, by merging cells based on the number of rays that hit these cells. Van Naver [1996] tries to circumvent this mixed-determined nature by considering the singular-values of the system. In this method, parameters are also added based on the velocity variation. However, when the singular-value of a point is too small, the point is omitted. In this way the inverse problem stays well-determined. Still, the method is influenced by the user, because the grid design is performed in an interactive way and the positions of the reflectors are assumed to be known.

3.6.2 Regularization by resolution dependent parameterization

In the inversion of focusing operators an attempt is made to combine both regularization methods: the resolution of the parameters in the optimization is translated to the spatial resolution in the parameterization. The resolution will not only determine whether parameters are removed, but also whether parameters are added. The parameterization will not be designed interactively or based on velocity gradients, but the driving force behind the parameterization will be the resolution, which is determined by the data itself.

The resolution and covariance are important tools to evaluate the solution of the optimization. Therefore, it is also useful to use them as criteria for the modification of the parameterization. In practice the resolution is used to modify the parameterization during the tomographic inversion process, as will be discussed in section 3.6.3. In the procedure for tomographic inversion of focusing operators (Figure 3.1) two processes should be added to obtain the new procedure in which resolution dependent parameterization is carried out (Figure 3.16):

[1] After optimization, the resolution of the model parameters should be calculated. The resolution matrix is defined as:

\[ R = A^\dagger (A). \]  \hspace{1cm} (3.6.33)

This matrix contains \( M \) diagonal elements that correspond to the resolution of the model parameters.
3.6 Regularization of the inversion

[2] The diagonal elements of the resolution matrix are used to modify the parameterization after optimization. Model parameters that have a low resolution can be removed, and in regions containing parameters that have a high resolution, extra parameters can be added.

By removing low-resolution parameters an attempt is made to obtain a well-determined inverse problem, in which the minimum is well-shaped. This should lead to accurate and stable results. By positioning extra parameters in high resolution regions, all the available information in the data will be translated to the model. Because the parameterization adapts to the data, a user-introduced bias is avoided. This way of parameterization is data-driven because the parameterization is determined by the resolution, while the resolution is determined by the data. As a consequence, this regularization not only attempts to generate accurate and stable results without user-introduced bias, but also to makes the method data-driven.

---

**Figure 3.16:** Procedure flowchart for tomographic inversion of focusing operators with resolution dependent parameterization. The extra processes with respect to Figure 3.1 are printed in bold letter-face.

---

In Figure 3.17 an example of resolution dependent parameterization in 2D is shown. For a velocity model the traveltime data corresponding to the rays related to a single focal point are available (Figure 3.17a). When the resolution matrix is calculated and the diagonal elements of the matrix are plotted for each gridpoint, this results in an image as shown in Figure 3.17b. In this image the light values indicate high resolution and dark values indicate low resolution. As mentioned before, low reso-
Figure 3.17: Example of resolution dependent parameterization. a) Delaunay triangulation in model defined by a regular grid, and the ray-paths from a single focal point to the surface. b) Resolution at the grid-points in the optimization with ray-paths in a). Light values indicate high resolution, dark values indicate low resolution. c) Modification of the parameterization based on the resolution in b).

Solution parameters can be removed and in regions containing parameters that have a high resolution, extra parameters can be added. This finally results in the resolution dependent parameterization shown in Figure 3.17c.

### 3.6.3 Calculation of resolution and covariance by LSQR

The resolution and covariance are important tools to evaluate the quality of the inversion result and will be used for the regularization in tomographic inversion of focusing operators. They can easily be computed by explicit optimization methods, but for iterative methods this is more complicated. In this section, the approximation of the resolution and covariance matrices by means of the LSQR algorithm is presented. In addition, it will be discussed how the resolution (and covariance) can be used in the resolution dependent modification of the parameterization.

#### Resolution

The resolution matrix can be regarded as a linear filter relating the true and the estimated model parameters. The resolution matrix can be calculated during optimization by [Menke, 1984]:

\[ R = A^T A, \]  \hspace{1cm} (3.6.34)

so \( R \) is a function of the matrix \( A \). Note that \( R \) is an \( M \times M \) matrix, if \( M \) is the number of model parameters. The \( M \) diagonal elements represent the resolution of the individual model parameters. A resolution value of 1 at the diagonal indicates that sufficient data exists to yield a unique estimate for that parameter. A resolution value close to zero reveals that the parameter is under-determined. The non-diagonal elements show the dependence between the different parameters.
In SVD the resolution matrix can be calculated by means of the model space matrix:

\[ \mathbf{R}^{[k]} = \mathbf{V}^{[k]} \mathbf{V}^{[k]}^T, \]  
(3.6.35)

where \( k \) indicates the first number of columns of matrix \( \mathbf{V} \) that are used. When an over-determined inverse problem is considered, and all columns \( (k = M) \) are used, the resolution matrix \( \mathbf{R} \) is an identity matrix \( (\mathbf{R} = \mathbf{I}) \). This indicates that sufficient data exist to solve for all parameters, but this does not indicate how well they are resolved. Note that in our problem \( N > M \), which indeed means an over-determined problem.

![Resolution Matrices](image)

**Figure 3.18:** Resolution matrices for a tomographic inversion problem with 31 parameters, calculated by means of LSQR after 7, 14, 21, 28 and 31 iterations. The matrix after 31 iterations equals the resolution matrix calculated by SVD.

For the LSQR algorithm the similarities with the SVD algorithm can be used to construct a resolution matrix. The resolution matrix can be calculated in a similar
Figure 3.19: Diagonal of the resolution matrices in Figure 3.18, representing the resolution for the 31 parameters calculated by means of LSQR after 7 (dotted), 14 (dashed dotted), 21 (dashed), 28 (solid) iterations. The resolution after 31 iterations equals one for every parameter.

way [Zhang and Mecham, 1995]:

$$R^{[k]} = \tilde{V}^{[k]} \tilde{V}^{[k]^T}. $$

(3.6.36)

Note that according to (3.5.30) the resolution matrix has to be calculated by

$$R = (\tilde{V}^{[k]} V_B) (\tilde{V}^{[k]} V_B)^T,$$

however, $V_B V_B^T$ is equal to the identity matrix, so only $\tilde{V}^{[k]} \tilde{V}^{[k]^T}$ contributes. The resulting resolution describes the resolution of the system when only $k$ basis-vectors are used. Again, the $M$ diagonal elements of resolution matrix $R^{[k]}$ represent the resolution of the separate parameters. The average amplitude of the diagonal elements of $R^{[k]}$ equals the ratio of the number of basis-vectors used in the solution $[k]$, to the total number of parameters ($M$):

$$\frac{\sum_{i=1}^{M} R^{[k]}_{ii}}{M} = \frac{k}{M}. $$

(3.6.37)

As more independent basis-vectors emerge with an increasing number of iterations, the average amplitude of the diagonal elements of the LSQR resolution matrix continues to approach 1. This is demonstrated in Figure 3.18 and Figure 3.19. The resolution matrices $R^{[k]}$ for $k = 7, 14, 21, 28$, and 31 iterations are calculated by means of the model space matrices obtained during optimization. For the maximum number of iterations ($k = M$) the resolution matrix will be a unity matrix ($R = I$). This is equivalent to the case of using all ($k = M$) SVD eigen-vectors to calculate the resolution matrix. Despite the fact that the final resolution matrix does not tell us how well the parameters are determined, the resolution matrix after a limited number of iterations reveals some features of the inverse system. This is shown in Figure 3.19, where the diagonal elements of the resolution matrices in Figure 3.18 are considered. Parameters with high resolution will be described first and low resolution parameters are described during later iterations. In addition, the parameters that have a high resolution in the first iteration will also have a high resolution
3.6 Regularization of the inversion

During the last iterations, in other words, the 'relative' resolution stays the same. So, the way the resolution evolves during the iterations is a qualitative indication of how well the parameters are determined.

Covariance

Another tool to evaluate the quality of the solution of the inverse problem is the covariance matrix. This matrix can be calculated after optimization by:

\[ \mathbf{C} = \mathbf{A}^\dagger (\mathbf{A}^\dagger)^T. \]  

(3.6.38)

Note that \( \mathbf{C} \) is an \( M \times M \) matrix. The diagonal of this matrix represents the variance of the model parameters, which is a measure of uncertainty in the parameters. The non-diagonal elements show the covariance of two parameters, which provides a measure of how strongly correlated these parameters are. In the case of using SVD the covariance matrix can be written as:

\[ \mathbf{C}^{[k]} = \mathbf{V}^{[k]} (\mathbf{\Lambda}^{[k]})^{-2} \mathbf{V}^{[k]T}. \]  

(3.6.39)

Note that in this equation besides the model space matrix the singular-values are incorporated. In a similar way, the covariance matrix can be calculated for the LSQR algorithm by:

\[ \mathbf{C}^{[k]} = \left( \mathbf{\tilde{V}}^{[k]} \mathbf{V}_\mathbf{B} \right) \mathbf{\Lambda}_\mathbf{B}^{-2} \left( \mathbf{\tilde{V}}^{[k]} \mathbf{V}_\mathbf{B} \right)^T. \]  

(3.6.40)

The resulting covariance does not only depend on the model space matrix \( \mathbf{V}^{[k]} \), but also on the 'pseudo' singular-values \( \mathbf{\Lambda}_\mathbf{B} \). The \( M \) diagonal elements of covariance matrix \( \mathbf{C} \) represent the variance of the separate parameters, while the non-diagonal elements represent the covariance between the parameters. This is demonstrated in Figure 3.20 and Figure 3.21. The covariance matrices \( \mathbf{C}^{[k]} \) for \( k = 7, 14, 21, 28 \), and 31 iterations are calculated by means of the model space matrices and the 'pseudo' singular-values obtained during optimization. The average amplitude of the diagonal elements of \( \mathbf{C}^{[k]} \) increases when more iterations are performed. For the maximum number of iterations \( (k = M) \) the covariance matrix will equal the real covariance matrix of the system.

It is very important to realize that the covariance after a limited number of iterations does not represent the real covariance. When an under-determined parameter is not described by the current basis-vectors, it incorrectly will show a low variance. This incorrectness will be revealed in one of the later iterations, when the parameter will be described by the basis-vectors that correspond to a smaller 'pseudo' singular-value. This can be observed in Figure 3.21, where the 2\(^{nd}\), 7\(^{th}\), and 31\(^{st}\) parameter have a very low variance after a limited number of iterations, but when all iterations are performed it turns out that they have a very high variance. So, the way the
Equation (3.27) is then solved by means of the LSQR algorithm [Vesnaer, 1996], and the resolution and covariance are calculated. After a number of iterations, the resolution is displayed for each parameter, as shown in Figure 3.20. The resolution and covariance are displayed for a tomographic inversion problem with 31 parameters, after iterations 7, 14, 21, 28, and 31. The final covariance equals the covariance calculated by SVD.

The properties of the resolution and covariance calculated by means of the LSQR algorithm are summarized in Table 3.2. Since in large tomographic systems it is hardly feasible to perform the maximum number of LSQR iterations, it is impractical to use the covariance to evaluate how well the parameters are determined. Along the same line of reasoning it is also impractical to use the singular-values for this evaluation [Vesnaer, 1996], as these are only available when a complete SVD (maximum number of LSQR iterations) is performed. In large tomographic systems only a limited number of iterations is feasible and therefore, only the resolution can be used to evaluate how well the
3.6 Regularization of the inversion

Figure 3.21: Diagonal of the covariance matrices in Figure 3.20, representing the variance for the 31 parameters calculated by means of LSQR after after 7 (dotted), 14 (dashed dotted), 21 (dashed), 28 solid, and 31 (thin solid) iterations.

<table>
<thead>
<tr>
<th>maximum number of iterations (=SVD)</th>
<th>limited number of iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>R ( R = I ) in over-determined case: this is an indication that all parameters ARE resolved, not how WELL they are resolved.</td>
<td>The way the resolution evolves during the iterations is an indication of how WELL the parameters are resolved.</td>
</tr>
<tr>
<td>C Final covariance is the real covariance</td>
<td>The way the covariance evolves during the iterations cannot be trusted</td>
</tr>
</tbody>
</table>

Table 3.2: Comparison of the properties of the resolution and covariance calculated by means of the LSQR algorithm and by means of SVD.

parameters are determined. Hence, in tomographic inversion of focusing operators, the resolution is used for modification of the parameterization.

Quality of the solution

Although covariances that are calculated after a limited number of iterations cannot be relied on, they can be used to describe the quality of the solution in another way. The properties of the resolution and covariance matrix together can be exploited to develop a quantitative tool for measuring the quality of the solution. Normally, when the resolution is high, the covariance is low and vice versa. However, as shown in the previous section, parameters that are not described by the current basis-vectors, contain both low resolution and low covariance. This apparent contradiction can be used to determine which parameters are resolved by the current number
of iterations. This is shown in Figure 3.22. In the beginning, after only seven iterations, most of the parameters are not described by the seven obtained basis-vectors (indicated by circles). After an increasing number of iterations, an increasing number of parameters is described and after the maximum number of iterations \((k = M = 31)\) all parameters have been determined. In other words, this procedure shows which parameters are determined after a limited number of iterations.

**LSQR optimization strategy**

In this section it has been demonstrated how the resolution and covariance can be approximated by the LSQR algorithm. The main question now is how this can be used in the regularization during tomographic inversion of focusing operators. In the LSQR algorithm the following tools were developed:

- The ‘pseudo’ singular-values can be used quantitatively to evaluate how well the system is determined;
- The ‘relative’ resolution can be used qualitatively as a criterion for the modification of parameterization;
- The comparison between resolution and covariance during the iterations can be used quantitatively to evaluate which parameters are determined by the current number of LSQR iterations.

There are two strategies for using these tools in the resolution dependent modification of the parameterization.

In the first strategy the number of LSQR iterations and the resolution threshold value, above which parameters can be added, are fixed. The addition of vertices is stopped when the singular-values become too small. However, when the number of vertices increase while the number of iterations is fixed, the relative resolution can become too small and, as a consequence, no new vertices are added. So, in this strategy the solution can be limited by both the number of iterations and the low singular-values.

In the second strategy, the number of LSQR iterations depend on the inverse problem as proposed by Paige and Saunders [1982]. The resolution threshold value is adapted to this changing number of iterations by means of equation (3.6.37). In this way, vertices can constantly be added. The addition of vertices is only stopped when the singular-values become too small. So, in this strategy the solution is only limited by the low singular-values.

In this research the first strategy is used, because it prevents the tomographic inversion from becoming too expensive when the number of LSQR iterations becomes very large.
Figure 3.22: Resolution (left) and variance (right) for the 31 parameters calculated by means of LSQR after 7, 14, 21, 28, and 31 iterations. The circles indicate parameters with both low resolution and low variance. Note that, for the sake of visibility, the vertical scale of the variance plots after 7 and 14 iterations is smaller.
3.7 Conclusions on the tomographic inversion strategy

In this chapter, the methodology of tomographic inversion of focusing operators has been discussed. The two objectives of the research that are addressed here are: 1) Developing a method for tomographic inversion of focusing operators, and 2) Making the method data-driven (section 1.3). The conclusions on the tomographic inversion strategy are subdivided in these two objectives.

3.7.1 Tomographic inversion of focusing operators

The tomographic problem is nonlinear, but in order to be able to invert for the model parameters, the tomographic problem is linearized. As a result, the total inversion is replaced by an iterative procedure. In each iteration the inversion of the traveltime differences between observed and modeled focusing operators leads to an update of the initial model parameters.

The subsurface is described by model parameters that influence the traveltimes: the velocities and the locations of the focal points.

- The velocities in the model are defined on irregular grid-points, which are connected by Delaunay triangulation, in which the quadratic slowness is linearly interpolated. Defining the velocities at the grid-points, and not within the triangles reduces the number of parameters, and results in a smooth model. Using the quadratic slowness makes it possible to derive a simple ray-trace system.

- The flexible velocity parameterization is required to make the data-driven method possible.

- The method of Delaunay triangulation allows for adding grid-points in an existing parameterization. It is not necessary to rebuild the entire parameterization when it is modified (see appendix A).

- The focal points are defined independently of the velocity grid-points. A focal point location can be described by a vector containing the coordinates. By parameterizing them independently, the velocity changes are not dependent on, or restrained to, the reflectors.

The forward modeling of focusing operators is performed by a ray-based method. The focal points are considered to be sources from which rays are calculated towards receivers at the surface.

- Both the traveltimes and the ray-paths of the seismic rays are calculated analytically. As a consequence, the method is efficient in memory and can calculate the traveltimes accurately.
3.7 Conclusions on the tomographic inversion strategy

- The ray-based method is a 'shooting method'. Hence, to reach a predefined receiver, interpolation by re-shooting is required. Although this interpolation is performed efficiently, the re-shooting makes the method slower. In addition, it is not always possible to reach the receiver due to shadow-zones.

- The ray-paths can be used to describe the relationship between the traveltimes and the model parameters. As a consequence, the partial derivatives of the traveltimes with respect to the quadratic slownesses and the focal point locations are also calculated analytically during the ray-tracing. They are used to update the model parameters during the optimization step.

The optimization of the model parameter updates is performed by the LSQR method.

- The LSQR method determines the generalized inverse in an iterative way. Therefore, it is efficient in memory, and by limiting the number of iterations it is also efficient in speed.

- By limiting the number of LSQR iterations the optimization is automatically stabilized, because the algorithm solves the largest basis-vectors first. It is similar to truncating the SVD.

- The system can be analyzed by evaluating the 'pseudo' singular-values. These can be obtained from the LSQR method by applying a SVD to the bi-diagonal matrix. This is more efficient than applying SVD directly to the optimization matrix, which is hardly feasible.

3.7.2 Data-driven procedure

The regularization of the inverse problem is performed by resolution dependent parameterization. As the parameterization depends on the resolution, and the resolution depends on the data, this inversion becomes data-driven.

- The relative resolution of the parameters can efficiently be calculated by means of the LSQR method, by using the model space matrix. This relative resolution can be used qualitatively as a criterion for the modification of the parameterization. In this way, an expensive explicit matrix inversion, that is normally needed to obtain the resolution, is not required.

- Model parameters that have a low resolution can be removed and in regions containing parameters that have a high resolution, extra parameters can be added. In this way an attempt is made to get a well-determined inverse problem in which the minimum is well-shaped.

- In addition, all the available information within the data can be translated into model parameters, because the parameterization adapts to the data. In this way, a user introduced bias is avoided.
Besides the resolution, the covariance can also be calculated by the LSQR method. Although the accurate covariance is only obtained when the maximum number of LSQR iterations is performed, the intermediate result can also be used. The comparison between the resolution and the covariance during the iterations can be used quantitatively to evaluate which parameters are resolved.

In Chapters 5 and 6, the methodology as described in this chapter is applied to both 2D and 3D synthetic and field data examples. In the next chapter, the methodology of the third objective is addressed: exploring the possibilities of the tomographic inversion of focusing operators.
Additional modules to the tomographic inversion of focusing operators

4.1 Introduction

The methodology of tomographic inversion of focusing operators, as described in the previous chapter, primarily aims at translating one-way traveltimes focusing operators to a velocity model of the subsurface. In this chapter, the possibilities of the tomographic inversion of focusing operators are explored, which is the third objective of this research (section 1.3). Three additional modules to this inversion are introduced:

[1] Inversion of the focusing operators using a priori information;

The first module is related to adding extra information in the inversion. Although this module will be presented for the inversion of focusing operators, it can be applied in any kind of traveltime inversion. The last two modules are specifically based on the inversion of one-way traveltime focusing operators. The theory of the three modules will be explained in this chapter.
4.2 Inversion of focusing operators using a priori information

In tomographic inversion of focusing operators only the traveltimes of the focusing operators are used to obtain a model of the subsurface. However, it might be desirable to include some a priori information, if available. This a priori information can be any kind of data or knowledge about the values of the model parameters that we are inverting for.

A priori information on the model parameters can be incorporated in the model by fixing the value of the parameter in the model. However, it is more appropriate to incorporate the a priori information in the optimization algorithm. Equation (3.2.1) then changes to:

\[
\begin{bmatrix}
A \\
I
\end{bmatrix}
\Delta m = \begin{bmatrix}
\Delta d \\
\Delta m_p
\end{bmatrix},
\]

(4.2.1)

where vector \(\Delta m_p\) contains updates dependent on the a priori information of the model parameters, i.e. is the difference between the a priori and the initial parameter values. In this way, the a priori information acts as an extra constraint in the inversion.

In this formulation the a priori information should be available at the location of the model parameters. Normally, a priori information is available at very specific locations like wells, water layers, salt bodies, etc, which do not necessarily coincide with the parameterization of the velocity model. A method to obtain the a priori information at the location of the model parameters is provided by geostatistics.

4.2.1 Transferring a priori information to the model parameterization by geostatistics

A priori information can be available at any location in the subsurface, and is therefore defined on separate points in the model. In this way, any kind of a priori information can be described. In Figure 4.1a points with the a priori information from four well-logs and from the water layer are defined. The points that contain the a priori information are called samples. The goal is now to transfer values of the a priori information defined at the samples, to the grid-points of the model parameterization. This is achieved by defining the value at a grid-point \(m_p\) as a linear combination of the value at the available \(S\) samples \(m_s^{[j]}\) (\(j = 1, 2, \ldots, S\)), so

\[
m_p = \sum_{j=1}^{S} w^{[j]} m_s^{[j]},
\]

(4.2.2)

where \(w^{[j]}\) forms the weight for sample \(j\). This process is called Kriging [Isaaks and Srivastava, 1989]. The required weights can be calculated based on the covariance
of the samples. First a covariance function of the a priori model information is determined by calculating the covariance between the known samples as a function of a specific distance $h_s$ between the locations of the samples. So, the covariance $C(h_s)$ between the samples $m_s^{[i]}$ and $m_s^{[j]}$ for specific distances is defined as:

$$C(h_s) = \left( \frac{1}{N(h_s)} \sum_{i,j} m_s^{[i]} \cdot m_s^{[j]} \right) $$

$$ - \left( \frac{1}{N(h_s)} \sum_i m_s^{[i]} \right) \cdot \left( \frac{1}{N(h_s)} \sum_j m_s^{[j]} \right),$$

(4.2.3)

where $h_s^{[i,j]}$ indicates the distance between sample $i$ and $j$ ($i = 1, 2, ...S, j = 1, 2, ...S$). The summation is carried out only over the $N(h_s)$ pairs of samples whose locations are separated by $h_s$. When the covariance values $C(h_s)$ are plotted against the several specific distances $h_s$, a covariance plot is obtained (Figure 4.1b). From the covariance plot a covariance function can be determined by fitting a predefined function, e.g. a linear, Gaussian, or exponential function, in a least squares sense (line in Figure 4.1b). This covariance function can be used to calculate the theoretical covariances for both the distances $h_s^{[i]}$ between all pairs of the known sample $(C(h_s^{[i]}))$ and for the distances $h_s^{[ip]}$ between the samples and the unmeasured point $(C(h_s^{[ip]})$. By means of these theoretical covariances the unknown vector of weights $w_s^{[j]}$ for point $p$ can be obtained by $\sum_j C(h_s^{[i,j]} \cdot w_s^{[j]} = C(h_s^{[ip]}).

With the help of a Lagrangian multiplier $\lambda$ the constraint equation $\sum_j w_s^{[j]} = 1$ is added to the equations. Now, the weights for point $p$ can be calculated by:

$$\begin{bmatrix} w_1^{[1]} \\ w_2^{[2]} \\ \vdots \\ w_s^{[S]} \\ \lambda \end{bmatrix} = \begin{bmatrix} C(h_s^{[11]}) & C(h_s^{[12]}) & \ldots & C(h_s^{[1S]}) \\ C(h_s^{[21]}) & C(h_s^{[22]}) & \ldots & C(h_s^{[2S]}) \\ \vdots & \vdots & \ddots & \vdots \\ C(h_s^{[S1]}) & C(h_s^{[S2]}) & \ldots & C(h_s^{[SS]}) \\ 1 & 1 & \ldots & 1 \end{bmatrix}^{-1} \begin{bmatrix} C(h_s^{[ip]}) \\ C(h_s^{[ip]}) \\ \vdots \\ C(h_s^{[ip]}) \\ 1 \end{bmatrix}.$$  

(4.2.4)

The weighting vector can be used for interpolating the unknown values from the a priori samples $m_p$ in (4.2.2). When $m_p$ is determined for all locations in the subsurface, the result is shown in Figure 4.1c. It is obvious that the closer the location of the model parameters is to the samples, the lower the uncertainty in the estimation will be. The error variances for the interpolated values are proportional to the covariances between the samples and the unmeasured point [Isaaks and Srivastava, 1989]:

$$c_p = \sum_{j=1}^{S} w_s^{[j]} C(h_s^{[i,j]}) + \lambda,$$

(4.2.5)
which can be converted to weights for the optimization, e.g. by: \( w_p = \frac{1}{c_p} \). The weights for the Kriging result in Figure 4.1c is shown in Figure 4.1d.

After the estimated values and their weight are calculated at the locations of the model parameters, they can both be incorporated in the optimization. The estimated values \( m_p^{[i]} \), with \( (i = 1, 2, ..., M) \) for \( M \) model parameters, serve as extra data in the data vector and they are properly included in the optimization matrix by means of weights \( w_p^{[i]} \):

\[
\begin{bmatrix}
A \\
W_p^{\frac{1}{2}}
\end{bmatrix}
\begin{bmatrix}
\Delta m \\
\Delta d
\end{bmatrix} =
\begin{bmatrix}
\Delta d \\
W_p^{\frac{1}{2}} \Delta m_p
\end{bmatrix},
\]

(4.2.6)

where

\[
\Delta m_p =
\begin{bmatrix}
m_p^{[1]} - m_{\text{init}}^{[1]} \\
m_p^{[2]} - m_{\text{init}}^{[2]} \\
\vdots \\
m_p^{[M]} - m_{\text{init}}^{[M]}
\end{bmatrix},
\quad
W_p =
\begin{bmatrix}
w_p^{[1]} & 0 & \ldots & 0 \\
0 & w_p^{[2]} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & w_p^{[M]}
\end{bmatrix}.
\]

(4.2.7)

Vector \( \Delta m_p \) contains the difference between the a priori information and the initial values of the model parameter \( m_{\text{init}} \). Matrix \( W_p \) describes the weighting of the parameters by means of the uncertainties; when the uncertainty in the data is high, it is reasonable to give priority to the a priori information and vice versa. In this way the updates in \( \Delta m \) will not only depend on the traveltimes of the focusing operators in \( \Delta d \), but also on the a priori information in \( \Delta m_p \). The method of including a priori information, as presented in this section, is completely compatible with the tomographic inversion method as presented in Chapter 3. The information can be available at irregular grid-points in the subsurface, and all sorts of a priori information can be transferred to these grid-points by Kriging. In addition, the a priori information will take part in the optimization of the parameters and proper weighting of this information can be included. Finally, the inclusion of a priori information will also play a role in the resolution dependent modification of the parameterization, as it will (locally) increase the resolution, resulting in a finer parameterization.

An application of the tomographic inversion using a priori information will be presented in section 5.2.8.
4.2 Inversion of focusing operators using a priori information

Figure 4.1: a) Velocity model with samples taken in this model (white dots). The samples are obtained from four well logs and knowledge about the water layer. b) The covariance plot (dots) and estimated covariance function (line). c) The result after Kriging for the entire subsurface. d) The error variance of the Krigeed result in c). Dark values indicate high variance, i.e. low resolution.
4.3 Joint inversion of P and S-wave focusing operators

Seismic waves are elastic waves that can be subdivided into two types: Pressure waves, denoted by a 'P', and shear waves, denoted by an 'S'. Generally, only the P-wave type is used, however, the P and S-waves propagate differently through the subsurface and, as a consequence, contain different information about the subsurface. For example, the S-waves provide valuable information on rock and fluid parameters, that cannot be inferred from conventional P-waves. Therefore, a fuller description of a complex earth can be achieved by using both P and S-waves. Generally, seismic exploration, both on- and off-shore, is carried out using P-waves, because they are easy to generate and to record. Although it is more difficult, onshore S-waves can also be generated and recorded. Offshore, however, this is not possible, as S-waves cannot be transmitted through water. A solution to this problem is the use of Ocean Bottom Station (OBS) recording. In this case P-waves are generated at seafloor and travel downwards. The P-waves reflect at boundary layers, where a mode conversion to S-waves takes place. These S-waves are recorded at the ocean bottom. These mode converted data will be considered in this section.

The CFP method is pre-eminently suited to image mode converted data accurately, because the CFP method addresses upward and downward traveling energy separately (see Chapter 2). Focusing of the sources and focusing of the receivers can be performed using a P or an S-wave operator [Berkhout and Verschuur, 2000; Verschuur et al., 2002].

In this section it will be explained how the P-waves and S-waves can be described by the WRW model. Next, it will be shown how the P and S-wave operators can be retrieved from seismic data using the CFP method. Finally, a scheme for joint tomographic inversion of P and S-wave operators is presented.

### 4.3.1 The WRW model for P and S-wave data

The WRW model was introduced in Chapter 2, to describe seismic data. It was defined by equation (2.2.1), which is repeated here for the sake of convenience:

\[
P(z_0, z_0) = D^-(z_0) \sum_{m=1}^{M} [W^-(z_0, z_m) R(z_m) W^+(z_m, z_0)] S^+(z_0).
\]  

(4.3.8)

In this equation P describes the data as obtained at the surface, S^+ contains the properties of the sources and D^- of the receivers, W^+ describes the downward and W^- the upward propagation, and R contains the reflection properties. Until now only data related to P-waves were considered. However, this equation can easily be generalized to any kind of wave type, so P as well as S, and to any kind of mode conversion between these wave types [Berkhout and Verschuur, 2000]. This is done by defining:
4.3 Joint inversion of P and S-wave focusing operators

\[ P(z_0, z_0) = \begin{bmatrix} P_{PP}(z_0, z_0) & P_{PS}(z_0, z_0) \\ P_{SP}(z_0, z_0) & P_{SS}(z_0, z_0) \end{bmatrix} \]

\[ S^+(z_0) = \begin{bmatrix} S^+_P(z_0) & 0 \\ 0 & S^+_S(z_0) \end{bmatrix} \]

\[ W^+(z_m, z_0) = \begin{bmatrix} W^+_{PP}(z_m, z_0) & W^+_{PS}(z_m, z_0) \\ W^+_{SP}(z_m, z_0) & W^+_{SS}(z_m, z_0) \end{bmatrix} \]

\[ R(z_m) = \begin{bmatrix} R_{PP}(z_m) & R_{PS}(z_m) \\ R_{SP}(z_m) & R_{SS}(z_m) \end{bmatrix} \]

\[ W^-(z_0, z_m) = \begin{bmatrix} W^-_{PP}(z_0, z_m) & W^-_{PS}(z_0, z_m) \\ W^-_{SP}(z_0, z_m) & W^-_{SS}(z_0, z_m) \end{bmatrix} \]

\[ D^-(z_0) = \begin{bmatrix} D^-_P(z_0) & 0 \\ 0 & D^-_S(z_0) \end{bmatrix} \]

The non-diagonal elements in \( W^+ \) and \( W^- \) indicate mode conversion during propagation, and the non-diagonal elements in \( R \) are related to mode conversion during reflection. If only mode conversion due to reflection is considered, equation (4.3.8) can be represented by four independent equations, which can be described by

\[ P_{\alpha\beta}(z_0, z_0) = D^-_{\beta}(z_0) \sum_{m=1}^{M} \left[ W^-_{\beta\beta}(z_0, z_m) R_{\beta\alpha}(z_m) W^+_{\alpha\alpha}(z_m, z_0) \right] S^+_{\alpha}(z_0), \quad (4.3.9) \]

in which \( \alpha \) and \( \beta \) stand for P or S. In the situation of Ocean Bottom data, the source only generates P-waves, so only the following versions of equation (4.3.9) should be used

\[ P_{PP}(z_0, z_0) = D^-_{P}(z_0) \sum_{m=1}^{M} \left[ W^-_{PP}(z_0, z_m) R_{PP}(z_m) W^+_{PP}(z_m, z_0) \right] S^+_{P}(z_0), \quad (4.3.10) \]

which represents the conventional P data, and

\[ P_{SP}(z_0, z_0) = D^-_{S}(z_0) \sum_{m=1}^{M} \left[ W^-_{SP}(z_0, z_m) R_{SP}(z_m) W^+_{PP}(z_m, z_0) \right] S^+_{S}(z_0), \quad (4.3.11) \]

which represents the converted P to S data. Note that \( z_b \) represents the ocean bottom, where the receivers of OBS data are located. These are the equations that are used in the remainder of this section.
\subsection{4.3.2 The CFP method for P and S-wave data}

Focusing of the mode converted data can be done with a P or an S-wave operator. This means that the principle of equal traveltime can be applied for both P and S traveltimes. This procedure will be explained (in a similar way as in section 2.3) by means of focusing of receivers and focusing of sources.

\section*{Focusing of the receivers}

When PP data are considered, focusing of the receivers occurs by applying a P-wave operator $\mathbf{F}_{P,j}^{\dagger}$ to the upward propagation side of the data matrix $\mathbf{P}_{PP}$:

$$\mathbf{P}_{P,j}^{\dagger}(z_m, z_0) = \mathbf{F}_{P,j}^{\dagger}(z_m, z_0) \mathbf{P}_{PP}(z_b, z_0)$$

$$= \mathbf{R}_{PP}(z_m) \mathbf{W}_{PP}^{\dagger}(z_m, z_0).$$

The focal point response in CFP gather $\mathbf{P}_{P,j}^{\dagger}(z_m, z_0)$ contains the downward propagation of the P-wave.

In the case of PS data, focusing of the receivers occurs by applying an S-wave operator $\mathbf{F}_{S,j}^{\dagger}$ to the upward propagation side of the data matrix $\mathbf{P}_{SP}$:

$$\mathbf{P}_{P,j}^{\dagger}(z_m, z_0) = \mathbf{F}_{S,j}^{\dagger}(z_m, z_0) \mathbf{P}_{SP}(z_b, z_0)$$

$$= \mathbf{R}_{SP}(z_m) \mathbf{W}_{PP}^{\dagger}(z_m, z_0).$$

The focal point response in CFP gather $\mathbf{P}_{P,j}^{\dagger}(z_m, z_0)$ contains the downward propagation of the P-wave.

This indicates that the focal point response in the CFP gather of the PP data ($\mathbf{W}_{PP}$ being removed at the receiver side) and in the CFP gather of the PS data ($\mathbf{W}_{SS}$ being removed at the receiver side) have the same traveltimes. In addition, these traveltimes equal the traveltimes of the P-wave operator.

\section*{Focusing of the sources}

The sources can only be focused by applying a P-wave operator, because the downward propagation part from the sources to the reflector only contains P-waves. So, focusing of the sources occurs by applying a focusing operator $\mathbf{F}_{P,j}^{\dagger}$ to the downward propagation side of the data matrix:

$$\mathbf{P}_{P,j}(z_b, z_m) = \mathbf{P}_{PP}(z_b, z_0) \mathbf{F}_{P,j}^{\dagger}(z_b, z_m)$$

$$= \mathbf{W}_{PP}(z_b) \mathbf{R}_{PP}(z_m),$$
The focal point response in CFP gather \( P_{P,j}^- \) contains the upward propagation of the P-wave:

\[
P_{S,j}^-(z_b, z_m) = P_{SP}(z_b, z_0) F_{P,j}^+(z_b, z_m) = W_{SS}^-(z_0, z_m) R_{SP}(z_m).
\]

The focal point response in CFP gather \( P_{S,j}^- \) contains the upward propagation of the S-wave. This indicates that the focal point response in the CFP gather of the PP data and the CFP gather of the PS data (in both cases \( W_{PP} \) is removed at the source side) differ in traveltime. However, this property can be used for the estimation of the P and S-wave operators, as will be shown in the next section.

### 4.3.3 Focusing operator updating for P and S-wave data

As presented in Chapter 2, the P-wave operator could be updated by applying the principle of equal traveltime to the PP data. In the same way the S-wave operator can be determined by means of the SS data. However, in OBS data the SS data is not available. It is possible to obtain the S-wave operator from converted wave data, but in this case another strategy is used. After the P-wave operator has been obtained from the PP data, it can be used to determine the S-wave operator, related to the same focal point, via the converted wave data [Verschuur et al., 2002]:

1. Use PP data to get the P-wave operator;
2. Focus the sources of the PS data by applying the P-wave operator;
3. 'Read' the S-wave operator from the obtained CFP-gather. The location in time of the corresponding focal point response can be found using the vertical two-way traveltimes of the PP and PS data, by \( t_S = t_{PS} - \frac{t_{PP}}{2} \).

This updating procedure leads to verified P and S-wave operators which are related to the same focal point. These operators automatically include anisotropy and dispersion effects.

### 4.3.4 The aperture of the P and S-wave focusing operators

In order to investigate the aperture of the P and S-wave operators, consider a plane interface as in Figure 4.2 [Borghols, 2003]. In general, the ray-parameter \( p = s \sin \theta \) of a ray propagation through a layer, has the same value for the incident, reflected, and refracted wave. When a P-wave with an angle of incidence \( \theta_P \) reflects at a boundary, the angle of reflection will also be \( \theta_P \), because the slowness \( s_P \) stays the same. However, when the same incident P-wave converts to an S-wave during the reflection at the boundary, the angle of reflection will change to \( \theta_S \), because the
S slowness \( s_S \) will be different than the P slowness \( s_P \). For example, an incident P-wave with a maximum incidence angle \( \theta_{P_{\text{max}}} \) result in a reflected S-wave with a maximum reflection angle of \( \theta_{S_{\text{max}}} \):

\[
s_P \sin \theta_{P_{\text{max}}} = s_S \sin \theta_{S_{\text{max}}},
\]

and as a result

\[
\theta_{S_{\text{max}}} = \sin^{-1} \left( \frac{s_P}{s_S} \sin \theta_{P_{\text{max}}} \right).
\]

Since the P-wave slowness is smaller than the S-wave slowness, the maximum angle of the S-wave (\( \theta_{S_{\text{max}}} \)) will be smaller that the maximum angle of the P-wave (\( \theta_{P_{\text{max}}} \)). As a consequence, for converted waves in seismic data with a maximum source-receiver offset, the aperture of the S-waves is much smaller than the aperture of the P-waves. Therefore, also the aperture of the S-wave operators will be smaller than the aperture of the P-wave operators.

\[\text{Figure 4.2: Conversion from P to S-wave, resulting in a limited aperture of the S-wave operator.}\]

### 4.3.5 Joint tomographic inversion of P and S-wave focusing operators

Once the P and S-wave operators are available, they can be used for tomographic inversion. In the inversion equation \( \Delta \mathbf{d} = \mathbf{A} \Delta \mathbf{m} \) (3.2.1) both a slowness part (superscript \( s \)) and a focal point location part (superscript \( x \)) is present. For the P-wave operators this can be written as

\[
\Delta \mathbf{d}_P = \begin{bmatrix} A^P_P & A^P_x \end{bmatrix} \begin{bmatrix} \Delta \mathbf{m}^P_P \\ \Delta \mathbf{m}^P_x \end{bmatrix},
\]

(4.3.18)
and for the S-wave operators as

\[
\Delta d_S = \begin{bmatrix} A_S^* & A_S \end{bmatrix} \begin{bmatrix} \Delta m_S^* \\ \Delta m_S \end{bmatrix}.
\]  
(4.3.19)

The P and S-wave operators are both the result of focusing at the same subsurface focal point, which is the conversion point. Therefore, both the P and the S-wave operators can be seen as a response from the same focal point in the subsurface to receivers at the surface, so \(m_S^* = m_S = m^*\). Consequently, in tomographic inversion the initial focal points have to be defined at identical locations in the P and S models, and the updates of the focal point locations in both models are also the same. This leads to the following simplification [Borghols, 2003]:

\[
\Delta m_S^* = \Delta m_S = \Delta m^*.
\]  
(4.3.20)

Combining equations (4.3.18) and (4.3.19) by using equation (4.3.20) leads to:

\[
\begin{bmatrix} \Delta d_P \\ \Delta d_S \end{bmatrix} = \begin{bmatrix} A_P^* & 0 \\ 0 & A_S^* \\ A_S \end{bmatrix} \begin{bmatrix} \Delta m_P^* \\ \Delta m_S^* \\ \Delta m_S \end{bmatrix}.
\]  
(4.3.21)

In this equation the slowness parts of the optimization matrices for the P and S-wave operators are still separated. Note, however, that both operator types are indeed coupled via the focal point locations. This equation can be used for the joint tomographic inversion of P and S-wave operators.

The joint inversion process can be visualized by extending the flowchart for the tomographic inversion given by Figure 3.16. This results in the flowchart as depicted in Figure 4.3, in which the P-wave system (left-hand side) and the S-wave system (right-hand side) are combined in a joint inversion (center part). The input is formed by both P and S-wave operators, and both an initial P and an initial S model. The inversion generates a final estimate of both the P and S model, in which the slownesses can differ, but the focal point locations are the same. An application of the joint inversion of P and S-wave operators will be presented in section 5.2.9.
Figure 4.3: Flowchart for the joint tomographic inversion of P and S-wave operators, based on the flowchart for the single tomographic inversion as shown in Figure 3.16. In this flowchart the P-wave system (left-hand side), and the S-wave system (right-hand side) are combined in a joint inversion (center part).
4.4 Evaluation of the inversion result using focal point clouds

Focusing operators are inverted to obtain an estimate of both the velocity in the subsurface and the locations of the focal points. The focal points are positioned in such a way that the traveltimes of the rays, that originate from these points, fit the traveltimes of the focusing operators. At the same time, the velocities represent the best fit for the traveltimes of the rays that are traveling through the subsurface. However, it is expected that the updated velocity model cannot fully explain the traveltimes of the focusing operators. In this case, the rays will still have traveltime residuals.

The traveltime residuals of the rays that originate from one focal point can be mapped around this focal point. This is done by making use of the rays of the focusing operator that are calculated in the updated velocity model. We can backtrack along the rays of this focusing operator until the traveltimes of the observed focusing operator are reached. By putting a point at these locations, the traveltime residual is translated into an estimated focal point location (see Figure 4.4). When this is done for all rays of one focal point, this results in a 'cloud' of points around the focal point: the focal point cloud (Figure 4.4c). The focal point cloud reveals how well the velocity model can explain the focusing operators. When the focusing operators are exactly matched, the focal point cloud will turn out to form a point at the focal point location. In this research, the focal point clouds are only used to evaluate the capacity of the model to explain the operators, not the accuracy of the operators themselves, which are assumed to be correct.

![Figure 4.4: a) The rays of the focusing operator in the updated velocity model. b) Backtrack along the rays in a) until the traveltimes of the observed focusing operator are reached. c) The points plotted at the location where the rays in b) ended form the focal point clouds.](image)
Additional modules to the tomographic inversion of focusing operators

Figure 4.5: a) A 3D model containing 3 layers, that have a velocity of $1.5\text{ km/s}$, $2.0\text{ km/s}$, and $2.5\text{ km/s}$ respectively. b) Each layer contains $20 \times 20$ focal points. c) Focal point clouds of a selection of focal points, using an erroneous (homogeneous) velocity ($2.0\text{ km/s}$), and erroneous focal point locations. d) Focal point clouds of the same focal points as in c), using the correct velocity and focal point locations.

An example of focal point clouds in 3D is shown Figure 4.5. The observed focusing operators originate from this 3D model that contains three layers with velocities of $1.5\text{ km/s}$, $2.0\text{ km/s}$, and $2.5\text{ km/s}$ respectively, and focal points at the reflectors between these layers. When the traveltime residual between the observed focusing operators and the focusing operators related to an erroneous velocity model are mapped, all kind of focal point clouds appear (Figure 4.5c). When an accurate velocity model is used, the clouds focus exactly in a point (Figure 4.5d). So, by means of the focal point clouds, the capacity of the velocity model to focus the operators is evaluated. Thanks to this, the focal point clouds can be used as a tool to:

1. Resolve the velocity-depth ambiguity during updating;
2. Indicate anisotropy;
3. Discover 3D effects in 2D tomographic inversion;
4. Influence the focusing,
These tools will be demonstrated on the model as presented in Figure 4.5.

4.4.1 Analyzing the focal point clouds to resolve the velocity-depth ambiguity during updating

The shape of the focal point clouds can indicate a specific problem in the estimated parameters. In general, when an error in the location of the focal point occurs, the cloud will have a curved 'horse shoe' shape (Figure 4.6). For example, when the curve of the 'horse shoe' is upward, the focal point is positioned too shallow, and when the curve is downward, the focal point is positioned too deep. When an error in velocity is present, the cloud will be flat. A flat cloud, that is positioned above the focal point, reveals a velocity that is too low, and when the flat cloud is positioned below this focal point, the velocity is too high.

The focal point clouds that indicate errors in velocity or depth in the 3D model (Figure 4.5), are shown in Figure 4.7. The use of the correct focal point locations and an erroneous homogeneous velocity of 2.0 km/s results in the clouds as presented in Figure 4.7a. Thanks to the flat focal point clouds, which are positioned too deep, it becomes clear that for the first layer the velocity is too high and the focal point locations are probably correct. For the second layer the focal point clouds form almost a point, which agrees with the fact that both the used velocity and the velocity of the second layer are 2.0 km/s. The third layer shows, despite the influence of the overburden, flat clouds that are positioned too shallow. This indicates a velocity that is too low. Figure 4.7b presents the focal point clouds when the velocity is correct and the focal point locations are too shallow. This error is clearly indicated by the 'horse shoes' that are curved upward.

In most cases the shape of the focal point cloud will not be influenced by either velocity or depth errors, but by a superposition of both errors (Figure 4.8). In addition, in a complex subsurface the shape of the clouds will be a superposition of the errors of the different layers. Despite all this, generally, a flatter cloud will indicate a larger error in velocity than in depth, while a more curved cloud indicates the opposite.

This knowledge can be used to unravel the velocity-depth ambiguity during updating. When the focal point clouds are flat, priority should be given to the updating of the velocity in the tomographic inversion. This concept is shown in Figure 4.9. Suppose, that the tomographic inversion procedure is started with a velocity model that results in the focal point clouds as presented in Figure 4.9a. These clouds reveal that the model that is used contains a velocity error. However, when the tomographic inversion is performed, both the velocity and focal point locations are updated (Figure 4.9b). In other words, the error in velocity is also translated to an update of the focal point locations, which is caused by the velocity-depth ambiguity in the update. It is most likely that in the next updates this erroneous translation
from velocity error to focal point location update will be neutralized, but it is undesirable that this translation is made in the first place. This ambiguity during the updating can be tackled by using the information revealed by the focal point clouds. The clouds showed that the emphasis should be put on the velocity updating during the tomographic inversion. This results in the new focal point clouds in which the error in velocity is not translated to an update of focal point location (Figure 4.9c). The traveltime residuals of the results in Figure 4.9b and c are the same, despite the correct translation from traveltime differences to velocity updates. This, again, emphasizes the presence of a velocity-depth ambiguity.

**Figure 4.6:** Shape of the focal point clouds for errors in either velocity or depth. The darkness of the line shows to which surface location the focal point is related; the darker part is related to rays coming from the left and the lighter part is related to rays coming from the right. The circle represents the focal point location in the updated model. The star represents the true focal point location.
4.4 Evaluation of the inversion result using focal point clouds

Figure 4.7: a) The focal point clouds using the correct focal point locations (open dots) and an homogeneous velocity of \(2.0\text{ km/s}\). b) The focal point clouds using the correct velocities and focal points that are all positioned too shallow (open dots).

Figure 4.8: Shape of the focal point clouds for a combination of errors in velocity and depth.
Figure 4.9: The use of the focal point clouds to solve the velocity-depth ambiguity during updating. a) The flat focal point clouds for the first update indicate that there is an error in velocity. b) The clouds after updating both velocity and focal point locations. c) The clouds after updating velocity only.
4.4 Evaluation of the inversion result using focal point clouds

4.4.2 Evaluating the focal point clouds to indicate anisotropy

Traveltime residuals do not only occur due to erroneous velocities or focal point locations. They also occur when isotropy in the velocity model is assumed, while anisotropy is present in the focusing operators that are inverted (and thus in the earth). In this situation, (isotropic) tomographic inversion is not able to explain the traveltimes completely. The resulting traveltime residuals will appear in the focal point clouds, showing a characteristic shape (Figure 4.10a,b). Not only the anisotropy itself, but also the direction of anisotropy can be revealed by the clouds (Figure 4.10c,d).

Figure 4.11 shows how the focal point clouds can be used to indicate anisotropy. Again, the 3D model is used, however, in this case the true model and the observed focusing operators in this model contain elliptical anisotropy. These anisotropic operators are tomographically inverted using isotropic velocities. The updating procedure is started using a homogeneous velocity and erroneous focal point locations. The corresponding focal point clouds (Figure 4.11a,b) show that the model used is incorrect. No anisotropy is visible in the clouds yet, because the error due to anisotropy is overruled by the error in velocity and focal point location. After several updates, the velocities and focal point locations are updated, the traveltime residual does not decrease anymore, and the inversion is stopped. However, the focal point clouds show that the focusing using the updated model is not accurate (Figure 4.11c,d). The characteristic shapes of the clouds reveal that anisotropy is present in the observed focusing operators, which has not been described by the isotropic velocity model. Based on this focal point cloud information it can be decided either to perform anisotropic tomographic inversion or to put anisotropy in the model manually, as both the anisotropy and its direction are revealed by the clouds.
Figure 4.10: Shape of an the focal cloud when anisotropy is neglected in the updating. The true model has an elliptical anisotropy in which a) the vertical direction is the fastest direction, or in which b) the horizontal direction is the fastest direction. In c) and d) the anisotropy is modeled with two different orientations. Note that the direction of anisotropy is revealed by the focal point clouds.
4.4 Evaluation of the inversion result using focal point clouds

Figure 4.11: The focal point clouds during updating of anisotropic focusing operators, using an isotropic inversion. a) Focal point clouds for the initial model. b) Selection of focal point clouds as presented in a). c) Focal point clouds for the final model. d) Selection of focal point clouds as presented in c).
### 4.4.3 Evaluating the focal point clouds to discover 3D effects

Traveltime residuals can also occur when 2D models are used to explain data that have traveled through a 3D earth. In this case, the resulting traveltime residuals will also appear in the focal point clouds. However, the shape can be complicated, dependent on the complexity of the 3D structure.

![Figure 4.12](image)

**Figure 4.12:** a) The 3D model used for testing the 3D effects in focal point clouds in 2D tomographic inversion. The stars and triangles indicate the sources and receivers respectively, of the 2D acquisition geometry. b) The focal point clouds after updating 2D focusing operators, which contain 3D effects. c) Selection of focal point clouds as presented in b).

An example of 3D effects in 2D focal clouds is presented in Figure 4.12. A data set with a 2D geometry in the x-direction at \( y = 0.0 \text{ km} \) is acquired in this 3D model. The shots start at \( x = 1.5 \text{ km} \), and end at \( x = 6.0 \text{ km} \). The length of the receiver array is 1.5 km. The focal points are related to the second reflector at 0.6 km depth. The principle of equal traveltime is used to estimate 2D focusing operators, which show a good focusing of the data. Next, these operators are inverted to a 2D model. However, the 2D model will not be able to explain the traveltimes of the rays that traveled outside the 2D plane of the model. Figure 4.12c shows that the focal clouds caused by 3D effects might be similar to the focal clouds caused by anisotropy (Figure 4.10). As a consequence, care should be taken not to mis-interpret the clouds.
4.4 Evaluation of the inversion result using focal point clouds

4.4.4 Using the focal point clouds to influence the inversion

The focal point clouds show how well the model can explain the traveltimes of the focusing operators. This can be analyzed visually, but also quantitatively by calculating the rms values of the traveltime residuals (\(\Delta t_{rms}\)) that belong to one focal point:

\[
\Delta t_{rms} = \sqrt{\frac{\sum_{i=1}^{N} (\Delta t[i])^2}{N}},
\]

in which the focal point has \(N\) rays, that have an error of \(\Delta t\). The focal point clouds and the \(\Delta t_{rms}\) for the three reflectors of the 3D model are shown in Figure 4.13. When the first reflector is considered, a peculiar focusing pattern appears: The \(\Delta t_{rms}\) values are higher in the central area (light color). However, this is also the area where the most rays are available. Although this seems contradictory, the presence of the large amount rays in this region causes a large focal point cloud. An explanation can be found in the fact that the velocity in the center of the first layer is also determined by rays from the deeper reflectors.

This feature is also obvious for the second reflector. In the thicker part of the second layer, the velocities are severely influenced by rays from the deeper reflector. This results again in large focal point clouds for the focal points of the lower part of the second reflector (light color on the left hand side). Although the velocities in the third layer are only influenced by the operators of the third reflector, the focal point clouds of this reflector have a rather high \(\Delta t_{rms}\) compared to the second layer. This is caused by the fact that this region is not well determined. Summarized, in areas that contain a lot of rays, \(\Delta t_{rms}\) increases. Even though this means that the residuals are higher, it does not mean that the velocities are not well determined.

The inversion result can be influenced and improved with the use of the focal point clouds. By weighting the rays during tomographic inversion by means of \(\Delta t_{rms}\), the inversion in the areas with large residuals, i.e., large focal point clouds, can be improved. This is shown in Figure 4.14. The tomographic inversion is started using the velocity of Figure 4.13, and four more tomographic updates are carried out. When the rays are weighted by the \(\Delta t_{rms}\) of the focal point they originate from (high \(\Delta t_{rms}\) means more weight), the focal point clouds become more similar for the different focal points (left hand side of Figure 4.14). But if the updating continues without any weighting, the \(\Delta t_{rms}\) patterns for the 3 layers stay the same (right hand side of Figure 4.14).

The focal point clouds will be used in sections 5.4, 5.5, and 6.2.
Figure 4.13: Left: the focal point clouds for the 3 layers after 6 updates. Right: the rms values of the traveltime misfit for the 3 layers after 6 updates. The light colors indicate a high misfit.
4.4 Evaluation of the inversion result using focal point clouds

Figure 4.14: The $\Delta t_{rms}$ for the 3 layers. Left: the after updating by using $\Delta t_{rms}$ weighting. Right: the result after updating without weighting.


4.5 Conclusions

In this chapter, the methodology of the third objective is addressed: explore the possibilities of the tomographic inversion of focusing operators (section 1.3). The principles of three different additional modules to the tomographic inversion are explained and the following conclusions can be drawn:

- It has been shown that a priori information on the model parameter values can be included in the tomographic inversion by using Kriging and taking them into account as extra constraints in the inversion.

- The inversion of P and S-wave operators is presented. The estimation of P and S-wave operators in mode converted data always results in operator pairs related to the same focal points. This aspect can be translated in a joint inversion approach. In this approach, a P and an S-velocity model are estimated simultaneously, in which the focal point coordinates are the same in both models.

- The concept of focal point clouds is introduced. This is a tool that can evaluate the adequacy of the velocity model to explain the data. The shape of the focal point clouds makes it possible to diagnose the origin of the residual between the observed focusing operators and the operators calculated in the velocity model. Velocity errors can be distinguished from focal point location errors, and anisotropy and 3D effects can be recognized.

The different additional modules, as described in this chapter, will be applied to both 2D and 3D synthetic and field data examples, which are presented in Chapters 5 and 6.
Evaluation of the methodology with synthetic data examples

5.1 Introduction

In the previous chapters the theory for the tomographic inversion of focusing operators has been discussed. In this chapter, this theory is applied to several 2D and 3D synthetic data sets. Testing the method on a synthetic data set generates a lot of insight into the performance of a new algorithm, because the availability of the ideal model makes it possible to evaluate the inversion results.

Tomographic inversion is first applied to two 2D models: the 'Turbidite' and the 'Marmousi' model. The 'Turbidite' model is suitable for testing the method, because layers and structures can easily be recognized. By means of this model the complete inversion procedure is discussed in detail. This procedure is also used in the next examples, without being addressed extensively. The 'Turbidite' model is also used to present the results of the additional features of the method that are presented in Chapter 4; the inclusion of a priori information, and the joint inversion of P and S operators are applied to this model. Whereas the structure of the 'Turbidite' model is quite clear, the 'Marmousi' model has a very high level of complexity. Therefore, the performance of the method will be examined on this data set.

After the method has been demonstrated on 2D examples, 3D synthetic data and models are considered. The first 3D example is a simple model, called the 'Antidune' model, on which the 3D method is tested. Next, the structurally complex 3D 'SEG/EAGE Salt' model is used to examine the method for more realistic cases. The concept of focal point clouds is also applied to both 3D datasets.
5.2 The 2D Turbidite model

5.2.1 Description of the model

The first 2D synthetic example is based on a velocity model containing a salt dome, fault structures beneath the dome, lateral and vertical velocity gradients within the layers, and a complex turbidite velocity structure with low velocities (Figure 5.1a). The model is 16 km wide and 4 km deep. The water layer at the surface has the lowest velocity: 1.5 km/s, and the salt dome has the highest velocity: 3.5 km/s. The velocity variation is not restricted to the layer interfaces in the model, but the structure variation can easily be recognized. In this way, estimated models can be visually gratified. Note that in the proposed tomographic inversion method the model is defined in terms of quadratic slownesses (section 3.4), but in this section, the input and the results of the inversion are presented as velocity models.

![Figure 5.1: True 'Turbidite' model. a) True velocity model (velocity scale in m/s), in which the 328 white dots indicate the focal points. b) Modeled focusing operators used as input for the inversion.](image)

5.2.2 Input data: the 2D focusing operators

The focusing operators are the input data to the inversion algorithm. When real field data are used, they are obtained from seismic by data-driven updating (section 2.4). In synthetic data, the 'observed' focusing operators are generated by ray-tracing in the velocity model. Rays are calculated, starting at the focal points in the subsurface (white dots in Figure 5.1a). For real seismic data the focal points are positioned at the reflectors, because they are related to reflection energy. Therefore, in the synthetic model, they are also located at the reflectors. Typically, there are fewer pickable reflections events below the salt and it would be realistic to assume that there are fewer focal points there. The 'observed' focusing operators are generated.
by ray-tracing from these focal points to a receiver grid the surface. The receivers have a spacing of 0.1 km and for each operator an aperture of 2 km is used. The calculated focusing operators, that form the input of the inversion algorithm, are shown in Figure 5.1b.

### 5.2.3 Initial velocity model

In order to start the tomographic inversion procedure, an initial velocity model is needed. It is assumed that no a priori knowledge about the subsurface is available, and we choose a homogeneous initial velocity of $2.0\text{ km/s}$ for the entire model. The initial velocity model is described by only a few grid-points, so that most points are presumably well-determined by the information available in the focusing operators. The initial velocity model is defined by 35 regularly distributed grid-points which are connected by Delaunay triangles. The velocity is interpolated within the triangles (Figure 5.2a).

![Initial velocity model](image)

**Figure 5.2:** Initial model. a) Initial homogeneous velocity model (velocity scale in $\text{m/s}$). The velocity is defined on the grid-points, which are connected by the Delaunay triangles. The velocity is linearly interpolated within the triangles. b) Initial homogeneous velocity model, in which the 328 white dots indicate (erroneously positioned) focal points. c) The initial focusing operators calculated in the model parameterization as presented in a), starting at the focal points as presented in b).
Also the initial focal point locations have been defined, indicated by the white dots in Figure 5.2b, but these points are mispositioned deliberately, as usually their location is unknown. The maximum error in the focal point location is 250 m, and the root mean square (rms) is 170 m for both the lateral position and the depth. This initial homogeneous velocity model with erroneous focal point locations is used to start the updating procedure. In this velocity model, initial focusing operators are calculated by ray-tracing from the focal points towards receivers at the surface (Figure 5.2c). The model parameters are updated by minimizing the difference between the 'observed' and the modeled focusing operators during the optimization.

### 5.2.4 Optimization of the model parameters

After building the initial model and generating the modeled focusing operators by ray-tracing, the updates of the model parameters can be calculated from the traveltime differences by means of the obtained system $\Delta \mathbf{d} = \mathbf{A} \Delta \mathbf{m}$ (see section 3.2.1). Instability of the optimization, due to low resolution of either the velocity or the focal point parameters, is avoided by slight damping of the equation (see equation (3.6.32)). The optimization is performed by the LSQR method. There are several aspects, which have already been denoted in section 3.5, in relation to this optimization:

- **Number of LSQR iterations;**
  By using the iterative LSQR method, the large basis-vectors, or 'modes', of the system are solved first (section 3.5). When the number of iterations equals the number of parameters, all modes of the system are resolved. By limiting the number of LSQR iterations, the optimization problem is inherently regularized as only the largest modes of the system will be resolved. As the system is dominated by the velocity parameters (Figure 5.3b, which will be discussed in one of the next points), it is decided that the number of LSQR iterations is dependent on the number of velocity parameters. The number of iterations in each update will be 60% of the number of velocity parameters. The maximum number of iterations will be 1000.

- **'Pseudo' singular-values;**
  The 'pseudo' singular-values can be obtained from the LSQR algorithm. These values represent the modes of the system that are solved. The 'pseudo' singular-value spectra for the different updates is shown in Figure 5.3a. The dark blue line indicates the singular-values for the first update. During this update, the number of LSQR iterations will be 60% of the number of velocity parameters. As only 31 velocity parameters are present in the optimization step, only 19 iterations are performed. As a result, only the 19 largest singular-values are obtained. The high singular-values show that the parameters are well determined. The other lines in Figure 5.3a refer to the next updates. The spectrum
changes for each update, because the optimization system changes for each update.

- **Resolution;**
  The resolution is also obtained from the LSQR algorithm. The resolution for the first update using 19 LSQR iterations is plotted in Figure 5.3b. This shows that the system is dominated by the velocity parameters, which are represented by the first 31 parameters. The next $2 \times 328$ parameters represent the focal point locations that play a secondary role, but will also be updated. In Figure 5.4a, the resolution of the velocity grid-points is plotted in the model, so that the distribution of the resolution becomes clear. It is shown that most parameters are over-determined (red color), but the parameters near the edges, at the bottom, and below the salt-dome have a low resolution (blue color).

---

**Figure 5.3:** Evaluation of the inverse problem. a) Pseudo-singular-value spectra for all updates. The spectrum changes most after updates 2, 3, 6, and 7. b) Resolution of the first update for 31 slowness parameters and $2 \times 328$ focal point parameters. c) Resolution of the final update for 875 slowness parameters and $2 \times 328$ focal point parameters.
5.2.5 Modification of the parameterization

The parameterization will be modified using the theory as presented in section 3.6. The strategy that is followed in the modification of the parameterization, is that the model will 'grow' towards its final solution. The initial velocity model is described by only a few grid-points that, as a consequence, are mainly well-determined. If there are under-determined points, they will be removed during the first update. During the next updates, grid-points will only be added to the model at regions with a high resolution. By doing this, the model can be described more accurately, while the parameters are still well-determined. There are three reasons for starting with a few grid-points and adding more points during the inversion: 1) the first updates in which the main trends of the velocity model are calculated are cheap, 2) it is more stable to add points than to remove points, as removing points will change the model considerably and will temporarily result in a deterioration of the model, and 3) the chance of ending the inversion in a local minimum is decreased by not using a fine parameterization from the beginning.

A question that arises is whether the parameterization should be modified after each update or not. When it is modified after each update, the final fine grid is obtained very quickly. This method turns out to converge to a minimum. However, an optimal minimum is not necessarily reached, as the non-linearity is not taken into account correctly by only performing one linearized update (see section 3.2.1). A more accurate strategy is to use the same parameterization during several updates, until the data misfit does not decrease anymore. When this minimum is reached for this (coarser) parameterization, it can be modified by adding grid-points. The final grid is obtained at the cost of more updates, but the non-linearity of the system is taken into account, resulting in a more accurate final model.

Grid-points are added based on the resolution obtained from the optimization algorithm. A resolution-threshold value is defined, and when the resolution of the three grid-points of one triangle is above the threshold value, a new grid-point can be added within the triangle. The threshold value in this research is defined to be 0.3. Figure 5.4 shows how the parameterization is modified during the 2nd, 3rd, 6th and 7th update. More and more grid-points are added in the regions where this is possible. Finally the model is described by 875 grid-points. The resolution values for these grid-points are plotted in Figure 5.3c. It is shown that the resolution relatively decreased and is more evenly spread for the 875 velocity parameters. Due to the fact that the velocity parameters are not highly overdetermined anymore, the focal point parameters start playing a more important role in the optimization.

The stop-criterion for adding grid-points is based on the singular-values, which are calculated during the optimization. Figure 5.3a shows how the singular-values develop during the updates. Major changes in the spectrum take place when the parameterization is modified. Due to the addition of grid-points, the highest singular-values start to approach a singular-value-threshold, which is defined to be 30, and the
5.2 The 2D Turbidity model

Figure 5.4: The modification of the parameterization for the ‘Turbidity’ model. On the left hand side the resolution of the model is plotted, and on the right hand side the resulting parameterization is plotted. From top to bottom the pictures show the modification of the initial parameterization, the parameterization of update 3, the parameterization of update 6, and the parameterization of update 7.

modification of the parameterization by adding grid-points to the model is stopped. When the singular-values of the final update (dark red line) are compared to the values of the first update (dark blue line), it is clear that the highest values decreased
and all values became more similar.

Summarized, the resolution is used for the qualitative aspect of the resolution dependent parameterization (where to add and remove grid-points), and the 'pseudo' singular-value spectrum is used for the quantitative aspect (when to stop adding points). So, the parameterization will be modified using the resolution and 'pseudo' singular-values, which are calculated during the optimization. As a consequence, this parameterization will be influenced by the different aspects of the optimization:

- **Number of LSQR iterations;**
  By limiting the number of LSQR iterations, only the largest modes of the system will be resolved, and the relative resolution values of the parameters will be relatively small. As a result, only a few grid-points can be added and the parameterization will stay coarse. This, again, reveals that only the largest modes of the system are solved, and the coarse model is the best result that can be obtained by this limited number of iterations. On the other hand, when the number of iterations keeps growing during the updates, the number of parameters keeps increasing and the parameterization becomes finer. However, this can not go on infinitely and therefore, the 'pseudo' singular-values serve as a stop-criterion. In this 'Turbidite' example, the maximum number of LSQR iterations is 1000, which did not result in a limitation of the number of parameters. The addition of parameters was stopped by the 'pseudo' singular-values.

- **'Pseudo' singular-values;**
  The 'pseudo' singular-values are used as a stop-criterion by defining a singular-value-threshold. When this value is high, the addition of grid-points is stopped in an early stage and the parameterization will stay coarse. When this value is low, more points will be added. It is important not to define too low a threshold, because this can result in an excessive addition of grid-points that may be under-determined.

- **Resolution;**
  Another threshold value that plays a role is the resolution value above which grid-points can be added. As the relative resolution is related to the number of LSQR iterations, this threshold value should be defined in relation to the number of LSQR iterations. When the number of iterations is 60% of the number of parameters the average resolution value will be 0.6 (see equation(3.6.37)). If in this case, the resolution-threshold value is defined to be 0.9, hardly any grid-points will be added. On the contrary, when the value is 0.1 a lot of points will be added and the risk exists that points added turn out to be under-determined.
5.2 The 2D Turbidite model

5.2.6 Result

The final result after 15 updates is shown in Figure 5.5. The parameterization of the final model is shown in Figure 5.5a. Many points are located in regions of high resolution and less points define the regions of low resolution (compare to Figure 5.4a). For example, around the salt dome and above the turbidite structure, a lot of rays are available. This results in a more dense parameterization. The bottom, the edges, and the sub-salt region only contain a few ray-paths, resulting in fewer grid-points and, as a consequence, larger triangles. In general, the resolution decreases with depth, because at larger depths, given the limited surface recording geometry, the angle coverage will decrease. The obtained parameterization reveals where it is necessary to add extra information. The coarse parameterization below the salt shows that in the sub-salt area focal points should be added. However, data that is related to this area is generally not available, because these parts might not have been illuminated by the acquisition geometry.

![Figure 5.5](image)

*Figure 5.5: The tomographic inversion result after 15 updates. a) The parameterization of the updated model. b) The updated velocity model containing the updated focal points. c) Comparison between the true focal point locations (open dots) and the updated focal point locations (closed dots). d) Result of a shot record migration of the data using the updated velocity model.*

The updated velocity model in Figure 5.5b shows that the salt dome, the turbidite structure, and the velocity gradients have been resolved. However, the entire lower
Evaluation of the methodology with synthetic data examples

![Graph of number of velocity parameters versus update number.](image)

Figure 5.6: Important quantities during the updates. a) The number of parameters. b) The log-plot of the rms of the data misfit (difference between the travel times of the 'observed' and the modeled focusing operators). c) The rms of the focal point errors for the lateral coordinate (green dashed line) and the depth coordinate (red solid line).

...part of the model is not resolved in detail, due to the coarse parameterization in this region. Indeed, this is caused by the lack of information in the data in this part, so that only a small number of grid-points can be defined. In the result it can also be observed that the edges between the different features are smoother than in the true model. This is inherent in tomographic inversion, as this will generate a low frequency approximation of a high frequency true model.

Besides the velocities in the subsurface, the focal point locations are also updated. Figure 5.6c shows how the location errors of the focal points decrease during the updates. When the updated focal point locations are compared with the true focal point locations (Figure 5.5c), it can be seen that they are very accurately positioned. The final rms error is 26 m, in both depth and lateral position. Usually, as the reflection angle decreases with depth, the depth-velocity ambiguity increases with depth. This means that after optimization a sufficient data-fit can be obtained by an updated model in which an error in depth is compensated by an error in velocity. This would have shown up as incorrectly positioned focal points in the updated result, but this effect is not present in this example. This might be due to the well determined optimization problem, where only a few grid-points are present in the deeper parts, due to the low resolution. An effect that plays a role in this example is that the area under the salt dome is not well illuminated by the acquisition geometry.
As a result, the focal points in this region are not well positioned. Although the focal points and the velocities are different parameters that describe the subsurface, together they optimally describe the structures. The turbidities, the salt-dome, and even the sub-salt faults are outlined by the focal points.

The data misfit evaluates the result quantitatively. Figure 5.6b shows how this misfit decreases during the updates. The rms difference between the traveltimes of the 'observed' and modeled focusing operators will finally be 3.7ms, which is within seismic accuracy (this is approximately 4ms, due to picking errors, wavelet length, etc.). The remaining discrepancy between the two operators is caused by the difference in parameterization between the true and the updated model. The rms data misfit finally converges to this level from which it does not decrease much anymore. When the iterative updating continues when this level is reached, it might even occur that the rms misfit increases again. This is caused the ray-shooting method which is used for the forward modeling of focusing operators. Continuing the updating will make the velocity anomalies more extreme. As a result, the rays will get more extreme curvatures, and it will become harder to reach a receiver. Finally, fewer receivers will be reached by using the shooting method and the rms value of the traveltine misfit can eventually increase.

As the purpose of velocity model estimation is not to reveal the exact geological features in the model, but to provide a way to explain seismic wavefield propagation, a conventional shot record migration using the updated velocity model is performed. The result is shown in Figure 5.5d. All structures are clearly visible. However, the sub-salt region and the deepest reflector are not well resolved due to a lack of accuracy in the velocity model, which is caused by a lack of illumination in this area. This is also revealed by the parameterization, which is coarse at these locations.

### 5.2.7 Post-updating step

The tomographic inversion of focusing operators, as presented so far, produces a result in a data-driven way, without any input of the user. Afterward, the result obtained can be reviewed by different methods. The imprint of the irregular parameterization can be evaluated by uniform re-gridding. Besides this, the structures revealed in the tomography result can be interpreted and further processed.

**Uniform re-gridding**

Any kind of parameterization causes an imprint on the model that it describes. Only models described by the finest gridding will have the highest resolution. Nevertheless, this is not realistic because such a problem will be under-determined and cannot be resolved. To consider the imprint of the current irregular parameterization of the obtained model (Figure 5.5a), we redefine the model by a different (in this case
regular) grid and perform one more update. This is shown in Figure 5.7. Note that the complete updating procedure is far more expensive than the updating procedure with the previous grid, as more parameters should be optimized. In addition, the problem is under-determined for a lot of parameters, because a lot of points have low resolution or even no crossing ray-paths at all (Figure 5.7b). This does not cause a problem, because the updated model (Figure 5.5a) serves as an initial model, which is already close to the accurate model. Furthermore, the algorithm will only update the points that are determined by the ray-paths. The other points will keep the velocity value they already had. The resulting velocity model after optimization using the regular parameterization is presented in Figure 5.7c. When it is compared to the result obtained using the irregular parameterization (Figure 5.5b, repeated in 5.7d), it can be seen that all the structures that were visible already, stay the same. All the solvable structures have already been obtained while using the irregular parameterization. However, the velocity contrasts due to the imprint of the larger triangles beneath the salt dome is reduced. Note that when this regular grid would have been used from the beginning of the tomographic inversion procedure, it would have been difficult to obtain an accurate result, due to the under-determinacy of the problem.

![Figure 5.7: Post-updating by re-sampling. a) The regular parameterization of the velocity model. b) The resolution for this parameterization. c) The result of updating by using the regular grid. d) The result by using the irregular grid (Figure 5.5b is repeated here).](image-url)
Interpretation of structures

The parameterization of the final result as shown in Figure 5.5a is totally determined by the available data. The result reveals the structures of the subsurface: the velocities indicating possible geological regions and the focal points showing the position of the reflectors in the subsurface. The structure revealed can now be used for interpretation of the subsurface. Normally, when a global parameterization is used, this structure has already been defined by the user before the optimization, which might lead to an incorrect initial parameterization. The geological structures in Figure 5.8a are interpreted by using the focal points, and the velocities within them are averaged, resulting in the model as shown in Figure 5.8b. Note that any kind of velocity interpolation and averaging can be done. Optionally, after interpretation, the optimization of the velocity model can be continued (Figure 5.8c and d). In the new result the interpreted structures are visible, and the traveltimes of all focusing operators can be explained.

Figure 5.8: Post-updating by interpretation. a) The result of optimization using the irregular grid (Figure 5.5b is repeated here). b) Regions defined based on interpretation of velocities and focal point locations in a. c) The new regular parameterization of the interpreted model in b. d) Result after updating using the regular parameterization of the interpreted model.
5.2.8 Including a priori information

Although in the proposed method all the available information in the data can be translated to a model, it might be desirable to include a priori information on the subsurface in the inversion. The theory of the inclusion of a priori information has been described in section 4.2. This theory is used in this subsection, which presents an example of using a priori information in the tomographic inversion for the 'Turbidite' model. Figure 5.9a shows the samples taken from the true velocity model of the subsurface, which correspond to four well log measurements and the a priori knowledge of the water layer. By applying geostatistics to these samples the a priori model and its accuracy are obtained (Figure 5.9b,c).

Figure 5.9: A priori model. a) Samples (white dots) taken from the true velocity model. b) Result after Kriging based on the samples in a. c) Accuracy of the Kriged result indicated by resolution. d) Resolution during tomographic inversion.

For the tomographic inversion the same strategy as used before is applied, but now, the a priori information and its accuracy are included in the inversion. The final result after 15 updates is shown in Figure 5.10. Due to the addition of extra information, the resolution is locally increased, and the parameterization is finer around the boreholes. In particular in the deeper areas, where originally only a little information from the focusing operators is present, the parameterization is refined. The
velocities around the borehole are adopted from the a priori information. However, the velocities in the other regions are determined by the conventional focusing operator data and show a lot of similarity with the model obtained by the inversion of traveltimes only. The improvement of adding a priori information becomes more obvious when the updated focal point locations (Figure 5.10c) are compared to the focal point locations obtained by the inversion of traveltimes only (Figure 5.10d). Especially in the deeper regions below the salt, next to the borehole at 1 km and 5 km, the focal points are located at the right position. The final rms traveltime misfit is 5 ms, which is higher than the misfit in the inversion without a priori information. This is caused by the fact that the optimization was not only done with respect to the traveltimes of the focusing operators, but also with respect to the a priori information about the velocities. This prevents the convergence to a minimum error in the traveltime data.

Figure 5.10: The tomographic inversion result by including a priori information, as shown in Figure 5.9. a) The parameterization of the updated model. b) The updated velocity model containing the updated focal points. c) Comparison between the true focal point locations (open dots) and the updated focal point locations (closed dots). d) The focal point locations after updating without using the a priori information (Figure 5.5c is repeated here).
5.2.9 Joint inversion of P and S-wave focusing operators

From converted seismic data two types of focusing operators can be obtained: the P-wave operators and the S-wave operators (section 4.3). These operators can be inverted separately in order to obtain a P and an S-velocity model. However, they can also be inverted simultaneously, as both operator types share the same focal point. This theory has been explained in section 4.3. The separate and joint inversion of P and S-wave operators of the 'Turbidite' model, will be presented in this subsection.

![Diagram](image)

**Figure 5.11:** a) True P-velocity model. The white dots represent the focal point locations. b) True S-velocity model. c) Initial P-velocity model with erroneously positioned focal points. d) Initial S-velocity model. Note that the focal points in the P and S-velocity model are the same.

The P-velocity model that is used is the same as in the previous subsections (Figure 5.11a). The corresponding S-velocity model is shown in Figure 5.11b. The S-wave velocity is at least a factor $\sqrt{2}$ lower than the P-wave velocity. The focal points are the same for both velocity models. From these focal points the 'observed' P and S-wave operators are calculated in the P and S-velocity model respectively. The aperture of the S-wave operators will be smaller than the aperture of the P-wave operators (subsection 4.3.4). Therefore, the S-wave operator aperture is defined to be 1.2 km, while the P-wave operator aperture is 2 km.
To start the inversion procedure, both an initial P and S-velocity model have to be defined. The initial models are defined by different homogeneous velocities (Figure 5.11c,d), because the S-wave velocity is expected to be lower than the P-wave velocity. Still, the initial focal points in both models are the same. In both velocity models the initial focusing operators are calculated by ray-tracing. The optimization is done by minimizing the difference between the ‘observed’ and the modeled focusing operators.

Figure 5.12: Results after a separate inversion of P waves and S waves. a) The final parameterization of the P-velocity model. b) The final parameterization of the S-velocity model. c) The P-velocity model including the focal point locations. d) The S-velocity model including the focal point locations.

First, the inversion is done for the P and S-velocity model separately. The results after 15 updates are presented in Figure 5.12. The updated P-velocity model has already been presented in subsection 5.2.6. The misfit of the focal points for the P-velocity model was 26m in both directions, and the final data misfit was 3.7ms. When the updated S-velocity model is compared to the P-velocity model it is obvious that the parameterization of the S-velocity model is coarser, due to the lower resolution. This is caused by the smaller amount of available information, which is a result of the smaller aperture of the S-wave operators. Nevertheless, like in the updated P-velocity model, all structures are also solved in the updated S-velocity model.
model. The misfit of the focal points is 39 m in the lateral direction and 40 m in depth, which is larger than in the P-velocity model. The final rms traveltime misfit is 3.3 ms.

![Diagram](image)

**Figure 5.13:** Result after joint inversion of P and S waves. a) The final parameterization of the P-velocity model. b) The final parameterization of the S-velocity model. c) The P-velocity model including the focal point locations. d) The S-velocity model including the focal point locations. Note that the focal point locations in c) and d) are the same.

Next, the joint inversion of the P and S-velocity model is done, in which the P and S-wave operators are coupled by the focal points. The results are shown in Figure 5.13. Again, for both models the velocity structure is resolved and the S-velocity model has a coarser parameterization than the P-velocity model. The obtained models look slightly different than the results from the separate inversion. In order to determine which result is better, the updated focal point locations are considered (Figure 5.14). The misfit of the focal points is 24 m in the lateral direction and 26 m in depth, which is only a small improvement compared to the separate P inversion, but a large improvement compared to the separate S inversion. When the points of the separate S inversion are compared to the points of the joint inversion, it is obvious that the P-wave operators contribute a lot of information on the focal points, from which the S-velocity model benefits. On the other hand, the contribution of the S-wave operators to the P-velocity model is mainly visible in the part below the salt
5.2 The 2D Turbidite model

dome. By introducing the S-wave operators, the focal points are more accurately positioned in the sub-salt region. The final data misfit is 3.8 ms, which is comparable to the results of the separate inversion.

![Image of focal point locations after the separate inversion of P and S waves, and after the joint inversion of P and S waves. Dots represent the updated locations, and the open circles represent the true locations.](image)

**Figure 5.14:** Focal point locations after the separate inversion of a) P and b) S waves, and c) after the joint inversion of P and S waves. Dots represent the updated locations, and the open circles represent the true locations.

In conclusion, the joint inversion of P and S-wave operators can result in better velocity models, than obtained from a separate inversion. In particular the S-velocity model benefits from this joint inversion.
5.3 The 2D Marmousi model

5.3.1 Description of the model

The second 2D synthetic example is related to the Marmousi model. The geometry of the Marmousi model is based on a profile through the North Quenguela trough in the Cuanza basin in Angola [Versteeg, 1993]. The Marmousi model contains velocity distributions with strong horizontal and vertical gradients (Figure 5.15). Some very thin layers, anticlinal structures, high velocity wedges, and complicated faults are present. The velocity is between 1.5 km/s and 5.5 km/s. The geometry and the velocity model result in complex seismic data which require advanced processing techniques to obtain a correct image of the subsurface. In this section the performance of the tomographic inversion of focusing operators is tested on the complex ‘Marmousi’ model. In total 571 focal points, which are chosen on the reflectors, were defined. The points are dense in the shallower regions, and more rare in the deeper regions, especially in the regions below the high velocity wedges. Rather than deriving the focusing operators from the synthetic seismic data, the focusing operators to be inverted were directly calculated in the model. The traveltimes from the focal points to the receivers at the surface are calculated by means of the finite difference eikonal solver [Vidale, 1988]. The receivers have a spacing of 0.1 km. In order to keep the operators realistic, a maximum aperture of 2 km is defined. The focusing operators form the data input of the inversion.

Figure 5.15: True ‘Marmousi’ velocity model in which the 571 white dots indicate focal points.
5.3.2 Initial velocity model

The other input that is required, to start the inversion procedure, is the initial velocity model. It is assumed that the velocity in the subsurface is increasing with depth from 1.5 km/s at the surface to 3.5 km/s at 3 km depth (Figure 5.16a). The initial velocity model is described by only a few grid-points, which are connected by the Delaunay triangulation. The velocity is linearly interpolated within the triangles. Also the initial estimates of the focal point locations have been defined. These are indicated by the white dots in Figure 5.16b. When we compare them to the true focal points (Figure 5.15), it is clear they are positioned incorrectly. The maximum error in both lateral position and depth is 250 m. The velocity model and the initial focal point locations are used to start the updating procedure. By calculating rays in this model, from the focal points towards the surface, the initial focusing operators are modeled. The optimization of the model parameters is done by minimizing the difference between the 'observed' and the modeled focusing operators.

![Figure 5.16: Initial model. a) Initial velocity model, with a velocity increasing with depth from 1.5 km/s to 3.5 km/s. b) Initial velocity model, in which the 571 white dots indicate the (erroneously positioned) focal points, with a maximum error of 0.25 km.](image)

5.3.3 Optimization of the model parameters

The optimization strategy is analogous to the method that was used for the 'Turbidite' model. The number of LSQR iterations will be 60% of the number of velocity parameters. As a result only 30 iterations are performed during the first update. The dark blue line in Figure 5.17 indicates the 'pseudo' singular-value spectrum for the first update, which evidently contains only 30 values. The high singular-values show that the parameters are well-determined. The large difference between the singular-values show that some parameters are much better determined than other parameters.
5.3.4 Modification of the parameterization

The initial velocity model is described only by few grid-points that are mainly well-determined, and the potentially under-determined points will be removed in the first update (Figure 5.18). During the next updates, the parameterization will only be modified when the data misfit, using the current grid, does not decrease anymore. If this is the case, grid-points will be added to the model at regions with a high resolution. For the inversion of the 'Marmousi' model, the parameterization is modified after update 2, 7, 10 and 12 (Figure 5.18). The parameterization of the final model is defined by 1295 grid-points. Many of these points are located in regions of high resolution and fewer points define the regions of low resolution. For example, the bottom of the model only contains a few ray-paths, which results in fewer grid-points and, as a consequence, larger triangles. The final resolution is more similar for all parameters. The obtained parameterization reveals where it is necessary to add extra information. For example, the coarse parameterization in the lower part shows that in this area focal points should be added.

The 'pseudo' singular-values, calculated during the optimization, serve as stop-criterion for the addition of grid-points. Figure 5.17 shows how the singular-values spectrum develops during the updates. Major changes in the spectrum take place when the parameterization is modified (update 2, 7, 10 and 12). When the singular-values of the final update (dark red line) are compared with the values of the first update (dark blue line), it is clear that the highest values decreased and finally start to approach a singular-value-threshold value, which is defined to be 30. At this point the modification of the parameterization by adding grid-points to the model is stopped. The final singular-values are more similar, which also indicates that all parameters are more equally well-determined.

Figure 5.17: 'Pseudo' singular-value spectra for all updates.
Figure 5.18: The modification of the parameterization for the Marmousi model. On the left hand side the resolution of the model is plotted, and on the right hand side the resulting parameterization is plotted. From top to bottom, the pictures show the modification of the initial parameterization, the parameterization of update 7, the parameterization of update 10, and the parameterization of update 12.
5.3.5 Result

The final result after 15 updates is shown in Figure 5.19. All geometries and velocity gradients have been resolved, and the updated model looks like a smoothened version of the true model (Figure 5.15). Even the thin layers with a thickness of 150 m appear as velocity anomalies. However, the entire lower part of the model is not resolved in detail, due to the lack of information from the data in this part. This particularly affects the appearance of the high velocity wedges. These wedges actually do appear, but the shape is not accurately resolved. This inaccurate shape is not only caused by the fact that they are positioned in the lower part of the model, but also caused by the fact that they are positioned at the edges of the model, which effectively gives a smaller aperture.

Figure 5.19: The inversion result after 15 updates. a) The parameterization of the updated model. b) The updated velocity model containing the updated focal points. c) Comparison between the true focal point locations (open dots), and the updated focal point locations (closed dots). d) Result of a shot record migration of the data using the updated velocity model.

Besides the velocities in the subsurface, the focal point locations are also updated. When the updated focal points are compared to the true focal point locations (Figure 5.19c), it can be seen that they are reasonably well positioned and nicely show the
structure of the reflectors in the model. Together, the focal points and the velocities give the most complete picture of the model. The faults, the thin layers, and the anticlinal structure are all outlined by the focal points. The final rms error is 52 m in the lateral position and 65 m in the depth position. This error is caused particularly by the deeper focal points, which are inaccurately positioned due to the depth-velocity ambiguity which increases with depth. Despite the coarse parameterization in the deeper regions this could not be prevented.

The velocity depth ambiguity can also be observed in Figure 5.20c. This figure shows how the location error of the focal points develops during the updates. The error of the lateral position of the focal points decreases continuously, but the error of the depth position shows a slow increase after the 3rd update. This is also an indication that the inversion cannot find a unique solution for the depth and the associated velocity. The data-misfit evaluates the result quantitatively. Figure 5.20b shows how the data misfit behaves during the updates. A slight increase of the misfit can be observed, when the parameterization is modified during updates 7 and 10. The root mean square (rms) difference between the traveltimes of the 'observed' and modeled focusing operators will finally be 3.0 ms.

![Figure 5.20](image)

**Figure 5.20:** Important quantities during the updates. a) The number of velocity grid-points. b) The log-plot of the rms of the data misfit. c) The rms of the focal point location error for the lateral coordinate (green dashed line) and the depth coordinate (red solid line).
Again, the velocity model resulting from the tomographic inversion is used for conventional migration of the seismic data. The result is shown in Figure 5.19d. All structures are clearly visible, but the deeper layers and anticlinal structure are less well resolved due to a lack of accuracy in the velocity model. When this result is compared to the migration using the true velocity model (Figure 1.2a), the result in the upper part is similar; the faults, the thin layers, and even the anticlinal structures at 2 km depth appear in the same way. Only the anticlinal structure in the deepest part, and the shape of the high velocity wedges are less well visible in the migration that is carried out using the velocity model from tomographic inversion.
5.4 The 3D Anticline model.

In the previous sections, the method has been applied successfully to 2D problems, so now it can be examined on 3D problems. The strategy that is used is similar for both problems. The 3D method will first be tested on a simple model in this section.

5.4.1 Description of the model

The first synthetic 3D example is based on a simple velocity model containing 5 reflectors (Figure 5.21a). The upper flat reflector represents the water bottom, which is followed by a dipping reflector, two curved reflectors, and finally a flat reflector at 2.8 km. The two curved reflectors have an anticlinal structure; hence, this model is called the 'Anticline' model. The model is 6 km wide, 6 km long and 3 km deep and is constructed by connecting the grid of the reflectors, which results in cubes. Each cube is divided in 6 tetrahedra (Figure 5.21b). The velocity is increasing with depth. The velocity of the water layer is 1.5 km/s, and the velocity of the deepest layer is 3.5 km/s (Figure 5.21c). This is the true model that is used to evaluate the tomographic inversion of focusing operators for 3D problems.

5.4.2 Input data: the 3D focusing operators

The focusing operators form the input data for the inversion algorithm. Normally they can be obtained from 3D seismic data by using the algorithm as presented in section 2.4.2. For this synthetic experiment they are calculated by forward modeling in the true model. The focal points are positioned at the reflectors in the model, on which they are regularly positioned, with a spacing of 300 m in both directions. The total number of focal points is 2000 (Figure 5.21d). Next, rays are traced from these focal points towards a regular receiver grid at the surface, where the receivers also have a spacing of 300 m (Figure 5.21e). This is obviously very coarse, but it will be used to test the tomographic inversion algorithm in 3D. The maximum aperture of the focusing operators is 2 km. The ray-tracing in the true model generates the 'observed' traveltimes of the focusing operators, which form the input for the inversion algorithm (Figure 5.21f).
Figure 5.21: The true ‘Antidote’ model. a) This model is 6 km wide, 6 km long, and 3 km deep, and contains 5 reflectors. b) The parameterization of the true model is defined by connecting the layers by cubes, which are divided into six tetrahedra. c) The velocity within the layers of the true model is increasing with depth from 1.5 km/s to 3.5 km/s. d) Each reflector contains 20 × 20 focal points with a spacing of 300 m in both directions. e) The rays from one focal point resulting in a focusing operator with a maximum aperture of 2 km, and f) the corresponding travel times at the surface in seconds.
5.4 The 3D Anticline model.

5.4.3 Initial velocity model

For the tomographic inversion procedure an initial velocity model is needed. The initial model is described by only a few grid-points, such that most initial points are well-determined during the optimization. In the case considered here, $7 \times 7 \times 4$ regularly distributed grid-points are defined, which are connected by cubes, that are divided in tetrahedra (Figure 5.22a).

![Initial velocity model](image)

**Figure 5.22:** The initial model. a) The initial model is defined by $7 \times 7 \times 4$ grid-points with 1 km spacing in each direction. These points are connected to form cubes, which are divided into tetrahedra. b) Slices through the initial model. Note that slicing the tetrahedra results in a triangle representation within the slices. The model has a homogeneous velocity of 2.2 km/s. c) The 2000 focal points are erroneously positioned in the initial model, with a maximum error of 25.0 m in each direction. d) The rays from one focal point resulting in a focusing operator, and e) the corresponding traveltimes at the surface in seconds.
The velocity is defined at the grid-points and interpolated within the tetrahedra. As we assume that no a priori knowledge about the model is available, we choose a homogeneous initial velocity of 2.2 km/s for the entire model (Figure 5.22a,b). The focal point locations are not known either, so they are erroneously positioned in the velocity model, with a maximum error of 250m in the x, y and z directions (Figure 5.22c). The error is random in the x and y direction, while the error is consistent in the z-direction. The focal points are positioned too deep, and this error increases with depth. This consistent depth error is introduced to examine the depth-velocity ambiguity. The initial velocity model and focal point locations are required to start the updating procedure. The initial operators are calculated by ray-tracing in this model (Figure 5.22d,e). The optimization of the model parameters is performed by minimizing the difference between the ‘observed’ and the modeled traveltimes of the operators.

5.4.4 Optimization of the model parameters

Also in the 3D case, the optimization is done by the LSQR method. The maximum number of iterations is defined to be 1000. During each update, the number of iterations will be 60% of the number of velocity parameters (196 velocity grid-points). Therefore, only 100 LSQR iterations are carried out during the first update. This results in 100 ‘pseudo’ singular-values (dark blue line in Figure 5.23). Apparently there is a large difference in singular-values.

Figure 5.23: ‘Pseudo’ singular-value spectra for all updates.
5.4 The 3D Anticline model.

5.4.5 Modification of the parameterization

The resolution that is calculated during the optimization will be used for the modification of the parameterization. Also in the 3D case the model will ‘grow’ towards its final solution. The initial velocity model is described by only a few grid-points that are, as a consequence, well-determined (see singular-value spectrum in Figure 5.23).

\[ \begin{array}{cc}
\text{Resolution} & \text{Parameterization} \\
\end{array} \]

![Resolution and Parameterization](image)

**Figure 5.24:** Slices through the model for the first update. The left hand side represents the resolution slices. The right hand side represents the slices through the resulting parameterization of the velocity model.

When the resolution at the grid-points in the initial model is considered (left hand side in Figure 5.24), the grid-points in the center of the model have generally high resolution, which is caused by a high data density. Near the boundaries of the model the grid-points have a low resolution. This is due to the fact that only a few rays are passing through these regions, resulting in a low data density. Using this resolution for the modification of the parameterization results in a new parameterization as shown on the right hand side in Figure 5.24. During the next iterations, the parameterization will only be modified when the data misfit, resulting from the current parameterization, does not decrease anymore. If this is the case, grid-points will be added in the high resolution regions. For this model, the parameterization is
Evaluation of the methodology with synthetic data examples

modified during update 2, 11, and 12. The resolution value for adding grid-points equals 0.3. Finally, 3417 grid-points are used to describe the model (Figures 5.25 and 5.28a).

![Parameterization and Resolution](image)

**Figure 5.25**: Slices through the model for the final update. The left hand side represents the slices through the velocity model. The right hand side represents the slices through the corresponding resolution.

The threshold value for the 'pseudo' singular-values is 100. When the largest value is smaller than 100, the addition of grid-points is stopped. The evolution of the singular-value spectrum is shown in Figure 5.23. Also in this case, major changes in the spectrum take place when the parameterization is modified. When the 'pseudo' singular-values of the final update (dark red line), are compared to the values of the first update (dark blue line), it can be observed that the variation of the values decreased due to a better distribution of grid-points.

Note that the shape of the spectrum differs from the 2D 'pseudo' singular-value spectra. In 3D more values reside at the higher end of the spectrum, while in the lower end only a few values are present. In the 2D case the opposite occurs. This implies that the 3D problem is basically better determined than the 2D problem.
5.4 The 3D Anticline model.

5.4.6 Results

The velocity model result after 25 updates is shown in Figure 5.27. A slice through each layer and the focal points associated with that layer are shown. The homogeneous initial velocity model has changed into a model that represents all velocity changes and the separate layers show up clearly. For all layers, the result for both velocity and focal point locations is quite similar to the true model.

When the focal points are considered separately (Figure 5.26) it turns out that the initially erroneously positioned focal points tend to align and show up as five separate layers. Figure 5.28c presents the root mean square values of the focal point location error. It decreases continuously during the updates for both the x, y and the z direction, until the final misfit is 23 m for the x and y direction and 25 m in depth. The location error increases with depth, which is due to the decrease of resolution with depth.

The traveltime misfits decrease during the updates. The final rms misfit is 5.7 ms. This is due to the parameterization of the model, that is not able to describe the layers exactly. As a consequence the precise ray-paths and traveltimes cannot be obtained in the updated velocity model.

![Figure 5.26](image-url)

*Figure 5.26: a) The initial erroneously positioned focal points. b) The focal points after one update. c) The focal points after the final update. d) The true focal points.*
Figure 5.27: Comparison of true model (left) with the model that is the result after 25 updates (right). The five layers of the model are considered. The blue dots represent the focal points.
5.4 The 3D Anticline model.

Figure 5.28: Important quantities during the updates. a) The number of parameters defining the velocity of the model. b) The log-plot of the rms of the data misfit (difference between the traveltimes of the ‘observed’ and the modeled focusing operators). c) The rms of the focal point location error for the x coordinate (green dashed line), the y coordinate (blue dashed-dotted line), and the depth coordinate (red solid line).
5.4.7 Focal point clouds

The focal point clouds can evaluate the adequacy of the velocity model to explain the data (section 4.4). When the clouds of the initial model are analyzed (Figure 5.29a), it can be observed that the velocity model does not explain the focusing operators, because the cloud does not appear as a point at the location of the focal point. All kind of errors occur, due to the random error in the initial focal point location. A structural error in velocity is revealed by the flat shape of the focal clouds of the shallowest and the deepest focal points. In the depth-range between 1 km and 2.2 km, more curved clouds appear, which shows that an initial estimate of $2.2 \, \text{km/s}$ is reasonably correct for this region. During the tomographic inversion, the velocity model is improved after each update. The focal point clouds of the final model (Figure 5.29b,c,d) show that this velocity model is capable of explaining the focusing operators. The clouds are concentrated at the focal point locations.

![Focal point clouds for the Anticline model](image)

**Figure 5.29:** Focal point clouds for the Anticline model. The red circles indicate the focal point locations. The blue dots form the clouds. a) A selection of focal point clouds for the initial model. b) A selection of focal point clouds for the final model. c) Focal point clouds of the fourth layer for the final model. d) Close-up of some focal point clouds of the fourth layer as shown in c).
5.5 The 3D SEG/EAGE Salt model.

5.5.1 Description of the model

The SEG/EAGE Salt model has been designed to test various seismic applications [Aminzadeh et al., 1996]. The model contains a salt-dome surrounded by sand-bodies (Figure 5.30). The main characteristics of the salt are the plunging stock and a fault next to the stock. The velocity of the salt is 4.0 km/s. The velocity of the sand-bodies ranges from 1.5 km/s to 2.5 km/s. Together, the sand-bodies form a smooth velocity area. They include some thin layering which can be reflective. At 3.5 km depth a flat sand layer is present which can serve as a reference. In the sand-bodies beneath the salt-dome two major faults occur.

Figure 5.30: a) The true salt model is 10 km wide, 10 km long, and 4 km deep, and contains a complex salt-dome. b) The velocity of the salt is 4.0 km/s and the velocities of the surrounding sand-bodies range from 1.5 km/s to 2.5 km/s. c) The 4997 focal points in this model are related to places where reflection can occur. Therefore, the separate reflectors and faults are visible.
5.5.2 Input data

The focal points, required for the generation of the focusing operators, are positioned at the reflectors in the model. They are regularly positioned at each reflector, with a spacing of 480 m in both directions. The total number of focal points is 4997 (Figure 5.30c). From these points forward modeling of the focusing operators is performed towards a regular receiver grid at the surface, where the receivers have a spacing of 360 m. This is carried out using the finite difference eikonal solver method [Vidale, 1988]. The maximum aperture of the focusing operators is 2 km. The ‘observed’ traveltimes of the operators form the input for the inversion algorithm.

5.5.3 Initial velocity model

The initial model is described by only 11 × 11 × 9 regularly distributed grid-points, which are connected by cubes, that are divided into tetrahedra (Figure 5.31a). The velocity is defined at the grid-points and interpolated within the tetrahedra. As we assume that no a priori knowledge on the model is available, we choose a homogeneous initial velocity of 2.2 km/s for the entire model (Figure 5.31a,b). The focal point locations are not known, so they are randomly positioned in the velocity model, with a maximum error of 250 m in both the x, y and z directions (Figure 5.31c). The error is random in the x and y direction, while the error is increasing in the z-direction. The initial focusing operators are calculated by ray-tracing in this model from the initial focal point locations to the receivers at the surface.

5.5.4 Optimization of the model parameters

The maximum number of LSQR iterations is defined to be 1000, in order to prevent the inversion from becoming too expensive. During the optimization in the first update, 653 LSQR iterations are performed, since the number of iterations should be 60% of the number of velocity parameters (1089 velocity grid-points). This results in 653 'pseudo' singular-values (dark blue line in Figure 5.32). Most singular-values are high, which indicates that the grid-points of the initial coarse grid are generally well determined. Nevertheless, a few parameters are underdetermined and will be removed during the first update.
5.5 The 3D SEG/EAGE Salt model.

**Figure 5.31:** The initial model. a) The initial model is defined by $(11 \times 11 \times 9)$ regularly distributed grid-points. These points are connected to form cubes, which are divided by tetrahedra. b) Slices through the initial model. The model has a homogeneous velocity of $2.2 \text{ km/s}$. c) The 4997 focal points are erroneously positioned in the initial model, with a maximum error of 250 m in each direction.

**Figure 5.32:** 'Pseudo' singular-value spectra for all updates.
5.5.5 Modification of the parameterization

The resolution that is calculated during the optimization will be used for the modification of the parameterization. When the resolution at the grid-points in the initial model is inspected (Figure 5.33), the grid-points in the center of the model have mainly a high resolution and near the boundaries of the model the grid-points have a low resolution. Using this resolution for the modification of the parameterization results in a new parameterization as shown in the upper part Figure 5.33.

During the next iterations the parameterization will only be modified when the data misfit does not decrease anymore (Figure 5.38b). For this model the parameterization is modified during updates 2, 13, and from update 15 to 20 (Figure 5.38a). As a consequence, the final model is described by 4411 grid-points, for which the resolution is more uniform (lower part Figure 5.33. The evolution of the singular-value spectrum is shown in Figure 5.32. The shape of the spectrum is similar to the 'pseudo' singular-value spectra shown for the 3D 'Anticline' model (Figure 5.23). The major changes in the spectrum take place when the parameterization is modified.

Parameterization

Resolution

Figure 5.33: Slices through the model for the first update (top) and the final update (bottom). The left hand side represents the slices through the velocity model. The right hand side represents the slices through the corresponding resolution.
The threshold value for the 'pseudo' singular-values is 100. So, when the largest value is smaller than 100, the addition of grid-points is stopped. However, this value is never reached, because the maximum number of LSQR iterations is limited to 1000. The relative resolution values calculated by the limited number of iterations will rarely exceed the resolution threshold in the final updates (15 to 20). As a result during these updates, the parameterization is only slightly modified. In this way, the limited number of iterations also limits the result. In order to get a more detailed velocity model, the number of LSQR iterations will have to be increased, which simultaneously will result in an increase of computer memory and computation time.

5.5.6 Results

The velocity model after 20 updates is compared to the true velocity model in Figures 5.34, 5.36, and 5.37. The homogeneous initial velocity model has changed into a model that represents the salt-dome and the sand-bodies. The shape of the salt-dome is obtained and also the fault in the salt-dome is visible. Note that the plots of the salt-domes in Figure 5.34 are obtained by making all velocities below a certain threshold value transparent. A true salt boundary is not estimated by the tomographic inversion process, as the estimated velocity model is always smooth. When the focal points are inspected separately (Figure 5.35) it turns out that the initially displaced focal points tend to align and show up as layers and faults. Also the flat reflector at 3.5 km becomes visible, however, it is not completely flat due to the lack of resolution in the deeper region. Figure 5.38c presents the rms values of the focal point location error. They decrease continuously during the updates for both the $x$, $y$ and the $z$ direction. The final rms location error is 40 m for the $x$ and $y$ direction and 60 m in depth. The location errors increase with depth, which is due to the decrease of resolution with depth (Figure 5.38d). The traveltine misfit decreases during the updates, and the final misfit is 4.7 ms. This is due to both the parameterization of the model, that is not able to describe the layers exactly, and the limited number of LSQR iterations. In order to get a more detailed velocity model, the number of iterations will have to be increased.
Evaluation of the methodology with synthetic data examples

**Figure 5.34:** Top: slices through the true velocity model (left), and the updated velocity model (right). Bottom: salt-dome of the true velocity model (left), and the updated velocity model (right).

**Figure 5.35:** The true focal point locations (left), and the updated focal point locations (right).
5.5 The 3D SEG/EAGE Salt model.

Figure 5.36: Top: vertical slices through the true ‘Salt’ velocity model, and the focal points (white dots) in the neighborhood of these slices. Bottom: vertical slices through the velocity model obtained after 20 updates, and the neighboring focal points.

Figure 5.37: Top: horizontal slices through the true velocity model. Bottom: horizontal slices through the velocity model obtained after 20 updates.
Figure 5.38: Important quantities during the updates. a) The number of velocity grid-points. b) The log-plot of the rms of the data misfit. c) The rms of the focal point location error for the x (green dashed line), the y (blue dashed-dotted line), and the z coordinate (red solid line). d) The rms of the focal point location error after the final update, with respect to depth.

■ 5.5.7 Focal point clouds

The focal point clouds evaluate the adequacy of the velocity model to explain the focusing operators (section 4.4). When the clouds of the initial model are analyzed (Figure 5.39a), it can be observed that the model does not explain the focusing operators. The shape of the focal point clouds above the salt-dome indicate that the velocity is too high and the focal points are located too deep (compare to Figure 4.9). The focal point clouds beneath the salt-dome show large errors, that cannot directly be translated in velocity or depth errors. The clouds are very complex due to the complicated ray-paths through the salt-dome. The focal point clouds shrink after each model update. The clouds of the final model (Figure 5.39b) show that this velocity model is capable of explaining the focusing operators. The clouds are concentrated at the focal point locations.
5.5 The 3D SEG/EAGE Salt model.

Figure 5.39: Focal point clouds for the Salt model. The red circles indicate the focal point locations. The blue dots form the clouds. a) Some focal point clouds for the initial model. b) Some focal point clouds for the final model.

### 5.5.8 2D versus 3D tomographic inversion

The importance of 3D seismic data and 3D tomographic inversion becomes clear, when the results of a 3D inversion are compared with the results obtained from a 2D inversion of the same (3D) subsurface. The input of the 3D tomographic inversion are 3D focusing operators resulting from the 3D model. The input of the 2D tomographic inversion are 2D focusing operators obtained from seismic data with a 2D geometry acquired in the 3D model. In Figure 5.40, a slice at 4.92 km in the x direction, and a slice at 4.92 km in the y direction are presented for both a 2D and a 3D inversion. In both cases, the salt-dome is resolved. However, the shape of the dome in 2D is not as accurate as in 3D. In particular the velocities of the sand-bodies and the location of the focal points in the deeper regions are better solved by the 3D tomographic inversion. In the 2D case the data misfit does not become smaller than 14 ms in the y slice and 8 ms in the x slice, while the data misfit in 3D is only 4.7 ms. In the 2D inversion, the rms of the focal point location errors reaches a minimum of 85 m in the lateral direction and 120 m in depth, versus 40 m and 60 m in the 3D inversion. The focal point clouds, constructed using the models obtained by 2D inversion, show that 2D velocity models are not capable to explain 3D effects in the data (Figure 5.41). These effects are visible in the focal clouds for both the x-slice and the y-slice. In particular the focusing around the salt-dome is difficult, due to the 3D structure of the dome.

Altogether, this indicates that 3D structures in model cause 3D effects in the data. As a result, 2D data, 2D focusing operators, and 2D tomography are not capable of obtaining an accurate velocity model, as the 3D effects cannot be explained by the 2D model.
Evaluation of the methodology with synthetic data examples

**Figure 5.40:** Comparison between the results of a 2D inversion and a 3D inversion for both slices in the x direction (left hand side) and the y direction (right hand side).

**Figure 5.41:** Focal point clouds for the Salt model constructed by using the models obtained by 2D inversion for both slices in the x direction (left hand side) and the y direction (right hand side).
5.6 Conclusions

The algorithm of tomographic inversion of focusing operators, as described in Chapter 3, has been demonstrated for both 2D and 3D synthetic examples. The additional modules, as described in Chapter 4, are also demonstrated in this chapter. Several conclusions can be drawn with respect to the three main objectives: 1) Developing a method for tomographic inversion of focusing operators, 2) Making the method data-driven, and 3) Exploring the possibilities of the method (section 1.3). The conclusions are subdivided in these three objectives.

5.6.1 Tomographic inversion of focusing operators

- The technique presented results in an accurate update of both velocities and focal point locations for both 2D and 3D synthetic cases. This accuracy is thanks to the regularization by resolution dependent parameterization. Parameters that are not well determined by the data are omitted. In addition, when gridpoints are added, this is only done at locations that are well determined, and only when the minimum traveltime misfit of the current parameterization is reached.

- The algorithm is capable of dealing with complex velocity models. All kind of structures can be solved accurately, as long as the resolution dependent parameterization is fine enough to describe it.

5.6.2 Data-driven procedure

- The initial model parameterization is coarse, and during the updates velocity grid-points will be added in the high resolution areas. In this way, the model will ‘grow’ towards its final solution in an accurate way.

- Whether the parameterization is modified, is also determined by the data. The parameterization is only modified when the data misfit of the current parameterization does not decrease any more.

- The resolution dependent parameterization will be influenced by the number of LSQR iterations, the 'pseudo' singular-values, and the resolution.

- As the method is data-driven, it requires minimal user-intervention. Only the threshold values for the 'pseudo' singular-values and the resolution, and the initial parameterization have to be defined.

- The data-driven result shows where information should be added. When the parameterization in an area is coarse, extra information is required to solve the area in more detail.
5.6.3 Exploring the possibilities of the method: The additional modules

- By using a priori information in the inversion the resolution increases in the region where the information is introduced. The extra information can yield a solution in areas that are not well resolved by the operators.

- The joint P and S-wave focusing operator inversion results in more accurate P and S-velocity models than separate inversions. In particular, the lower resolution S-velocity model can benefit from the information that the P-velocity model contributes.

- The focal point clouds show how well the velocity model explains the data.

5.6.4 Additional remarks

Besides the conclusions on the three objectives, some additional remarks can be made:

- The tomographic inversion generates a low frequency image of a subsurface that may contain high frequency features.

- Together, the updated velocities and focal point locations give an image of the subsurface in terms of velocities and horizons. This can be used during the 'post-updating' for interpretation and further processing.

- The shape of the 'pseudo' singular-value spectra show that the 3D problem is intrinsically better determined than the 2D problem.

- When significant 3D structures are present in the subsurface, the 3D seismic data and 3D tomographic inversion are required to obtain an accurate model of the subsurface. In these cases, the 2D data and 2D inversion are not capable of generating a correct model, because the 3D effects cannot be explained by the 2D model.
6

Evaluation of the methodology with real data examples

6.1 Introduction

In Chapters 3 and 4 the theory for the tomographic inversion of focusing operators has been discussed. In the previous chapter, this theory was applied successfully to synthetic data sets.

The next step is to apply the method to actually measured data. In this chapter, the tomographic inversion of focusing operators is applied to a 2D 'North Sea' data set, kindly provided by Norsk Hydro. As this marine data set contains both P and S wave recordings, both P and S wave focusing operators are generated. Hence, both the tomographic inversion of P operators and the joint inversion of P and S operators are evaluated. Also the concept of focal point clouds is applied to this dataset.
6.2 The 2D North Sea model

6.2.1 Description of the 2D field data set

In this section the performance of the tomographic inversion of focusing operators is tested on 2D field data from the North Sea. This OBS dataset contains recordings of both conventional PP data and converted PS data. In the first part of this section, only the PP data are considered. In the second part, the PS data are used as well. An initial stack of the PP dataset shows layering that is mainly horizontal, but it also shows some dips and faults (Figure 6.1). In this section, the reflection data from nine main reflectors are considered. In particular the region below the seventh reflector is interesting, as a hydrocarbon reservoir might be present in that area.

![Figure 6.1: Initial stack of the 'North Sea' dataset. The focusing operators that are related to the nine main reflectors (indicated by the numbers) are considered. Note that the vertical axis represents the two-way traveltimes.](image)

6.2.2 Input data: the 2D focusing operators

The focusing operators for the OBS 'North Sea' data set are obtained by means of the CFP method. Application of this method, requires sources and receivers to be at the same acquisition level. Therefore, Berkhout and Verschuur [2000] applied a redatuming of the sources towards the ocean bottom using water depth and velocity. Using this redatumed data set, the focusing operators can be obtained. In Verschuur et al. [2002] this procedure has been described. First, the PP data is considered. At the nine reflectors 1592 focal points are defined. At these points, the focusing
operator updating procedure as described in section 2.4 is used in order to obtain the P-wave focusing operators (Figure 6.2). The aperture of the shallow focusing operators is small due to the limited reflection angle, but for the deeper operators the maximum aperture is about 2 km. The receiver spacing at the surface is 50 m. The focusing operators form the input data for the inversion.

![Figure 6.2: P-wave focusing operators of the North Sea dataset. For visualization purpose only 10% of the 1592 operators is shown. Note that the vertical axis represents the one-way traveltimes.](image)

### 6.2.3 Initial velocity model

To start the inversion procedure, an initial velocity model, having a homogeneous velocity of $2.2 \text{km/s}$, is defined. This model is described by only 36 grid-points which are connected by the Delaunay triangulation (Figure 6.3a). Also the initial focal point locations have been defined. They are located at the lateral position of the apex of each operator, at a depth that is consistent with the minimum traveltimes. The initial focal point locations are indicated by the white dots in Figure 6.3b that reveal the nine reflectors to which they are related.

The initial focusing operators are modeled by calculating rays through the velocity model from the focal points towards the receivers at the surface. By minimizing the difference between the observed and the modeled focusing operators the model parameters are optimized.
6.2.4 Optimization of the model parameters

The optimization strategy is analogous to the method that was used for the synthetic models. The number of LSQR iterations in each update will be 60% of the number of velocity parameters that are optimized. As a result only 18 iterations are performed during the first update. The dark blue line in Figure 6.4a indicates the 'pseudo' singular-value spectrum for the first update. The high singular-values show that the parameters are well-determined. The large differences between the singular-values show that some parameters are much better determined than other parameters.

6.2.5 Modification of the parameterization

The initial velocity model is described by only a few grid-points that are mainly well-determined and the under-determined points will be removed in the first update (Figure 6.5b). During the next updates the parameterization will only be modified when the data misfit does not decrease anymore. If this is the case, grid-points will be added to the model at regions with a high resolution. For the inversion of the 'North Sea' model, the parameterization is modified after update 2, 8, 10 and 12 (Figure 6.5 and 6.4b). The parameterization of the final model is defined by 1320 grid-points.

The 'pseudo' singular-values that are calculated during the optimization, serve as the stop-criterion for the addition of grid-points. Figure 6.4a shows how the singular-value spectrum develops during the updates. Major changes in the spectrum take place when the parameterization is modified (update 2, 8, 10 and 12). When the singular-values of the final update (dark red line) are compared with the values of the first update (dark blue line), it is clear that the highest values decreased and
finally start to approach a singular-value-threshold value, which is set at 30. At this point, the modification of the parameterization by adding grid-points to the model is stopped. The final singular-values are more similar, which indicates that all parameters are more or less equally well-determined.

The data-misfit evaluates the result quantitatively. Figure 6.4c shows how the data misfit behaves during the updates. A slight increase of the misfit can be observed when the parameterization is modified during update 8. After 12 updates the rms difference between the traveltimes of the observed and modeled focusing operators is 1.6 ms.
Figure 6.5: The modification of the parameterization for the 'North Sea' model. On the left hand side the resolution of the model is plotted, and on the right hand side the resulting parameterization is plotted. From top to bottom, the pictures show the modification of the initial parameterization, the parameterization of update 8, the parameterization of update 10, and the parameterization of update 12.
6.2.6 Result of the P-wave focusing operator inversion

The final result after 12 updates is shown in Figure 6.6. Both the velocities and the focal point locations are updated. The updated focal points still represent the nine reflectors and reveal the faults and dips (Figure 6.6b). However, when they are compared to the initial focal points (Figure 6.6c), it becomes clear that they are shifted upward. In particular the deeper reflectors show this change. The updated model also shows that velocity changes are not necessarily restricted to the layers that are bounded by the reflectors. The velocity varies everywhere in between the reflectors, and in particular below the seventh reflector. Although this reflector shows lateral variations due to faults, these variations are not followed by the velocity, but are more constant in the lateral direction. The flat transition from lower to higher velocity occurs below the reflector. This change in velocity might result from a change in pore fluid, as the presence of hydrocarbons almost always results in a lowering of velocity.

To consider the imprint of the irregular parameterization on the updated model (Figure 6.6a), we redefine the model by a regular grid (Figure 6.6d) and perform a few extra updates. The resulting velocity model is presented in Figure 6.6e. When it is compared to the result obtained using the irregular parameterization (Figure 6.6b), it can be seen that all the visible structures in the model has not changed, but the small laterally varying velocities, caused by the variation in triangles, are smoothed. The data misfit of this model is 1.2 ms. Especially the velocity transition below the seventh reflector is smoother in the horizontal direction. As a result, the flat nature of the low velocity area becomes clearer.

The velocities resulting from tomographic inversion have been used to migrate the seismic data. The result is shown in Figure 6.7b, in which the faults and the layers are clearly visible. Also a continuous flat event appears below the seventh reflector, which might be a hydrocarbon indicator. From the depth migrated shot gathers, Common Image Gathers can be generated. A number of CIG’s along the line are displayed in Figure 6.8. A CIG reveals the accuracy of the estimated velocity model at a lateral position. Flat events indicate the correctness of the model. The gathers show generally flat events, and especially the events that correspond to the reflector on which the focusing operators are determined are quite flat. In order make the other events also flat, a more detailed velocity model has to be generated. This can be obtained by using more focusing operators, that are related to the events that have to be flattened. By using more operators, the resolution becomes higher, the parameterization finer, and the model more detailed. In all CIG’s, the event below the seventh reflector is not aligned, but curves downward. This indicates that a too high velocity is used for the migration of this event. In other words, this event has a lower velocity, which, again, might indicate the presence of hydrocarbons.
Evaluation of the methodology with real data examples

Figure 6.6: The result of tomographic inversion of P-wave operators after 12 updates. 

a) The parameterization of the updated model. 

b) The updated velocity model containing the updated focal points. 

c) Comparison between the initial focal point locations (open dots) and the updated focal point locations (closed dots). 

d) Regular parameterization of the updated model as presented in a). 

e) The result of updating by using the regular parameterization.
Figure 6.7: a) The updated velocity model containing the updated focal points. b) Result of a shot record migration of the PP data using the updated velocity model in a). The nine main reflectors are indicated by the numbers. The flat event below the seventh reflector is indicated by a dotted line.
Evaluation of the methodology with real data examples

Figure 6.8: A number of Common Image Gathers along the line, generated from the depth migrated shot gathers. The gathers show the accuracy of the updated velocity model that is used for the migration (Figure 6.7). The events that are related to the nine main reflectors are indicated by the numbers.
6.2.7 Joint inversion of P and S-wave focusing operators

Now that the P-wave operators have been inverted successfully, the S-wave operators are considered as well. They are obtained from the converted wave data by making use of the P-wave operators (section 4.3.3). The P-wave operators are used to focus the sources of the PS data. Next, the S-wave operators can be 'read' from the obtained CFP gather [Berkhout and Verschuur, 2000; Verschuur et al., 2002]. The S-wave operators are related to the same 1592 focal points as the P-wave operators (Figure 6.9). However, it is clear that the aperture of the S-wave operators are much smaller, due to the lower S velocity (section 4.3.4).

![Figure 6.9: S-wave focusing operators of the North Sea dataset. For visualization purpose only 10% of the 1592 operators is shown.](image)

In a first experiment, only the S-wave operators were tomographically inverted. The initial S-velocity model has a homogeneous velocity of 0.7 km/s, and the initial focal points are at the same locations as the initial focal points of the P-wave operator inversion (Figure 6.2). After a few updates, the data misfit reduced, but the minimum misfit will only reduce to 40 ms and the focal points are positioned at suspicious locations. Therefore, a second experiment is performed, in which a joint inversion is applied to the P and S-wave operators. In this case, a better result is expected (section 4.3). However, during the joint inversion, the misfit also remains high and the poor resolving power of the S-wave operators is transferred to the P-velocity model via the inaccurate focal point updates.
Evaluation of the methodology with real data examples

Figure 6.10: The result of tomographic inversion of S-wave operators after 31 updates. a) The parameterization of the updated model. b) The updated velocity model containing the focal points.

Therefore, it was decided to perform a joint inversion where the influence of the well-behaving P-wave operators can be transferred to the S-wave operators, but not the other way round. In other words, the focal point locations in both the P and S-velocity model are forced to be determined by the P-wave operators only. In this way, the focal points are expected to be positioned at plausible locations in the S-velocity model. During the updates the parameterization of the S-velocity model is modified, but it stays rather coarse due to the low information content in the S-wave operators (Figure 6.10a, 6.11a). Also here, the data misfit stays large and does not become smaller than 60ms (Figure 6.11b).

Figure 6.11: Important quantities during the updating of the S-wave operator inversion. a) The number of parameters defining the velocity of the model. b) The log-plot of the rms of the data misfit.
6.2 The 2D North Sea model

Figure 6.12: a) The focal point clouds of the P-wave operators, using the updated velocity model as shown in Figure 6.6. b) The focal point clouds of the S-wave operators, using the updated velocity model as shown in Figure 6.10.

In order to understand why the S-wave operator cannot be explained by the tomographic inversion algorithm, the focal point clouds are considered (see section 4.4). The focal point clouds can evaluate the adequacy of the velocity model to explain the focusing operators. The focal point clouds of the updated S-velocity model (Figure 6.12b) reveal that this model is incapable of explaining the focusing operators. In addition, the focal point clouds diagnose the origin of the residual between the observed focusing operators and the operators calculated in the velocity model. The
clouds of the updated S-velocity model reveal the presence of anisotropy in which the velocity in the horizontal direction is higher than the velocity in vertical direction (see section 4.4.2). Note that only the S-wave operators are severely affected by anisotropy, while the P-wave operators are well described by the updated isotropic P-velocity model. This is shown by focal point clouds of the updated P-velocity model, which form points at the focal point locations (Figure 6.12a). In order to explain the S-wave operators, anisotropy should be included in the S-velocity model.
6.3 Conclusions

The algorithm of tomographic inversion of focusing operators, as described in Chapter 3, has been applied to a real 2D data set. Joint inversion of P and S operators and focal point cloud analysis were also applied to this data set. Several conclusions can be drawn with respect to the three main objectives of this research (section 1.3).

### 6.3.1 Tomographic inversion of focusing operators

- For real data, where the true velocity model is not known, the only criteria for the performance of the method are the data misfit and the focal point clouds. The traveltime misfit between the observed and modeled P operators is minimized within acceptable limits. The P velocity model is capable of explaining the P wave focusing operators, which is indicated by the clouds that form points at the focal point locations.

- Another quality check is the migration that is performed using the estimated velocity model. This migration shows clear dips and sharp faults, and reveals a flat event that may be a direct hydrocarbon indicator. The Common Image Gathers, generated from the migrated shot gathers, show flat events. In particular for the events that correspond to the reflector on which the focusing operators are determined.

- The algorithm for tomographic inversion of focusing operators is also accurate for this real data example.

- This real data example, which contains clear reflectors, shows that the velocity changes are not necessarily restricted to these reflectors.

### 6.3.2 Data-driven procedure

- The initial model parameterization is coarse, and during the updates velocity grid-points will be added in the high resolution areas. In this way, the model will 'grow' towards its final solution in a stable way.

- Also in the real data example the method is data-driven with very little user interaction. Only the threshold values for the 'pseudo' singular-values and the resolution have to be defined, which were chosen to be the same as in the synthetic example.

- The data-driven result shows where information should be added. The resolution is very high in the middle of the model, while at the bottom and at the edges resolution is lower.
• In the areas where the parameterization is fine, the result is not very smooth. By re-parameterization of these areas, using a regular grid, the velocity changes become more smooth. This does not necessarily mean that the result obtained with the regular grid is better.

§ 6.3.3 Exploring the possibilities of the method: The additional modules

• In the joint inversion of P and S operator, the S operators have an adverse effect on the P operator inversion. This influence is transferred through the shared focal points.

• An S velocity model can be obtained by transferring only the information from the P to the S model, and not the other way round. Yet, the data misfit for the S operators stays very large.

• The focal point clouds of the S operator inversion result reveal that the subsurface, and hence the corresponding focusing operators, contain anisotropy. Therefore, it is impossible to invert the S operators using an isotropic velocity model.
Conclusions and Recommendations

In this thesis a method for tomographic inversion of focusing operators is presented. The data input consists of verified focusing operators, which are estimated from seismic data by using ‘the principle of equal traveltime’. These operators contain the kinematic effects of wave propagation in the subsurface, and can be viewed as one-way travel times from a virtual source point in the subsurface, the focal point, to receivers at the surface. The advantage of one-way focusing operators is that they are less complicated than conventional two-way data. Therefore, these focusing operators are very suitable for inversion to a velocity model.

Tomography is a known concept, but when applied to focusing operators it has some unique features. One of these is the parameterization of the velocity model. In two-way tomography propagation velocities and reflecting boundaries are estimated. In one-way tomography, however, the propagation velocities and the focal point locations have to be estimated. In this respect, the tomographic inversion of focusing operators can be compared to earthquake tomography, in which the velocity structure of the earth and the location of the earthquake are estimated by means of the one-way travel times generated by the earthquakes.

The objectives of this research were: 1) Develop a method for 3D tomographic inversion of focusing operators, 2) Make the method fully data-driven, and 3) Explore the advantages of the tomographic inversion of focusing operators, like the inclusion of a priori information, the joint inversion of P and S-wave operators, and the evaluation of the inversion result by focal point clouds. In this chapter, conclusions that can be drawn from this research are summarized and recommendations for future research are given. This is done separately for each of the three objectives.
7.1 Tomographic inversion of focusing operators

\subsection*{7.1.1 Conclusions}

The method developed contains four main steps: parameterization, forward modeling, optimization, and regularization. Detailed conclusions on this theory are given in section 3.7. The method has been evaluated on synthetic as well as field data, in both 2D and 3D cases. Specific conclusions on this evaluation are presented in section 5.6 and 6.3. This section summarizes the main advantages and limitations of the method:

Advantages

- The subsurface is described by model parameters that directly influence the travel times: the velocities and the spatial locations of the focal points. The velocities are defined on irregular grid points, which are connected by Delaunay triangulation in which the quadratic slowness is linearly interpolated. Defining the velocities at the grid points, and not inside the triangles, reduces the number of parameters and results in a smooth model. The use of quadratic slowness facilitates the involved forward modeling algorithm. The focal points are defined independently of the velocity grid points. They can 'float around' in the velocity model. As a consequence, the velocity changes are not restrained to the reflector. By means of this irregular parameterization and the independence of the focal points, any model of the subsurface can be described. In addition, this flexibility in the velocity parameterization is required to make the data-driven method possible.

- The forward modeling of focusing operators is performed by a ray-based method. The focal points are considered to be sources from which rays are calculated towards receivers at the surface. The ray-tracing is done analytically and is, consequently, efficient in memory. The ray-paths can be used to derive the partial derivatives of the traveltime with respect to the quadratic slowness and the focal point location. Hence, the partial derivatives are also calculated analytically.

- The optimization of the model parameter updates is carried out by the LSQR method. The LSQR method determines the generalized inverse in an iterative way and is, therefore, efficient in memory. The LSQR algorithm can also calculate 'pseudo' singular-values in a practical way, by which the inversion system can be analyzed.

- The regularization is performed by using resolution dependent parameterization: the resolution of the parameters in the optimization is translated to the
spatial resolution in the parameterization. Parameters that are not well determined by the data are removed and in high resolution regions parameters can be added. In this way, an attempt is made to obtain a well-determined inverse problem, in which the minimum is well-shaped. Besides, a user-introduced bias is avoided.

- The proposed algorithm has been evaluated on synthetic data in both 2D and 3D, and on real data in 2D. The algorithm is capable of dealing with complex velocity models. All kinds of structures can be resolved accurately, as long as the final parameterization is fine enough to describe it.

**Limitations**

- For the forward modeling a ‘shooting’ method is used. Hence, to reach the predefined receivers, interpolation by re-shooting is required. Although this interpolation is done efficiently, the re-shooting makes the method slower. Moreover, it is not always possible to reach all the receivers, especially when the model is complex. This can, for example, cause problems when the model exhibits higher contrasts during the updates. In this case more rays are ‘lost’ during the updates, which might result in under-determined problem. Nevertheless, this is rather a limitation of the implementation used in this thesis, than a fundamental problem.

- The maximum number of LSQR iterations is limited in order to prevent the inversion from becoming too expensive. However, this limited number of iterations may confine the inversion result. Also this is not a fundamental problem, but occurs because of the restricted computation time and computer memory.

### 7.1.2 Recommendations

In view of the foregoing, the following recommendations are made:

- As mentioned above, the shooting method requires re-shooting to hit the receivers, and does not necessarily reach each receiver. This practical problem can be circumvented on the short term by using a 'bending' method for ray-tracing [Moser et al., 1992]. This method retains the advantages of the ‘shooting’ method, because it is also an analytical ray-tracer in which the partial derivatives can be generated automatically. This 'bending' method will always find a ray that connects the source to a receiver. On the long term, one should focus on more advanced forward modeling methods that make use of recursive wave theory [Berkhout, 1982; Yilmaz and Taner, 1997].

- It is preferable that the number of LSQR iterations is not limited by computation time or memory, but dependent on the inverse problem as advised
by Paige and Saunders [1982]. So, in stead of using the first strategy for the
tomographic inversion, as proposed on page 70, the second strategy should be
used. Of course, one should keep in mind that this will require more computer
memory and computation time.

An extra recommendation that can be made is not related to the practical limitations
of the method, but is concerned with a possible extension of the method:

- The inversion of amplitudes might be a valuable addition to the inversion,
because they can reveal more information on the elastic properties of the sub-
surface. Also the use of multi-valued arrivals can contribute to the existing
inversion package, as in very complex structures the major energy may not be
contained in the first arrival times. Note that the method, as proposed in this
thesis, is already suitable for the inversion of multi-valued arrival times.

7.2 Data-driven procedure

7.2.1 Conclusions

An important aspect of the presented method is that it is fully data-driven. This
aspect is realized by the regularization of the inverse problem, which is accomplished
by a resolution dependent parameterization. As the parameterization depends on
the resolution, and the resolution depends on the data, the method becomes fully
data-driven. No user intervention is needed to guide the model parameterization.
The theory has been presented in section 3.6, and the method has been applied
in Chapters 5 and 6. Detailed conclusions on the method and its application are
given in section 3.7, and 5.6 and 6.3 respectively. A summary of the advantages and
limitations is presented in this section:

Advantages

- In the resolution dependent parameterization, low resolution parameters are
removed and parameters are added in high resolution regions. This way of
parameterization is data-driven because the parameterization is determined
by the resolution, while the resolution is determined by the data. As a con-
sequence, this regularization not only attempts to generate accurate and sta-
ble results without user-introduced bias, but also to makes the method data
-driven. For the resolution dependent parameterization, the relative resolution
of the parameters is used as a qualitative criterion. This resolution can effi-
ciently be calculated by means of the LSQR method.

- In the tomographic inversion of focusing operators the initial model parame-
terization is coarse, and during the updates velocity grid-points will be added
in the high resolution areas. There are three reasons for this strategy: 1) the first updates in which the main trends of the velocity model are calculated are fast, 2) it is more stable to add points than to remove points, as removing points will change the model considerably and will temporarily result in a deterioration of the model, and 3) the chance of ending the inversion in a local minimum is significantly decreased by not using a fine parameterization from the beginning. By using this strategy the model will 'grow' towards its final solution in a stable way, driven by the data. Whether the parameterization is modified is also determined by the data. The parameterization is only modified when the data misfit resulting from the current parameterization does not decrease anymore. This procedure leads to robust results.

- As the method is data-driven, there is minimal user interference; all the information that is available in the data is translated into the model, without intervention of the user. Eventually, the result can be analyzed by the user. The data-driven result shows both the estimated velocities and the focal point locations. Together they give a complete view of the subsurface structure. This can be used during the 'post-updating' for interpretation and further processing.

- The data-driven result shows where information should be added. When the parameterization in an area is coarse, extra information is required to resolve that area in more structural detail.

Limitations

- Although the tomographic inversion of focusing operators is designed to be, as much as possible, data-driven, even in this method interaction of the user has to take place at some stage. In this method, the interaction appears in defining the initial model and the threshold values for the 'pseudo' singular-values and the resolution, which is a very indirect way of user-interaction. The sensitivity of the final result to the threshold values have been described on page 108. Nevertheless, the sensitivity to the initial model is hard to quantify. Even though different initial velocity models that were used led to similar accurate results, no definite conclusions on the dependence of the final result on the initial model can be drawn.

- In this research a local optimization method is used and, as a consequence, there is a risk that the final solution ends in a possible local minimum. The risk has been reduced by using a parameterization that 'grows' towards the final solution, however, the risk has not been eliminated by using this strategy. Even though different initial velocity models led to similar results, the uniqueness of the result can not be claimed.
7.2.2 Recommendation

- The sensitivity of the final result to the user-defined initial model needs to be investigated by examining whether the solution space contains local minima. The presence of local minima can be checked by making use of the global inversion techniques [Rothman, 1985; Mosegaard and Vestergaard, 1991; Sambridge and Drijkoningen, 1992].

7.3 Exploring the advantages of the method: The additional modules

7.3.1 Conclusions

The final objective of this research was to explore further possibilities of the method. This research has focused on three additional modules: 1) Inclusion of a priori information, such as water velocity and well-log velocities, 2) Joint inversion of P and S-wave operators, and 3) Evaluate the inversion result by using focal point clouds. The theory of these modules is presented in Chapter 4, and they have been applied in Chapters 5 and 6. This results in the following conclusions on the advantages and limitations of these modules:

Advantages

- It has been shown that a priori information about the model parameter values can be included in the tomographic inversion methodology by using Kriging and taking this data into account as extra constraints in the inversion. By using a priori information in the inversion the resolution increases and the extra information can bring a solution in the areas that are not well resolved by the operators. This information may improve the quality of the estimated model in terms of depth conversion and geological information.

- The inversion of P and S-wave operators is presented. As the estimation of P and S-wave operators in mode converted data always results in operator pairs related to the same focal points, this aspect can be translated into a joint inversion. In this inversion, a P and an S-velocity model are estimated simultaneously, in which the focal point coordinates are the same in both models.

- The joint P and S-wave operator inversion can result in more accurate velocity models than a separate inversion. In practice, the S-wave operators results a lower resolution than the P-wave operators. Therefore, the lower resolution S-velocity model can benefit from the information that the P-velocity model contributes.
7.3 Exploring the advantages of the method: The additional modules

- The concept of focal point clouds is introduced. This is a tool that can evaluate the adequacy of the used velocity model to explain the focusing operators. The shape of the focal point clouds makes it possible to see what causes the defocusing. Velocity errors can be distinguished from focal point location errors, and anisotropy and 3D effects can be recognized.

Limitations

- In the joint inversion of P and S-wave operators also adverse effects are possible. When it is impossible to fit one of the operator types it will immediately affect the other operator type via the focal points. Therefore, attention should be paid to both operator types separately, before inverting them jointly.

- Although the focal point clouds give information on the ability of the velocity model to explain the operators, in the current implementation this information has not been translated to the velocities yet.

7.3.2 Recommendations

The third objective of this research was to explore the further possibilities of the algorithm developed. The following recommendations can be made:

- For the joint inversion of P and S-wave operators it is recommended to analyze and invert both operator types separately, before inverting them jointly, in order to detect the negative influence from one of the wave types. The focal point clouds can be a useful tool for this detection. In addition, this negative influence should be removed. This can be achieved by performing a joint inversion where the influence of the well-behaving operators can be transferred to the operators that have a negative influence, but not the other way round. In other words, the focal point locations in both the P and S-velocity model are forced to be determined by the well-behaving operators only.

- Further investigation of the focal point clouds is required, to see what information can be extracted from these clouds. Besides, a lot of progress could be made by applying pattern recognition to the clouds in order to extract the different kind of information automatically. Another point of interest is the translation of the cloud information to the velocity model. It is, for example, possible to find anisotropy with the help of the focal point clouds. However, this anisotropy can not be resolved, as the tomographic inversion scheme does not include anisotropy yet. The anisotropy has to be included in several aspects of the scheme; Anisotropy parameters have to be included in the model parameterization and the forward modeling algorithm should honor anisotropy. Until now, the focal point clouds have only been used to see how
well the model can explain the focusing operators, which were assumed to be correct. In addition, it should also be investigated whether the clouds can be used to give a warning about the validity of the operators.
Parameterization by Delaunay triangulation

A.1 Delaunay triangles and tetrahedra

The principle of Delaunay triangulation is simple but powerful. Given a set of data vertices, the Delaunay triangulation is a set of lines connecting each vertex to its natural neighbors (Figure A.1). In 2D the Delaunay triangulation is defined such that every circumcircle of a triangle is an empty circle, i.e. it does not include any other vertices. This is related to the Voronoi tessellation, as the circumcircle about a Delaunay triangle has its center at the vertex of a Voronoi polygon. In 3D, the triangulation procedure is the same, but circumcircles become circumspheres, Voronoi polygons become Voronoi polyhedra, and Delaunay triangles become Delaunay tetrahedra [Delaunay, 1924; Virieux, 1998].

A.2 Delaunay triangulation

■ A.2.1 Different methods

Several algorithms for generation of Delaunay triangulation exist [Fleischmann, 1999]. Generally the algorithms can be divided into three groups:

[1] Tetrahedra are calculated by means of convex hulls [Barber et al., 1995];

[2] Tetrahedra are attached to the current boundary of the triangulation [Lohner and Parikh, 1988];

Figure A.1: Geometry of Delaunay triangulation. a) In 2D each Delaunay triangle has a circumcircle that does not include any other vertices. The center of the circumcircle defines one vertex of the Voronoi polygon. b) In 3D a Delaunay tetrahedron is defined by means of a circumsphere.

The algorithm of the first kind is commonly used, because it is quick. This method uses a randomized algorithm based on the convex hull to generate the Delaunay triangulation. However, this algorithm can generate facets of tetrahedra, that do not correspond to the facets of the neighbor tetrahedra (Figure A.2). For some purposes, like ray-tracing within tetrahedra, it is required that the facets of the tetrahedra connect properly. So, the convex hull algorithm is not suitable for the tomography method presented in this thesis.

In the second algorithm, one starts with one tetrahedron and new tetrahedra are attached to the outside until a complete triangulation is obtained in which all vertices are connected [Lohner and Parikh, 1988]. However, when a new vertex is added inside the existing triangulation, the model has to be rebuilt from the beginning. This makes the method unsuitable for building the dynamic velocity models used in the tomography algorithm.

In this research an algorithm of the third group is used, because it produces consistent facets between tetrahedra and can add vertices to an existing triangulation. This method will be explained in the next part.
A.2 Delaunay triangulation

![Figure A.2](image_url)

**Figure A.2:** Problem using the convex hull algorithm: facets of left-hand side tetrahedron \( \mathbf{x}[1], \mathbf{x}[2], \mathbf{x}[3], \mathbf{x}[4] \) and \( \mathbf{x}[5], \mathbf{x}[6], \mathbf{x}[7], \mathbf{x}[8] \) do not necessarily correspond to the facets of the neighbor right-hand side tetrahedron \( \mathbf{x}[1], \mathbf{x}[2], \mathbf{x}[3], \mathbf{x}[4] \) and \( \mathbf{x}[5], \mathbf{x}[6], \mathbf{x}[7], \mathbf{x}[8] \).

### A.2.2 Delaunay triangulation used in the tomographic inversion

The method that is used in this research is based on the algorithm of Watson [1981]. The Delaunay triangulation is done in two phases: placement of the vertices, followed by triangulation. First, an initial triangulation is generated (see section A.3). Within this background triangulation new vertices are inserted. The tetrahedra whose circumspheres contain the inserted vertex are removed, and the resulting cavity is triangulated by linking the inserted vertex to all vertices of the cavity boundary. Figure A.3 shows this in 2D for visualization purposes. The code of this algorithm is as follows:

<table>
<thead>
<tr>
<th>Insert new vertex</th>
</tr>
</thead>
<tbody>
<tr>
<td>For circumspheres of all tetrahedra</td>
</tr>
<tr>
<td>if vertex in circumsphere</td>
</tr>
<tr>
<td>save 4 facets of tetrahedron</td>
</tr>
<tr>
<td>For all saved facets</td>
</tr>
<tr>
<td>if facet occurs once:</td>
</tr>
<tr>
<td>make tetrahedron from facet and new vertex</td>
</tr>
<tr>
<td>if facet occurs twice:</td>
</tr>
<tr>
<td>delete facet</td>
</tr>
</tbody>
</table>
Figure A.3: a) The new vertex (white dot) is positioned in the background triangulation (defined by black dots and solid triangles). b) It is determined whether the new vertex is within one of the existing circumcircles. c) The vertex is enclosed by the circumcircle of triangles a, b, and c. The edges of these triangles are considered. If an edge occurs twice (dashed line) it is an internal edge and is deleted. If an edge occurs once (dotted line) it is an outer edge. d) From the vertices of the outer edge to the new vertex the new edges are generated, which results in the final Delaunay grid.

A.2.3 Problems occurring in 3D Delaunay triangulation

The presented algorithm is quite simple and works perfectly in 2D. However, Delaunay triangulation in 3D shows some new difficulties. First, the very same domain may be triangulated with different numbers of tetrahedra. For example, a cube can be triangulated with either five or six tetrahedra (see section A.3). Second, tetrahedra can be poorly shaped in more ways than triangles. Needles, wedges and spindles can occur [Sack and Urrutia, 1999, chapter 6]. The number and the shape of the tetrahedra are not really a problem in the method presented in this research.

Nevertheless, there are tetrahedra that have a very specific bad shape, which do cause problems. These are the flat tetrahedra, also called slivers. A sliver has a very small or even no volume. As a result, they form a vacuum within the triangulation, which makes forward modeling impossible. Slivers are generally formed when vertices are added in the plane of an existing facet of the background triangulation (Figure A.4a and b). In addition, this facet should be an outer facet that will be connected to the inserted vertex. The resulting sliver can be removed by adding extra vertices.
A.3 Initial Delaunay triangulation

[Li, 2000]. However, this is not desirable, as in this research vertices can only be added based on a certain criterion. Therefore, another method is developed. It this method it is assumed that the neighbor tetrahedron on the other side of the outer facet is also affected by the inserted vertex (although it was originally not affected, see Figure A.4c). This transforms the outer facet that formed the sliver, into an inner facet that will be deleted.

![Diagram](image)

**Figure A.4**: a) The new vertex (white dot) is situated within the circumsphere of the tetrahedron $A, B, C, D$ (defined by black dots) b) Facet $A, B, C$ is the outer facet, which is connected to the new vertex to form a new tetrahedron. This results in a tetrahedron without a volume: a sliver. c) To remove this sliver, the neighbor tetrahedron on the other side of the outer facet (defined by the outer facet and the gray dot) is assumed to affect the new vertex too. This makes the outer facet an inner facet. In d)- f), the new outer facets of the neighbor tetrahedron are connected to the new vertex, resulting in well-shaped tetrahedra.

---

A.3 Initial Delaunay triangulation

For the Delaunay triangulation method that is presented here, an initial triangulation that encloses all vertices is required. In this method we use a rectangular box, because the subsurface of the earth is generally represented this way. The maximum and minimum coordinates of the available vertices are considered and a bounding box based on these coordinates is constructed. As a tetrahedral model is needed, this box is divided in tetrahedra, which can be done in two different ways; division into five or six tetrahedra (Figure A.5). The advantage of five tetrahedra is that
less tetrahedra are needed. The advantage of six tetrahedra is that all tetrahedra have equal volume. The initial triangulation and the division of cubes have several applications. It can be used for conversion from a regular grid to a tetrahedral grid. In this case the regular grid is connected by cubes, and these cubes are divided into tetrahedra. It can also be used for the construction of a tetrahedral layer model in which the layers are connected vertically. The resulting rectangular boxes are again divided into tetrahedra.

Figure A.5: a) Division of a cube in five tetrahedra. b) Division of the same cube in six tetrahedra, with equal volume.
A.4 Refining Delaunay triangulation

In this thesis, vertices will be added to an existing triangulation when the following criterion is met: the four vertices of the tetrahedron in which the vertex is positioned should have a high resolution (see section 3.6.2). The new vertex will be added within the tetrahedron defined by four vertices $x_i^{[k]}$, with $k = 1, 2, 3, 4$ and $i = 1, 2, 3$ in 3D. The new vertex $x_{i}^{\text{new}}$ will be added at the centroid of the tetrahedron (Figure A.6), which can be calculated by:

$$x_{i}^{\text{new}} = \frac{\sum_{k=1}^{4} x_i^{[k]}}{4}. \tag{A.4.1}$$

The triangulation will be adjusted by means of the triangulation as presented in subsection A.2.2.

Figure A.6: The centroid of a triangle corresponds to the center of mass (white dot), and can be calculated by the intersection of the three triangle medians $(x^{[1]}, m^{[1]})(x^{[2]}, m^{[2]})(x^{[3]}, m^{[3]})$
Forward modeling by the ray-based method

B.1 Eikonal equation

The eikonal equation is derived from the wave equation [Berkhout, 1982]. The homogeneous wave equation for pressure in a constant density medium is formulated as follows:

\[
\frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} + \frac{\partial^2 \phi}{\partial x_3^2} = q(x) \frac{\partial^2 \phi}{\partial t^2},
\]  

(B.1.1)

where \( \phi \) is the pressure in the coordinate system defined by \( x = (x_1, x_2, x_3) \), \( t \) is the traveltime, and \( q(x) \) is the quadratic slowness \((s(x))^2\) at location \( x \). The solution of the wave equation can be obtained by means of the equation for a plane wave:

\[
\phi(x, t) = A(x) e^{i(k \cdot x - \omega t)},
\]  

(B.1.2)

where \( A(x) \) is the wave amplitude, \( \omega \) is the frequency, and \( k \) is the wavenumber vector that points in the direction of propagation and thus, by definition, is a ray. \( k \cdot x \) can be replaced by \( \omega \tau(x) \), where \( \tau(x) = s(x) \cdot x \) represents the traveltime field. This traveltime field is called the eikonal function. Making use of \( \tau(x) \) and substituting (B.1.2) into equation (B.1.1), this results in

\[
\nabla^2 \left[ A(x) e^{i\omega[\tau(x)-t]} \right] = q(x) \frac{\partial^2}{\partial t^2} \left[ A(x) e^{i\omega[\tau(x)-t]} \right],
\]  

(B.1.3)
which gives two sets of equations. One describing the real part of the solution:

$$\nabla^2 A(x) - A(x)\omega^2 \left[ \left( \frac{\partial \tau(x)}{\partial x_1} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_2} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_3} \right)^2 \right] q(x) = 0, \quad (B.1.4)$$

and the other describing the imaginary part:

$$2 \left( \frac{\partial \tau(x)}{\partial x_1} \frac{\partial A(x)}{\partial x_1} + \frac{\partial \tau(x)}{\partial x_2} \frac{\partial A(x)}{\partial x_2} + \frac{\partial \tau(x)}{\partial x_3} \frac{\partial A(x)}{\partial x_3} \right) + A(x)\nabla^2 \tau(x) = 0. \quad (B.1.5)$$

The terms of the real part can be rearranged as

$$\left( \frac{\partial \tau(x)}{\partial x_1} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_2} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_3} \right)^2 q(x) = \frac{1}{A(x)\omega^2} \nabla^2 A(x). \quad (B.1.6)$$

The right hand side of this equation is a ratio of the spatial Laplacian of the amplitude divided by $\omega^2$. For high frequencies this term is small (let it be approximately zero), and thus (B.1.6) reduces to:

$$\left( \frac{\partial \tau(x)}{\partial x_1} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_2} \right)^2 + \left( \frac{\partial \tau(x)}{\partial x_3} \right)^2 = q(x), \quad (B.1.7)$$

which can be briefly written as

$$\nabla \tau(x) \cdot \nabla \tau(x) = q(x). \quad (B.1.8)$$

This is called the eikonal equation. It shows that if this equation applies, then the traveltine function $\tau(x)$ is independent of frequency, i.e. no dispersion will occur. So, solutions to the eikonal equation are not exact solutions to the wave equation due to the high frequency assumption. However, the eikonal equation will approximate the wave equation well if the fractional change in velocity gradient over one seismic wavelength is small compared to the velocity [Lay and Wallace, 1995]. For many regions inside the real earth, the necessary restrictions on spatial variations are satisfied, so solutions of the eikonal equations are useful. Note that the eikonal equation provides traveltine information only. To obtain amplitude information, equation (B.1.5) should be included as well.

The physical meaning of the eikonal equation is visualized in figure B.1. The eikonal equation relates wavefronts to a seismic velocity distribution, because the constant traveltine surfaces in the traveltine field correspond to the wavefronts of a propagating wave. A ray-path in a wavefield is defined by the curve for which $\nabla \tau$ is tangent at each point. Hence, a ray-path is perpendicular to the wavefronts. The eikonal equation is usually solved in terms of ray-paths. These paths can be described by an independent ray-trace parameter along the path. This can for example be the
arc-length $l$. The travelt ime from point $A$ to point $B$ along a path with arc-length $l$ can be written as:

\[
\tau_{AB} = \int_A^B \frac{\partial \tau}{\partial l}dl(x) = \int_A^B \nabla \tau \cdot dl(x) = \int_A^B |\nabla \tau| \cos \phi dl(x) = \int_A^B s \cos \phi dl(x),
\]

for any path between $A$ and $B$. Along a ray-path $\cos \phi$ equals unity. Hence, for any path between $A$ and $B$ the following yields:

\[
\int_A^B s dl(x) \geq \tau_{AB}
\]

The equal sign applies if the path between $A$ and $B$ is a ray-path. Equality formulates Fermat’s well-known minimum time principle.
B.2 The ray-tracing system

The eikonal equation is solved in terms of ray-paths described by an independent ray-trace parameter along the ray-path. In inhomogeneous media, the simplest analytical ray-tracing system is obtained using the 'natural variable along the ray'. This variable is denoted as $\sigma$ and its dimension is $[m^2/s]$. The traveltime from point $A$ to point $B$ along a ray-path with ray-trace parameter $\sigma$ can be written as:

$$\tau_{AB} = \int_A^B q d\sigma,$$  \hspace{1cm} (B.2.12)

where $q$ is the quadratic slowness, which is related to ray-trace parameter $\sigma$ by the traveltime. Using this equation, the traveltime of the ray-tracing system is described by:

$$\frac{dt}{d\sigma} = q(x).$$  \hspace{1cm} (B.2.13)

The location of the ray-tracing system can be described by making use of $\frac{dx}{d\sigma}$:

$$\frac{dx_i}{d\sigma} = \frac{dx_i}{dt} \frac{dt}{d\sigma} = \frac{1}{p_i(x)} q(x) = p_i(x).$$  \hspace{1cm} (B.2.14)

The ray-parameter of the ray-tracing system is described by making use of $\frac{dp}{d\sigma}$:

$$\frac{dp_i}{d\sigma} = \frac{\partial p_i(x)}{\partial x} \frac{dx}{d\sigma} = p_i(x) \frac{\partial p_i(x)}{\partial x} = \frac{1}{2} \frac{\partial q^2(x)}{\partial x} = \frac{1}{2} (\nabla q)_i(x).$$  \hspace{1cm} (B.2.15)

By means of this system the variation of $x$, $p$, and $t$ along the ray can be calculated. The initial conditions for a single ray are defined in the coordinates of the initial point ($\sigma^{[0]}$):

$$x_i^{[0]} = x_i(\sigma^{[0]}), \quad p_i^{[0]} = p_i(\sigma^{[0]}), \quad t^{[0]} = t(\sigma^{[0]}) = 0,$$  \hspace{1cm} (B.2.16)

for $i = 1, 2, 3$.

Because the ray-tracing system is described by means of the 'natural variable along the ray' $\sigma$, the velocity in the model has to be defined by the quadratic slowness $q$. The quadratic slowness is a linear function of the coordinates $x_i$:

$$q = q^{[0]} + \sum_i (\nabla q)_i (x_i - x_i^{[0]}).$$  \hspace{1cm} (B.2.17)
where $(\nabla q)_i$ is the quadratic slowness gradient in direction $x_i$. The ray-parameter $p(\sigma)$ of the ray-tracing system can be solved by substituting (B.2.17) in (B.2.15):

$$
p_i(\sigma) = p_i^{[0]} + \frac{1}{2} \int_{\sigma_0}^{\sigma} \nabla_q(q_{\sigma} x_i) d\sigma
$$

$$
= p_i^{[0]} + \frac{1}{2} \left[ \nabla_q(q_{\sigma} x_i) \right]_{\sigma_0}^{\sigma_1}
$$

$$
= p_i^{[0]} + \frac{1}{2} \Delta \sigma,
$$

(B.2.18)

where $\Delta \sigma = \sigma_{11} - \sigma_{01}$. This is an analytic solution of the ray-parameter. The location $x(\sigma)$ of the ray-tracing system can be obtained by substituting (B.2.18) in (B.2.15):

$$
x_i(\sigma) = x_i^{[0]} + \int_{\sigma_0}^{\sigma} p_i d\sigma
$$

$$
= x_i^{[0]} + \int_{\sigma_0}^{\sigma} p_i^{[0]} + \frac{1}{2} \nabla_q x_i d\sigma
$$

$$
= x_i^{[0]} + p_i^{[0]} \sigma + \frac{1}{4} (\nabla_q x_i)^2 \right|_{\sigma_0}^{\sigma_1}
$$

$$
= x_i^{[0]} + p_i^{[0]} \Delta \sigma + \frac{1}{4} (\nabla_q x_i)^2 \Delta \sigma^2,
$$

(B.2.19)

which is the ray-path. The traveltime $t(\sigma)$ of the ray-tracing system can be calculated by substituting (B.2.19) in (B.2.17), and substituting the resulting equation in (B.2.13):

$$
t(\sigma) = \tau^{[0]} + \int_{\sigma_0}^{\sigma} q_i^{[0]} + \sum_i (\nabla_q) x_i d\sigma
$$

$$
= \tau^{[0]} + \int_{\sigma_0}^{\sigma} q_i^{[0]} + \sum_i \left\{ (\nabla_q x_i^{[0]} + (\nabla_q x_i^{[0]} p_i^{[0]}) \sigma + \frac{1}{4} (\nabla_q x_i^{[0]} \sigma^2) \right\} d\sigma
$$

$$
= \tau^{[0]} + q_i^{[0]} \sigma + \sum_i \left\{ (\nabla_q x_i^{[0]} \sigma + \frac{1}{2} (\nabla_q x_i^{[0]} p_i^{[0]} \sigma^2 + \frac{1}{12} (\nabla_q x_i^{[0]} \sigma^3) \right\} \right|_{\sigma_0}^{\sigma_1}
$$

$$
= \tau^{[0]} + q_i^{[0]} \Delta \sigma + \sum_i \left\{ (\nabla_q x_i^{[0]} \Delta \sigma + \frac{1}{2} (\nabla_q x_i^{[0]} p_i^{[0]} \Delta \sigma^2 + \frac{1}{12} (\nabla_q x_i^{[0]} \Delta \sigma^3) \right\}
$$

(B.2.20)

In this way the traveltime, the ray-path, and the ray-parameter within a tetrahedron can be calculated analytically. When the ray reaches the facet of a tetrahedron, it
follows the law of refraction. Snell’s law states that:

\[ s^{[1]} \sin \theta_1 = s^{[2]} \sin \theta_2, \]  

(B.2.21)
in which \( s^{[1]} \) and \( s^{[2]} \) are the slownesses on either side of the facet and \( \theta_1 \) and \( \theta_2 \) are the angles from the normal \( \mathbf{n} \) to the facet, of the incident and refracted waves respectively. By considering Figure B.2 and using Pythagoras’ law it can be derived that the quadratic values of the left hand and the right hand side can be written as:

\[
(s^{[1]} \sin \theta_1)^2 = s^{[1]^2} - (\mathbf{p}^{[1]} \cdot \mathbf{n})^2 = q^{[1]} - (\mathbf{p}^{[1]} \cdot \mathbf{n})^2, \quad \text{and} \\
(s^{[2]} \sin \theta_2)^2 = s^{[2]^2} - (\mathbf{p}^{[2]} \cdot \mathbf{n}) = q^{[2]} - (\mathbf{p}^{[2]} \cdot \mathbf{n})^2. \]  

(B.2.22)

By substituting (B.2.22) in (B.2.21) the following equation is obtained

\[ q^{[1]} - (\mathbf{p}^{[1]} \cdot \mathbf{n})^2 = q^{[2]} - (\mathbf{p}^{[2]} \cdot \mathbf{n})^2, \]  

(B.2.23)

which can be written as

\[ (\mathbf{p}^{[2]} \cdot \mathbf{n}) = \sqrt{q^{[2]} - q^{[1]} + (\mathbf{p}^{[1]} \cdot \mathbf{n})^2}. \]  

(B.2.24)

For direction \( i \) in the ray tracing system the following equation holds:

\[ p_{i}^{[2]} - p_{i}^{[1]} = \left\{ (\mathbf{p}^{[2]} \cdot \mathbf{n}) - (\mathbf{p}^{[1]} \cdot \mathbf{n}) \right\} n_{i}, \]  

(B.2.25)

By rewriting this equation and substituting (B.2.24), this becomes:

\[ p_{i}^{[2]} = p_{i}^{[1]} - \left\{ (\mathbf{p}^{[1]} \cdot \mathbf{n}) - \text{sign}(\mathbf{p}^{[1]} \cdot \mathbf{n}) \sqrt{q^{[2]} - q^{[1]} + (\mathbf{p}^{[1]} \cdot \mathbf{n})^2} \right\} n_{i}. \]  

(B.2.26)

The \( \text{sign}(\mathbf{p}^{[1]} \cdot \mathbf{n}) \) is introduced in this equation to rectify the angle of incidence. It equals 1 if the incident wave makes an acute angle with \( \mathbf{n} \), and \(-1\) if the angle of incidence is obtuse.

By using this equation, the elements of the ray parameter on the other side of the facet \( (p_{i}^{[2]}, i = 1, 2, 3) \) can be calculated. In order to include both reflected and transmitted waves, \( \pm \) is added to this equation; the ’+’ sign applies to reflected waves, and the ’−’ sign to transmitted waves:

\[ p_{i}^{[2]} = p_{i}^{[1]} - \left\{ (\mathbf{p}^{[1]} \cdot \mathbf{n}) \pm \text{sign}(\mathbf{p}^{[1]} \cdot \mathbf{n}) \sqrt{q^{[2]} - q^{[1]} + (\mathbf{p}^{[1]} \cdot \mathbf{n})^2} \right\} n_{i}. \]  

(B.2.27)
Figure B.2: The incident rays and the resulting transmitted or reflected ray. \( s^{[1]} \) and \( s^{[2]} \) are the slownesses on either side of the facet and \( \theta_1 \) and \( \theta_2 \) are the angles from the normal \( n \) to the facet. \((p^{[1]} \cdot n)\) is the projection of the incident wave on the normal and \((p^{[2]} \cdot n)\) is the projection of the refracted wave on the normal.
The derivation of the partial derivatives with respect to quadratic slowness

The partial derivative of traveltime with respect to the quadratic slowness of vertex \(k\) \((k = 1, 2, 3, 4)\) of a tetrahedron is (equation (3.4.9) is repeated here for convenience):

\[
\frac{\partial t(\sigma)}{\partial q[k]} = \frac{\partial t[0]}{\partial q[k]} + \frac{\partial t[0]}{\partial q[k]} \Delta \sigma + \sum_i \frac{\partial (\nabla q)_i}{\partial q[k]} x_i[0] \Delta \sigma + \sum_i \frac{1}{2} \frac{\partial (\nabla q)_i}{\partial q[k]} y_i[0] \Delta \sigma^2 + \sum_i \frac{1}{6} \frac{\partial (\nabla q)_i}{\partial q[k]} (\nabla q)_i \Delta \sigma^3.
\]

(C.0.1)

In order to calculated these partial derivatives, \(\frac{\partial (\nabla q)_i}{\partial q[k]}\) and \(\frac{\partial q[0]}{\partial q[k]}\) have to be derived. For the derivation of \(\frac{\partial (\nabla q)_i}{\partial q[k]}\), \((\nabla q)_i\) is required, which can be calculated by means of:

\[
\begin{bmatrix}
  x_1^2 - x_1^1 & x_2^2 - x_2^1 & x_3^2 - x_3^1 \\
  x_1^3 - x_1^1 & x_2^3 - x_2^1 & x_3^3 - x_3^1 \\
  x_1^4 - x_1^1 & x_2^4 - x_2^1 & x_3^4 - x_3^1 \\
\end{bmatrix}
\begin{bmatrix}
  (\nabla q)_1 \\
  (\nabla q)_2 \\
  (\nabla q)_3 \\
\end{bmatrix}
= \begin{bmatrix}
\end{bmatrix},
\]

in which the \(x_i^k\) is the location of vertex \(k\) in direction \(i\), and \(q^k\) is the quadratic slowness of this vertex.

By rewriting the system, the partial derivatives of the quadratic slowness gradient
in direction \( i \) ((\( \nabla q_i \))\( i \)) with respect to the quadratic slowness of vertex \( k \) \( (q^{[k]}) \) can be obtained.

When the determinant of the system is defined by:

\[
D = \det \begin{vmatrix} x_2^{[2]} - x_1^{[1]} & x_2^{[1]} - x_1^{[1]} & x_3^{[2]} - x_1^{[1]} \\ x_1^{[3]} - x_1^{[1]} & x_2^{[3]} - x_1^{[1]} & x_3^{[3]} - x_1^{[1]} \\ x_1^{[4]} - x_1^{[1]} & x_2^{[4]} - x_1^{[1]} & x_3^{[4]} - x_1^{[1]} \end{vmatrix}, \tag{C.0.2}
\]

the derivatives of \( (\nabla q)_{i} \) with respect to the quadratic slowness of vertex 1 \( (q^{[1]}) \) can calculated by:

\[
\frac{\partial (\nabla q)_1}{\partial q^{[1]}} = \frac{(x_2^{[3]}x_3^{[3]} + x_2^{[2]}x_3^{[3]} - x_2^{[3]}x_3^{[3]} - x_2^{[2]}x_3^{[3]} + x_2^{[2]}x_3^{[2]} - x_2^{[2]}x_3^{[3]})}{D},
\]

\[
\frac{\partial (\nabla q)_2}{\partial q^{[1]}} = \frac{-1(-x_2^{[2]}x_3^{[3]} + x_1^{[2]}x_3^{[3]} + x_1^{[2]}x_3^{[3]} + x_1^{[2]}x_3^{[3]} + x_1^{[2]}x_3^{[3]} - x_1^{[2]}x_3^{[3]} - x_1^{[2]}x_3^{[3]}))}{D},
\]

\[
\frac{\partial (\nabla q)_3}{\partial q^{[1]}} = \frac{(x_1^{[2]}x_2^{[3]} - x_1^{[2]}x_2^{[3]} + x_1^{[2]}x_2^{[3]} + x_1^{[2]}x_2^{[3]} + x_1^{[2]}x_2^{[3]} - x_1^{[2]}x_2^{[3]}))}{D}.
\]

By making use of these equations, the partial derivative of the initial quadratic slowness \( (q^{[0]}) \) with respect to the quadratic slowness of vertex 1 \( (q^{[1]}) \) is derived by:

\[
\frac{\partial q^{[0]}}{\partial q^{[1]}} = 1 - \left( \frac{\partial (\nabla q)_1}{\partial q^{[1]}} x_1^{[1]} + \frac{\partial (\nabla q)_2}{\partial q^{[1]}} x_2^{[1]} + \frac{\partial (\nabla q)_3}{\partial q^{[1]}} x_3^{[1]} \right). \tag{C.0.3}
\]

By inserting these equations in (C.0.1), the partial derivatives of traveltime with respect to quadratic slowness of vertex 1 can be calculated. These derivations are
also valid for the partial derivatives with respect to the quadratic slowness of vertex 2 $(g^{[2]})$:

\[
\frac{\partial (\nabla g)_1}{\partial q^{[2]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_2}{\partial q^{[2]}} = \frac{1}{D} \left( -x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_3}{\partial q^{[2]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 \right),
\]

\[
\frac{\partial q^{[0]}}{\partial q^{[2]}} = \frac{1}{D} \left( \frac{\partial (\nabla g)_1}{\partial q^{[2]}} x_2^3 + \frac{\partial (\nabla g)_2}{\partial q^{[2]}} x_2^3 + \frac{\partial (\nabla g)_3}{\partial q^{[2]}} x_2^3 \right),
\]  \hfill (C.0.4)

with respect to the quadratic slowness of vertex 3 $(g^{[3]})$:

\[
\frac{\partial (\nabla g)_1}{\partial q^{[3]}} = \frac{1}{D} \left( -x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_2}{\partial q^{[3]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_3}{\partial q^{[3]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 \right),
\]

\[
\frac{\partial q^{[0]}}{\partial q^{[3]}} = \frac{1}{D} \left( \frac{\partial (\nabla g)_1}{\partial q^{[3]}} x_2^3 + \frac{\partial (\nabla g)_2}{\partial q^{[3]}} x_2^3 + \frac{\partial (\nabla g)_3}{\partial q^{[3]}} x_2^3 \right),
\]  \hfill (C.0.5)

and finally, with respect to the quadratic slowness of vertex 4 $(g^{[4]})$:

\[
\frac{\partial (\nabla g)_1}{\partial q^{[4]}} = \frac{1}{D} \left( -x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_2}{\partial q^{[4]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 \right),
\]

\[
\frac{\partial (\nabla g)_3}{\partial q^{[4]}} = \frac{1}{D} \left( x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 - x_2^3 x_3^3 + x_2^3 x_3^3 \right),
\]

\[
\frac{\partial q^{[0]}}{\partial q^{[4]}} = \frac{1}{D} \left( \frac{\partial (\nabla g)_1}{\partial q^{[4]}} x_2^3 + \frac{\partial (\nabla g)_2}{\partial q^{[4]}} x_2^3 + \frac{\partial (\nabla g)_3}{\partial q^{[4]}} x_2^3 \right),
\]  \hfill (C.0.6)
Notation, Symbols, and Abbreviations

D.1 Notation

For the notation in this thesis the following convention has been applied:

- Scalars are given by a lower-case normal font, e.g. $x$.
- Vectors are given by a lower-case bold font, e.g. $\mathbf{x}$. One element of a vector is denoted by a subscript, e.g. $x_i$ means the $i^{th}$ element of vector $\mathbf{x}$.
- Matrices are denoted using an upper-case bold font, e.g. $\mathbf{X}$. One column of this matrix is denoted by a subscript, e.g. $\mathbf{X}_j$ means the $j^{th}$ column of matrix $\mathbf{X}$. The $j^{th}$ row of this matrix is denoted by $\mathbf{X}_j^\top$. One element of a matrix is denoted by a upper-case normal font with a double subscript, e.g. $X_{ij}$, means the $(i,j)$ element of matrix $\mathbf{X}$.
- A superscript in a matrix, vector, or scalar refers to the variable that it describes, e.g. $x^{[k]}$ indicates the scalar of point $k$, and $\mathbf{x}^{[k]}$ indicates the vector of point $k$. Note that the superscripts are surrounded by brackets in order to avoid confusion with the power of a number.

D.2 Symbols

In what follows, the first column contains the symbol or abbreviation, the second column contains the page where it is first introduced, and the third column contains
the explanation. Symbols that are used only once are explained in the text and
omitted here.

Matrices

<table>
<thead>
<tr>
<th>symbol</th>
<th>page</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>30</td>
<td>optimization matrix that describes the relation between data and model parameters</td>
</tr>
<tr>
<td>$A^+$</td>
<td>54</td>
<td>generalized inverse of the optimization matrix that describes the relation between model parameters and the data</td>
</tr>
<tr>
<td>B</td>
<td>56</td>
<td>pseudo singular value matrix</td>
</tr>
<tr>
<td>C</td>
<td>67</td>
<td>covariance matrix</td>
</tr>
<tr>
<td>$D^-$($z_0$)</td>
<td>14</td>
<td>receiver matrix</td>
</tr>
<tr>
<td>$F^+_j(z_0, z_m)$</td>
<td>14</td>
<td>focusing operator, for focusing of the sources at $(x_j, z_m)$</td>
</tr>
<tr>
<td>$F^-_j(z_m, z_0)$</td>
<td>14</td>
<td>focusing operator, for focusing of the receivers at $(x_j, z_m)$</td>
</tr>
<tr>
<td>A</td>
<td>54</td>
<td>singular value matrix</td>
</tr>
<tr>
<td>$A_B$</td>
<td>58</td>
<td>singular value matrix of matrix B</td>
</tr>
<tr>
<td>$P(z_0, z_m)$</td>
<td>14</td>
<td>data matrix</td>
</tr>
<tr>
<td>$P^+_j(z_0, z_m)$</td>
<td>14</td>
<td>CFP-gather, the sources are focused at $(x_j, z_m)$</td>
</tr>
<tr>
<td>$P^-_j(z_m, z_0)$</td>
<td>14</td>
<td>CFP-gather, the receivers are focused at $(x_j, z_m)$</td>
</tr>
<tr>
<td>$R(z_m)$</td>
<td>14</td>
<td>reflection matrix</td>
</tr>
<tr>
<td>R</td>
<td>62</td>
<td>resolution matrix</td>
</tr>
<tr>
<td>$S^+(z_0)$</td>
<td>14</td>
<td>source matrix</td>
</tr>
<tr>
<td>U</td>
<td>54</td>
<td>eigenvector matrix of the data space</td>
</tr>
<tr>
<td>$\hat{U}$</td>
<td>56</td>
<td>basis vector matrix of the data space</td>
</tr>
<tr>
<td>$U_B$</td>
<td>58</td>
<td>eigenvector matrix of the data space of matrix B</td>
</tr>
<tr>
<td>V</td>
<td>54</td>
<td>eigenvector matrix of the model space</td>
</tr>
<tr>
<td>$\hat{V}$</td>
<td>56</td>
<td>basis vector matrix of the model space</td>
</tr>
<tr>
<td>$V_B$</td>
<td>58</td>
<td>eigenvector matrix of the model space of matrix B</td>
</tr>
<tr>
<td>$W^-(z_0, z_m)$</td>
<td>14</td>
<td>upward propagation matrix</td>
</tr>
<tr>
<td>$W^+(z_m, z_0)$</td>
<td>14</td>
<td>downward propagation matrix</td>
</tr>
<tr>
<td>$W_P$</td>
<td>77</td>
<td>weighting matrix for uncertainties in the a priori information</td>
</tr>
</tbody>
</table>
## Vectors

<table>
<thead>
<tr>
<th>symbol</th>
<th>page</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta d$</td>
<td>30</td>
<td>data (traveltime) difference vector</td>
</tr>
<tr>
<td>$k$</td>
<td>181</td>
<td>wave number vector $[\frac{1}{m}]$</td>
</tr>
<tr>
<td>$\Delta m$</td>
<td>30</td>
<td>model parameter update vector</td>
</tr>
<tr>
<td>$\Delta m_p$</td>
<td>30</td>
<td>model parameter updates according to a priori information</td>
</tr>
<tr>
<td>$n$</td>
<td>45</td>
<td>normal vector</td>
</tr>
<tr>
<td>$p^{[k]}$</td>
<td>44</td>
<td>ray parameter vector; $k$ indicates the vertex the location refers to $[\frac{m}{m}]$</td>
</tr>
<tr>
<td>$p^{[0]}$</td>
<td>44</td>
<td>initial ray parameter vector $[\frac{m}{m}]$</td>
</tr>
<tr>
<td>$(\nabla q)$</td>
<td>44</td>
<td>quadratic slowness gradient $[\frac{m}{m}]$</td>
</tr>
<tr>
<td>$\hat{u}^{[k]}$</td>
<td>56</td>
<td>basis vector of the data space for the $k^{th}$ iteration</td>
</tr>
<tr>
<td>$\hat{v}^{[k]}$</td>
<td>56</td>
<td>basis vector of the model space for the $k^{th}$ iteration</td>
</tr>
<tr>
<td>$x^{[k]}$</td>
<td>43</td>
<td>location vector; $k$ indicates the vertex the location refers to $[m]$</td>
</tr>
<tr>
<td>$x^{[0]}$</td>
<td>44</td>
<td>initial location vector $[m]$</td>
</tr>
<tr>
<td>$\Delta x^{[f]}$</td>
<td>49</td>
<td>update in location of (focal) point $f$ $[m]$</td>
</tr>
</tbody>
</table>

## Scalars

<table>
<thead>
<tr>
<th>symbol</th>
<th>page</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A(x)$</td>
<td>181</td>
<td>wave amplitude as a function of location</td>
</tr>
<tr>
<td>$C(h^{[ij]})$</td>
<td>77</td>
<td>covariance between sample $i$ and $j$</td>
</tr>
<tr>
<td>$C(h_s)$</td>
<td>77</td>
<td>covariance for specific distance $h_s$</td>
</tr>
<tr>
<td>$F$</td>
<td>49</td>
<td>number of focal points</td>
</tr>
<tr>
<td>$h^{[ij]}$</td>
<td>77</td>
<td>distance between sample $i$ and $j$ $[m]$</td>
</tr>
<tr>
<td>$h_s$</td>
<td>77</td>
<td>specific distance between samples $[m]$</td>
</tr>
<tr>
<td>$L$</td>
<td>49</td>
<td>number of quadratic slowness parameters</td>
</tr>
<tr>
<td>$M$</td>
<td>49</td>
<td>number of model parameters</td>
</tr>
<tr>
<td>$m_j^{[s]}$</td>
<td>77</td>
<td>model parameter value at sample $j$</td>
</tr>
<tr>
<td>$m_p^{[j]}$</td>
<td>77</td>
<td>a priori value at model parameter $j$</td>
</tr>
<tr>
<td>$N$</td>
<td>49</td>
<td>number of data</td>
</tr>
</tbody>
</table>
Notation, Symbols, and Abbreviations

\[ q \quad 43 \quad \text{quadratic slowness} \left[ \frac{a^2}{m^2} \right] \]
\[ q^{[0]} \quad 44 \quad \text{initial quadratic slowness} \left[ \frac{a^2}{m^2} \right] \]
\[ q^{[k]} \quad 50 \quad \text{quadratic slowness at vertex } k \left[ \frac{a^2}{m^2} \right] \]
\[ \Delta q^{[l]} \quad 49 \quad \text{update in quadratic slowness for vertex } l \left[ \frac{a^2}{m^2} \right] \]
\[ S \quad 76 \quad \text{number of a priori samples} \]
\[ s \quad 44 \quad \text{slowness} \left[ \frac{a^2}{m^2} \right] \]
\[ \sigma \quad 44 \quad \text{ray tracing parameter} \left[ \frac{a^2}{m^2} \right] \]
\[ \sigma^{[0]} \quad 44 \quad \text{initial ray tracing parameter} \left[ \frac{m^2}{a^2} \right] \]
\[ \Delta \sigma \quad 45 \quad \text{change in ray tracing parameter} \left[ \frac{m^2}{a^2} \right] \]
\[ t \quad 44 \quad \text{(travel)time} \quad [s] \]
\[ t^{[0]} \quad 44 \quad \text{initial (travel)time} \quad [s] \]
\[ \Delta t^{[n]} \quad 49 \quad \text{misfit in (travel)time for ray } n \quad [s] \]
\[ \Delta t_{\text{rms}} \quad 97 \quad \text{rms of the traveltime residuals} \quad [s] \]
\[ \tau (x) \quad 43 \quad \text{eikonal: traveltime as a function of location} \quad [s] \]
\[ w^{[j]} \quad 77 \quad \text{weight for sample } j \]
\[ \omega \quad 181 \quad \text{frequency} \quad \left[ \frac{1}{s} \right] \]
\[ z_0 \quad 14 \quad \text{interface representing surface, where sources and receivers are located} \]
\[ z_b \quad 81 \quad \text{interface representing the ocean bottom, where the receivers of OBS data are located} \]
\[ z_m \quad 14 \quad \text{interface in the subsurface, where the focusing is done} \]

### D.3 Abbreviations

<table>
<thead>
<tr>
<th>abbreviation</th>
<th>page</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>9</td>
<td>two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>9</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>AVP</td>
<td>13</td>
<td>Angle versus Ray parameter</td>
</tr>
<tr>
<td>CFP</td>
<td>13</td>
<td>Common Focal Point</td>
</tr>
<tr>
<td>CG</td>
<td>55</td>
<td>Conjugate Gradients</td>
</tr>
<tr>
<td>CIG</td>
<td>4</td>
<td>Common Image Gather</td>
</tr>
<tr>
<td>DTS</td>
<td>9</td>
<td>Differential Time Shift</td>
</tr>
<tr>
<td>LSQR</td>
<td>55</td>
<td>Large Sparse QR factorization</td>
</tr>
<tr>
<td>MVA</td>
<td>4</td>
<td>Migration Velocity Analysis</td>
</tr>
<tr>
<td>OBS</td>
<td>104</td>
<td>Ocean Bottom Station</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Page</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>P-wave</td>
<td>80</td>
<td>Pressure wave</td>
</tr>
<tr>
<td>rms</td>
<td>14</td>
<td>root mean square</td>
</tr>
<tr>
<td>SVD</td>
<td>54</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>S-wave</td>
<td>80</td>
<td>Shear wave</td>
</tr>
<tr>
<td>WRW</td>
<td>15</td>
<td>refers to down-going ( W^+ ), reflected ( R ), and up-going wavefield ( W^- ) (see symbols)</td>
</tr>
</tbody>
</table>
Bibliography


——— 1997b, Pushing the limits of seismic imaging, part II: Integration of prestack migration, velocity estimation, and avo analysis: Geophysics, 62, no. 3, 954–969.


Li, X. J., 2000, Spacing Control and Sliver-free Delaunay mesh: see http://www.csam.iit.edu/~xli/publications.html.


Nolet, G., 1985, Solving or resolving inadequate and noisy tomographic systems: Journal of computational physics, 61, 463–482.


Trinks, I., 2001, Traveltime tomography using irregular parameterised grids: see bullard.esc.cam.ac.uk/ trinks/.


Versteeg, R. J., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, 58, no. 6, 873–879.


BIBLIOGRAPHY
Summary

Tomographic inversion of focusing operators

Seismic images of the structure of the earth are a prerequisite for finding new hydrocarbon reservoirs and monitoring the production of these hydrocarbons. The quality of a seismic image is highly dependent on the accuracy of the velocity model of the subsurface. Traditional imaging consists of an iterative process between obtaining the image using a velocity model, and updating this model by investigating the property of the image.

The Common Focal Point (CFP) method makes it possible to analyze and treat seismic data in a fundamentally different way. The CFP method is able to determine focusing operators from conventional seismic data by using the 'principle of equal traveltime'. A focusing operator can be seen as a response from a secondary source (called a focal point) in the subsurface towards receivers at the surface. By means of these operators the two-way seismic data can be focused at corresponding focal points in the subsurface. Although the focusing operators can focus the data, the position of the focused data is yet unknown. The positioning of the events in the subsurface requires a velocity model. In the CFP method, this velocity model is estimated directly from the focusing operators.

In summary, the CFP method uses a two-step approach in which 1) two-way reflection data are transformed into one-way data by estimating focusing operators, and 2) the focusing operators are used to estimate the velocity model by tomographic inversion. This second step, the tomographic inversion of focusing operators, is the subject of this thesis.

An important difference between conventional seismic data and focusing operators
is that the seismic data obviously are in two-way traveltimes, and the focusing operators represent one-way traveltimes. The advantage of the one-way focusing operators is that they are less-complicated than the two-way data. This aspect makes them very suitable for velocity model estimation.

A very important aspect of velocity estimation problems is the way the velocity model is represented. Conventional methods either introduce a bias in the solution by using a geology based parameterization, or result in over-parameterization by using a dense spatial grid. In this thesis an attempt is made to circumvent these shortcomings by making use of an irregular spatial grid, in which the number of parameters is determined by the data, the data being the one-way traveltimes of the focusing operators.

Summarized, this research contains two important new aspects. First, the use of focusing operators in (3D) velocity model estimation, and second the data-driven approach of the method.

In this thesis, three main objectives are considered:

[1] Develop a method for (3D) tomographic inversion of focusing operators.


In order to satisfy the first objective, a tomographic inversion method is developed, that contains four main steps; parameterization, forward modeling, optimization, and regularization. The subsurface is described by model parameters that influence the traveltimes: the velocity and the location of the focal points. The velocity is defined on irregular grid-points, which are connected by Delaunay triangulation, in which the quadratic slowness is linearly interpolated. In 3D this results in a model consisting of tetrahedra. The use of quadratic slowness facilitates the involved forward modeling algorithm. The focal points are defined independently of the velocity grid-points. This flexible parameterization makes it possible to describe any model of the subsurface and to make the method data-driven. The forward modeling of focusing operators is performed by a ray-based method. The focal points are considered to be sources from which rays are calculated towards receivers at the surface. The ray-tracing is carried out analytically, so both the traveltimes and the partial derivatives are calculated analytically. The optimization of the model parameters is performed by the LSQR method. This is an efficient iterative method, that can calculate 'pseudo' singular values by which the inversion system can be analyzed.

An important aspect of the presented method is that it is fully data-driven, which is related to the second objective of this research. This aspect is realized by the regularization of the inverse problem, which is accomplished by a resolution dependent parameterization. The resolution of the parameters in the optimization is translated
to the spatial resolution in the parameterization. This means that the model parameters that have a low resolution are removed and in regions containing parameters that have a high resolution extra parameters are added. As the parameterization depends on the resolution, and the resolution depends on the data, the method becomes data-driven. In this way, the inverse problem becomes well-determined and no bias is introduced by the user. For the resolution-dependent parameterization the relative resolution of the parameters, which can be calculated by means of the LSQR method, is used as a qualitative criterion.

The presented method has been evaluated on synthetic data in both 2D and 3D, and on real data in 2D. The method results in an accurate update of both velocities and focal point locations, and is capable of dealing with complex velocity models. A variety of structures can be solved accurately as long as the resolution-dependent parameterization is fine enough to describe it. Minimal user interaction is required, because the method is driven by the data.

The final objective of this research is to explore the advantages of the proposed method. Obviously, not all the possibilities are explored, but this research focuses on three additional modules: 1) Inclusion of a priori information, 2) Joint inversion of P and S-wave operators, and 3) Evaluation of the inversion result by focal point clouds.

It has been shown that a priori information about the model parameter values can be included in the tomographic inversion. This can be achieved by interpolating this information by means of Kriging and taking the interpolated information into account as an extra constraint in the inversion. The a priori information, e.g. from well-log information, can bring a solution in areas that are not well resolved by the focusing operators.

Regarding the second module, a method is developed that can jointly invert P-wave and S-wave operators. The CFP method is pre-eminently suited to image mode converted data, because this method addresses upward and downward traveling energy separately. The estimation of P and S-wave operators in multi-component seismic data always results in operator pairs related to the same focal points. As a consequence, these operator pairs can be jointly inverted, in which a P and S-velocity model are estimated simultaneously, and the focal point locations are the same in both models.

Finally, an analysis tool is introduced being the focal point cloud, which is a map of the residual traveltime errors translated into position errors at the focal point. The focal point clouds can evaluate the adequacy of the velocity model to explain the data. The shape of the focal point clouds makes it possible to diagnose the origin of the residual between the observed focusing operators and the operators calculated in the velocity model. Velocity errors can be distinguished from focal point location errors, and anisotropy and 3D effects in 2D models can also be recognized.

Barbara Ellen Cox
Samenvatting

Tomografische inversie van focussingoperatoren

Seismische afbeeldingen van de structuur van de aarde zijn een vereiste voor het vinden van olie- en gas-reservoirs en het monitoren van de productie uit deze reservoirs. De kwaliteit van de seismische afbeelding is afhankelijk van de nauwkeurigheid van het snelheidsmodel van de ondergrond. Normaal gesproken worden seismische afbeeldingen verkregen uit een iteratief proces tussen het maken van een afbeelding met behulp van een snelheidsmodel en het verbeteren van dit model met behulp van de eigenschappen van de afbeelding.

De Gemeenschappelijke Brandpunt (Common Focal Point, CFP) methode maakt het mogelijk om seismische data op een fundamenteel andere manier te analyseren en behandelen. Uit conventionele seismische data kan de CFP methode focussingoperatoren verkrijgen door gebruik te maken van het "principe van gelijke looptijd". Een operator kan beschouwd worden als een respons van een secundaire bron (ook wel focussingpunt genoemd) in de ondergrond naar ontvangers aan het aardoppervlak. Met behulp van deze operatoren kunnen seismische twee-weg data gefocuseerd worden in de bijbehorende focussingpunten. Alhoewel de operatoren de data kunnen focuseren blijft de positie van de gefocuseerde data onbekend. Om die te kunnen bepalen is een snelheidsmodel nodig. In de CFP methode wordt dit snelheidsmodel geschat uit de focussingoperatoren.

Eigenlijk worden met de CFP methode seismische afbeeldingen in twee stappen verkregen, waarbij 1) twee-weg data getransformeerd worden naar één-weg data door het schatten van focussingoperatoren en 2) de operatoren gebruikt worden voor het schatten van een snelheidsmodel. Deze tweede stap, de tomografische inversie van focussingoperatoren, is het onderwerp van dit proefschrift.
Een belangrijk verschil tussen conventionele seismische data en focuseringsoperatoren is dat de seismische data twee-weg looptijden bevatten en de focuseringsoperatoren één-weg looptijden repräsentieren. Het voordeel van één-weg operatoren is dat ze minder complex zijn dan twee-weg data. Dit maakt de operatoren erg geschikt voor snelheidsanalyse.

Een belangrijk aspect van snelheidsanalyse is de manier waarop het snelheidsmodel gerepresenteerd wordt. Conventionele methoden introduceren vaak een bias in de oplossing door gebruik te maken van een parameterisatie die door de gebruiker gedefinieerd is, of resulteren in over-parameterisatie door gebruik te maken van een fijnbemonsterd spatiaal grid. In dit proefschrift wordt dit probleem aangepakt door gebruik te maken van een onregelmagig spatiaal grid waarin het aantal parameters door de data bepaald wordt. De data is in dit geval de verzameling looptijden van de één-weg focuseringsoperatoren.

Kortom, dit onderzoek bevat twee belangrijke nieuwe aspecten. Het eerste aspect is het gebruik van focuseringsoperatoren voor het schatten van een (3D) snelheidsmodel. Het tweede aspect is de data-gedreven aanpak van de methode.

In dit proefschrift worden drie doelstellingen beschouwd:

1. Ontwikkel een methode voor (3D) tomografische inversie van focuseringsoperatoren.
2. Zorg dat de methode gedreven wordt door de data.
3. Onderzoek de extra mogelijkheden van tomografische inversie van focuseringsoperatoren.

Om de eerste doelstelling te bereiken is er een inversie methode ontwikkeld die uit vier stappen bestaat: parameterisatie, voorwaartse modelleren, optimalisatie en regularisatie. Het model van de ondergrond wordt beschreven met parameters die de looptijden bepalen: de snelheid en de locatie van de focuseringspunten. De snelheid wordt gedefinieerd op onregelmatige gridpunten die verbonden worden met behulp van Delaunay triangulairisatie, waarbinnen het snelheidsveld bepaald wordt met een lineair geïnterpooleerde kwadratische traagheid. In 3D resulteert deze parameterisatie in tetraëders. De kwadratische traagheid is gebruikt zodat het voorwaartse modellerings algoritme eenvoudig analytisch af te leiden is. De focuseringspunten worden onafhankelijk van de snelheids-gridpunten gedefinieerd. Deze flexibele parameterisatie maakt het mogelijk om bijna ieder model van de ondergrond te beschrijven en om de methode datagedreven te maken. De voorwaartse modellering van de focuseringsoperatoren wordt gedaan met behulp van een stralenmethode. De focuseringspunten worden beschouwd als bronnen waarvandaan stralen berekend worden naar ontvangers aan het aardoppervlak. De berekening worden analytisch gedaan, dus zowel de looptijden als de partiële afgeleiden kunnen analytisch berekend worden. Voor de optimalisatie van de modelparameters wordt de LSQR
methode gebruikt. Dit is een efficiënte iteratieve methode die 'pseudo' singulièr waarden kan berekenen, waarmee het inversiesysteem geanalyseerd kan worden.

Een belangrijk aspect van de gepresenteerde methode is dat het compleet datagedreven is, hetgeen gerelateerd is aan de tweede doelstelling van dit onderzoek. Dit aspect komt naar voren in de regularisatie van het inversieprobleem, hetgeen be- werkstelligd wordt door resolutieafhankelijke parameterisatie. De resolutie van de parameters in de optimalisatie wordt vertaald naar de spatiële resolutie van de pa- ramaeterisatie. Dit houdt in dat de modelparameters die een lage resolutie hebben verwijderd worden en dat in gebieden die parameters met een hoge resolutie be- vatten extra parameters toegevoegd worden. Aangezien de parameterisatie afhangt van de resolutie en de resolutie op zijn beurt weer afhankt van de data, wordt de methode daarmee dus datagedreven. Op deze manier is het inversieprobleem goed bepaald en wordt er geen bias geïntroduceerd door de gebruiker. In de resolutieaf- hankelijke parameterisatie wordt de relatieve resolutie van de parameters gebruikt als een kwalitatief criterium. Deze kan berekend worden met de LSQR methode.

De gepresenteerde methode is geëvalueerd op synthetische data in zowel 2D en 3D en op echte data in 2D. De methode resulteert in een correcte update van snelheden en focuseringspunt-locaties en is in staat om complexe snelheidsmodellen te schatten. Verschillende ondergrond structuren kunnen correct opgelost worden, zolang de resolutieafhankelijke parameterisatie fijn genoeg is om ze te beschrijven. Er is bijna geen gebruiker-interactie nodig, omdat de methode gestuurd wordt door de data.

De laatste doelstelling van dit onderzoek is om de verdere mogelijkheden van de voorgestelde methode te onderzoeken. Natuurlijk zijn niet alle mogelijkheden be- schouwd, maar dit onderzoek neemt drie toegevoegde modules onder de loep: 1) toe- voeging van a-priori informatie, 2) gezamenlijke inversie van P- en S-golf operatoren en 3) analyse van het inversieresultaat door gebruik te maken van focuseringspunt- wolken.
Het onderzoek laat zien dat er a-priori informatie over modelparameters toegevoegd kan worden aan de tomografische inversie. Dit kan gedaan worden door deze infor- matie te interpoleren door middel van Kriging en deze geïnterpoleerde informatie als extra voorwaarde in de inversie mee te nemen. A-priori informatie, bijvoorbeeld uit boorputgegevens, kan een oplossing geven in gebieden die niet goed beschreven zijn door de operatoren. Met betrekking tot de tweede module is er een methode ontwikkeld die P- en S-golf operatoren gemeenschappelijk kan inverteren. De CFP methode is erg geschikt om geconverteerde data te analyseren, omdat deze methode de opwaartse en neerwaartse energie apart beschouwt. De schatting van P- en S-golf operatoren in meervoudige- componenten seismische data resulteert altijd in operatorparen die gerelateerd zijn aan hetzelfde focuseringspunt. Hierdoor kunnen de operatorparen gezamenlijk geïntegreerd worden, waarbij het P- en S-snelheidsmodel gelijktijdig worden geschat en de focuseringspunt-locaties in beide modellen hetzelfde zijn.
Tenslotte is er een analysemethode ontwikkeld, zijnde de focuseringspunt-wolken.
Deze wolken zijn een projectie van de residuele looptijdfouten, vertaald in locatiefouten rond het focuseringspunt. Met deze focuseringspunt-wolken kan geëvalueerd worden of het snelheidsmodel toereikend is om de data te verklaren. De vorm van de focuseringspunt-wolken maakt het mogelijk om de oorsprong van het residu tussen de waargenomen en de focuseringsoperatoren gemodelleerd in het snelheidsmodel te identificeren. Snelheidsfouten kunnen onderscheiden worden van fouten in de locaties van de focuseringspunten. Ook anisotropie en 3D effecten in 2D modellen kunnen herkend worden.

Barbara Ellen Cox
Barbara Cox was born in Melick and Herkenbosch on January 31, 1975. She attended secondary school at 'B.C. Schöndeln' in Roermond, where she received the 'VWO diploma' in 1993. In September of that year she started the study of Geophysics at the Utrecht University and received her M.Sc. degree in 1998 (graded 'with distinction'). Her M.Sc. thesis received the BP-Amoco award for "Most outstanding Geoscience student 2000".

After graduation she joined the DELPHI consortium at the Faculty of Applied Sciences at the Delft University of Technology. From June 1999 to November 2003 she has been conducting her Ph.D. research, which led to this thesis. During this period, she has made a 3 month research visit to Stanford University. In 2004 she will join Shell International Exploration and Production.
Dankwoord

Dit dankwoord is een kans die ik zelden krijg en misschien het meest gelezen onderdeel van het proefschrift. Hoewel ik alleen mijn eigen naam op de omslag heb gezet, wil ik deze belangrijke pagina’s gebruiken om de mensen te bedanken die een belangrijke bijdrage aan dit proefschrift hebben geleverd.


The sponsors of the DELPHI consortium are gratefully acknowledged for their encouragement, discussions, and financial support. In particular, I thank Norsk Hydro for providing me with the dataset presented in Chapter 6. I am grateful to the people of the S Ma rt joint venture for giving me the opportunity to visit Stanford University. I would like to express my appreciation to the people of the Stanford Exploration Project for showing me both their scientific and their Californian world.

Verder ben ik mijn oud collega’s, Kees, Karin, John, Ewoud, Arno en Johan, veel dank verschuldigd voor de sopele introductie in de wereld die DELPHI heet. Natuurlijk wil ik ook mijn huidige collega’s, Paul, Maurice, Sandra, Pascal, Ayon, Mustafa, Remco, Gerrit en Edith, bedanken voor het aangenaam maken van de 40+ uur in de week. Bovendien bedank ik Edo, Paul K. en Henry voor het draaiende houden van het computersysteem.
En dan mijn familie en vrienden, die voor het leven naast mijn promotieonderzoek zorgden. Mijn vader wil ik graag bedanken voor het leggen van de basis, mijn moeder in het bijzonder voor de voortzetting ervan. Zowiezo deel je het grootste deel van je leven met broers en zussen, maar ook de afgelopen vier jaar waren ze er: Annemarie en Fransje, bedankt daarvoor.

Met een grote glimlach op mijn gezicht bedank ik mijn vrienden voor de telefoontjes, eet-dates, JdeW bezoeken, salsa-avonturen, weekendjes weg, ski-vakanties, etc. die broodnodig waren als tegenhanger voor mijn uurtjes op de TU achter m’n compu. Dit dankwoord wil ik beëindigen met de voor mij belangrijkste persoon, waarmee ik zowel mijn thuis als ‘het promoveren’ kon delen: Roald, bedankt dat je er was.