

# Time Domain 2D Elastic Full Waveform Tomography

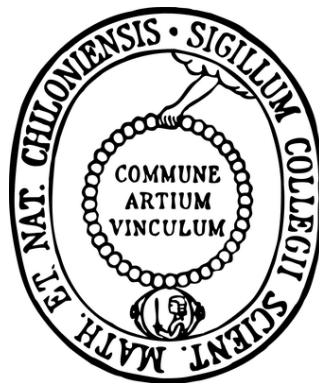
Dissertation  
zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Christian-Albrechts-Universität  
zu Kiel

vorgelegt von

Daniel Köhn

Kiel

März 2011





Referent:  
Koreferent:  
Tag der mündlichen Prüfung:  
Zum Druck genehmigt:

Prof. Dr. Wolfgang Rabbel  
Prof. Dr. Thomas Bohlen  
08.07.2011  
Kiel, 08.07.2011

Der Dekan





Für Sava-Birgit Germaine Denise De Nil  
אני אוהב אותך ואני מתגעגע אליך



---

## Abstract

With the increasing performance of parallel supercomputers full waveform tomography (FWT) approaches can reduce the misfit between recorded and modelled data, to deduce a very detailed physical model of the underground. In recent years acoustic waveform tomography became a very popular tool to image underground structures. However, acoustic waveform inversion has the disadvantage, that only P-waves can be inverted. It can not invert for S-waves, surface or interface waves.

The primary objective of this thesis is to implement a parallel 2D time domain elastic FWT code for an independent inversion of the two isotropic elastic material parameters and the density. The resulting code is benchmarked and tested against a wide range of test problems, from very simple geologically unrealistic problems to study the resolution and ambiguity of FWT to geological complex media.

The choice of model parameters to be inverted has a large impact on the resolution and ambiguity of the inversion results. To investigate the influence of different model parametrizations two elastic models for the parameter sets using Lamé parameters and seismic velocities are compared. The models consist of an elastic layer and a half space. Embedded in the layer are different geometrical bodies for the different model parameters. Synthetic multicomponent datasets are calculated for the test problem and inverted using a starting model with the correct elastic material parameters for the layer and the half space but without the geometrical structures. The Lamé parameters, respectively seismic velocities could be reconstructed very well without any ambiguity. Surprisingly the choice of parameters has a huge impact on the density results. Using Lamé parameters the density model could be recovered, but it is disturbed by extremely large shear module artefacts. When using seismic velocities as model parameters a stronger ambiguity is present, but the artefacts have the same values as the true density model.

Beside the model parametrization the starting point in the parameter space (starting model) has a large impact on the direction of the inversion procedure and therefore the quality of the inversion result. To find plausible starting models for a simple geological model, an evolution strategy is used to scan the long wavelength parameter space. The inversion results for different starting models are compared.

To test the feasibility of 2D elastic FWT to invert "observed" data, a synthetic dataset is generated for a simple layered subsurface model, which exhibits 3-D geometrical wave propagation effects. While the observed data shows a 3D geometrical spreading, the 2D FWT code can only model 2D geometrical spreading. A simple approach to transform the spreading from 3-D to 2-D is used to evaluate the applicability of FWT.



---

## Zusammenfassung

Mit zunehmender Leistung moderner Supercomputer werden Wellenformtomographie Ansätze (FWT) zur Minimierung der Differenz zwischen gemessenen und modellierten Daten verwendet, um ein detailliertes physikalisches Modell des Untergrundes abzuleiten. In den letzten Jahren entwickelte sich die akustische Wellenform Tomographie zu einem weitverbreiteten Verfahren zur Abbildung von Untergrundstrukturen. Ein Nachteil der akustischen FWT ist jedoch die Beschränkung auf P-Wellen. S-Wellen, Oberflächen- oder Grenzflächenwellen sind nicht invertierbar.

Hauptziel dieser Arbeit ist die Implementierung eines parallelen elastischen 2D FWT Zeitbereichscodes für die unabhängige Inversion der isotropen elastischen Materialparameter sowie der Dichte. Der Code wird gegen ein großes Spektrum von Testproblemen getestet und gebenchmarkt, von sehr einfachen Auflösungstests bis hin zur Inversion von komplexen geologischen Medien.

Die Wahl der zu invertierenden Modellparameter hat einen großen Einfluss auf die Auflösung und Mehrdeutigkeit der Inversionsergebnisse. Um diese Mehrdeutigkeit im Detail zu untersuchen werden zwei elastische Modelle für die Lamé Parameter und seismischen Geschwindigkeiten miteinander verglichen. Die Modelle bestehen aus einer elastischen Schicht über einem Halbraum. Eingebettet in die Schicht befinden sich, abhängig vom verwendeten Materialparameter, verschiedene geometrische Objekte. Als Startmodell für die Inversion wird das Hintergrundmodell ohne die geometrischen Störkörper verwendet. Die Lamé Parameter, bzw. seismischen Geschwindigkeiten können ohne Mehrdeutigkeit rekonstruiert werden. Die Wahl der Materialparameter hat jedoch einen großen Einfluss auf das invertierte Dichtemodell. Mit der Lamé Parametrisierung wird das Dichtemodell zwar rekonstruiert, jedoch durch dominante Schermodulartefakte verschlechtert. Bei Verwendung der Parametrisierung mit seismischen Wellengeschwindigkeiten tritt eine größere Mehrdeutigkeit auf, allerdings besitzen diese Artefakte den gleichen Wert wie das wahre Dichtemodell. Neben der Modellparametrisierung spielt die Wahl des Startmodells eine entscheidende Rolle bei der Auflösung der FWT. Um mögliche Startmodelle für ein einfaches geologisches Modell zu finden, wird der langwellige Parameterraum mit einer Evolutionsstrategie abgesucht und die Inversionsergebnisse für unterschiedliche Startmodelle miteinander verglichen.

Schließlich wird ein Aspekt der FWT für die reale Datenanwendung genauer untersucht. Dazu wird für ein einfaches geschichtetes Medium ein Datensatz mit 3D geometrischen Spreading Effekten berechnet. Der 2D FWT Code kann jedoch nur Daten mit 2D Spreading modellieren. Mit einer einfache 3D-2D Transformation wird die Anwendbarkeit der 2D FWT auf 3D Daten untersucht.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Objectives of this thesis . . . . .	2
1.2	Structure of the thesis . . . . .	2
<b>2</b>	<b>The Seismic Forward Problem</b>	<b>5</b>
2.1	Equations of motion for an elastic medium . . . . .	5
2.1.1	Conservation of Momentum . . . . .	5
2.1.2	Infinitesimal Deformation in the rock . . . . .	8
2.1.3	Analysis of the stress state . . . . .	10
2.1.4	Linear stress-strain relationship . . . . .	13
2.1.5	The elastic wave equations . . . . .	15
2.2	Solution of the elastic wave equation by finite differences . . . . .	16
2.2.1	Discretization of the wave equation . . . . .	16
2.2.2	Accuracy of FD operators . . . . .	19
2.2.3	Initial and Boundary Conditions . . . . .	20
2.3	Numerical Artefacts and Instabilities . . . . .	24
2.3.1	Grid Dispersion . . . . .	24
2.3.2	The Courant Instability . . . . .	28
2.4	Parallelization of the 2D elastic FD Code . . . . .	30
<b>3</b>	<b>The adjoint problem</b>	<b>33</b>
3.1	What is an "optimum" model ? . . . . .	33
3.2	How to find an optimum model . . . . .	35
3.3	Calculation of the gradient direction $\frac{\partial E}{\partial \mathbf{m}}$ . . . . .	37
3.4	Estimation of an optimum step length $\mu_n$ . . . . .	44
3.5	Nonlinear Conjugate Gradient Method . . . . .	52
3.6	The elastic FWT algorithm . . . . .	54

## CONTENTS

---

<b>4</b>	<b>Simple example: A spherical low velocity anomaly</b>	<b>57</b>
<b>5</b>	<b>The influence of model parametrization</b>	<b>69</b>
5.1	The gradient direction $\frac{\partial E}{\partial \mathbf{m}}$ in terms of different model parametrizations	69
5.2	The CTS test problem . . . . .	70
<b>6</b>	<b>A geological test problem - the elastic Marmousi2 model</b>	<b>75</b>
6.1	The complex Marmousi2 model . . . . .	75
6.1.1	Acquisition geometry and FD model . . . . .	78
6.1.2	Elastic wave propagation in the complex Marmousi model . . . . .	78
6.1.3	FWT of the complex Marmousi model . . . . .	80
6.2	The simple Marmousi2 model . . . . .	88
6.2.1	Acquisition geometry and FD model . . . . .	88
6.2.2	Elastic wave propagation in the simple Marmousi model . . . . .	88
6.2.3	FWT of the simple Marmousi model . . . . .	91
<b>7</b>	<b>1D Starting model estimations with Evolution Strategy</b>	<b>97</b>
7.1	Basic starting models . . . . .	97
7.2	Evolution Strategy . . . . .	99
7.3	Starting model estimations by (10+45)-ES . . . . .	101
<b>8</b>	<b>Reducing the nonlinearity of the elastic inverse problem I: Time Windowing</b>	<b>117</b>
<b>9</b>	<b>Reducing the nonlinearity of the elastic inverse problem II: Frequency Filtering</b>	<b>123</b>
<b>10</b>	<b>The Influence of the starting model</b>	<b>127</b>
<b>11</b>	<b>Inversion of 3D data with a 2D FWT code</b>	<b>133</b>
11.1	3D to 2D data transformation . . . . .	134
11.2	The simple 1D Marmousi model . . . . .	136
11.3	The complex 2D Marmousi model . . . . .	142



## CONTENTS

---

12 Summary, Conclusions and Outlook	149
A 1 <sup>st</sup> order perturbation of the elastic equations of motion	155
B Green's Function Solution	159
C Bibliography	161
D Nomenclature	167
E Danksagung (Acknowledgements)	173
Eidesstattliche Erklärung	177



# 1

## Introduction

The aim of Full Waveform Tomography (FWT) is to estimate the elastic material parameters in the underground. This can be achieved by minimizing the misfit energy between the modelled and field data using a gradient optimization approach. Because the FWT uses the full information content of each seismogram, structures below the seismic wavelength can be resolved. This is a tremendous improvement in resolution compared to traveltime tomography (Pratt et al. (2002)).

The concept of full waveform tomography was originally developed by Albert Tarantola in the 1980s for the acoustic, isotropic elastic, and viscoelastic case (Tarantola (1984b,a, 1986, 1988)). First numerical implementations were realized at the end of the 1980s (Gauthier et al. (1986), Mora (1987), Pica et al. (1990)), but due to limited computational resources, the application was restricted to simple 2D synthetic test problems and small near offset datasets. At the beginning of the 1990s the original time domain formulation was transferred to a robust frequency domain approach (Pratt and Worthington (1990), Pratt (1990)). With the increasing performance of supercomputers moderately sized problems could be inverted with frequency domain approaches.

A spectacular result to prove the application of acoustic FWT on laboratory scale was presented by Pratt (1999) for ultrasonic tomography measurements on a simple block model. In a numerical blind test Brenders and Pratt (2007) achieved a very good agreement between their inversion result and the unknown true P-wave velocity model. The parallelization and performance optimizations of the frequency domain approach (see e.g. Sourbier et al. (2009a), Sourbier et al. (2009b)) lead to a wide range of acoustic FWT applications for problems on different scales, from the global scale, crustal scale over engineering and near surface scale, down to laboratory scale (Pratt (2004)).

Beside the application to geophysical problems, the acoustic FWT is also used to improve the resolution in medical cancer diagnostics (Pratt et al. (2007)). However,

## CHAPTER 1. INTRODUCTION

---

all these examples are restricted to the inversion of the acoustic material parameters: P-wave velocity, density and additionally the viscoacoustic damping  $Q_p$  for the P-waves. Even today the independent 2D FWT of all three isotropic elastic material parameters is still a challenge. Most elastic approaches invert for P-wave velocity only and use empirical relationships to deduce the distribution of S-wave velocity and density (Shipp and Singh (2002); Sheen et al. (2006)). Recently some authors also investigated the independent multiparameter FWT in the frequency domain (Choi et al. (2008a,b); Brossier (2009)).

### 1.1 Objectives of this thesis

---

The primary objective of this thesis is to implement a parallel 2D time domain elastic FWT code for an independent inversion of the two isotropic elastic material parameters and the density. The resulting code is benchmarked and tested against a wide range of test problems, from very simple geologically unrealistic problems to study the resolution and ambiguity of FWT to geological complex media. Two problems which are important for the later data application are investigated in more detail: The estimation and influence of the starting model and the inversion of 3D data with a 2D FWT code.

### 1.2 Structure of the thesis

---

Beside this introduction and the final conclusion this thesis is divided in 10 chapters.

- I begin this thesis with a very detailed description of the theory of elastic waves. This is essential to understand the elastic inverse problem and its limitations. Chapter 2 covers the derivation of the conservation of momentum equation for a continuum, defines important terms like the strain tensor, the description of the stress state and the linear relationship between stress and strain to finally derive the equations of motion for an elastic medium. To solve these equations for complex media, numerical approaches are required. A robust and easy to implement way to solve the forward problem is the finite-difference approach. The theory, applications and limitations of this approach are discussed in chapter 2.2.
- The third chapter covers the solution of the elastic inverse problem. This highly nonlinear problem is solved by minimizing the misfit energy between the

---

## 1.2. STRUCTURE OF THE THESIS

---

modelled and field data using a local gradient method. After the derivation of the gradient direction for the elastic adjoint problem, the influence of the adaptive step length estimation and the conjugate gradient method on the convergence speed of the FWT code are tested.

- The different steps of the FWT algorithm are illustrated in chapter 4 using a very simple test problem: A spherical low velocity anomaly in a homogeneous fullspace.
- The model parametrization has a major impact on the resolution and ambiguity of the elastic model parameters. In chapter 5, a detailed study using a nonlinear model consisting of different geometrical bodies for the elastic model parameters shed some light on what is the optimum model parametrization for the elastic FWT.
- Chapter 6 deals with the application of the FWT on realistic geological test problems - a very complex 2D geological and a more simpler approximately 1D geological model.
- In the preceding chapters it is assumed, that the starting point of the model optimization is near the minimum of the misfit function describing the true model. In chapter 7 a global search strategy will be applied to estimate possible starting models in the parameter space for an approximately 1D geology.
- The starting models estimated by the global search algorithm are far away from the optimum minimum of the misfit function. The chapters 8 and 9 discuss how the nonlinearity of the adjoint problem can be reduced by the application of time windows and frequency filtering of the seismic data.
- Finally the influence of the different starting models on the FWT result are compared in chapter 10.
- A major problem of the 2D FWT implementation is the 2D geometrical spreading, however the seismic wavefield recorded in the field experiences a 3D geometrical spreading. One possible approach to solve this problem and its limitations are discussed in chapter 11.

## CHAPTER 1. INTRODUCTION

---

This thesis covers the whole experience and knowledge I earned during the last 5 years in the topic of FWT. I tried to find a good balance between an accurate mathematical description of the FWT and its numerical application. I hope you enjoy reading this introduction to elastic FWT. It is only a starting point to an unlimited amount of possible applications.

Daniel Köhn

Kiel, 27th May 2010

# 2

## The Seismic Forward Problem

### 2.1 Equations of motion for an elastic medium

---

In this chapter I give a short introduction to the theory of linear elasticity.

#### 2.1.1 Conservation of Momentum

What happens in a rock, during the passage of a seismic wave? To answer this question we investigate an infinitesimal small cube in the rock. This cube contains atoms, ions, molecules or large ensembles of crystals, which show an unknown type of motion. If we could describe the motion of these particles we could calculate the macroscopic particle displacements and particle velocities in the rock and therefore predict the propagation of seismic waves. To achieve this goal, we calculate the conservation of momentum for this cube. How large is the change of the total momentum in the cube

$$\frac{\partial \mathbf{P}}{\partial t} = ?. \quad (2.1)$$

The total momentum is the sum of the momentum of the  $N$  particles in the cube

$$\mathbf{P} = \sum_{i=1}^N m_i \mathbf{v}_i, \quad (2.2)$$

where  $m_i$  and  $\mathbf{v}_i$  denote the mass and velocity of the  $i$ -th particle, respectively. If there are a lot of particles in the cube, we can introduce the mass density  $\rho = \frac{dm}{dV}$  and replace the sum in Eq. (2.2) by an integral. With Eq. (2.1) we get

$$\frac{\partial \mathbf{P}}{\partial t} = \frac{\partial}{\partial t} \left\{ \int_V \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dV \right\} = ?. \quad (2.3)$$

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

Here  $\mathbf{x}$  denotes the position vector and  $t$  the time. In the simplest case there is no change of the momentum at all, i.d.:

$$\frac{\partial \mathbf{P}}{\partial t} = \frac{\partial}{\partial t} \left\{ \int_V \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dV \right\} = 0. \quad (2.4)$$

One possibility to change the total momentum in the cube is by adding or subtracting particles from the cube. By subtracting particles, the total momentum decreases, the addition of particles leads to an increase of the total momentum. By measuring how many particles are streaming through the surface of the cube per time unit, i.e. the momentum flux density  $\phi$ , we can calculate the momentum change in the cube. The total momentum flux  $\Phi$  through the surface of the cube equals the integral of the momentum flux density of the individual surface elements  $dA$  of the cube:

$$\Phi = \int_A \phi dA = \int_A \rho \mathbf{v}(\mathbf{v} \cdot d\mathbf{A}). \quad (2.5)$$

The combination of Eq. 2.5 with 2.4 leads to the following equation for the conservation of momentum:

$$\frac{\partial}{\partial t} \left\{ \int_V \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dV \right\} + \int_A \rho \mathbf{v}(\mathbf{v} \cdot d\mathbf{A}) = 0. \quad (2.6)$$

Finally the total momentum in the cube can be changed by forces. There are two types of forces, surface- and bodyforces. Surface forces are acting through the surface of the cube. An example would be the hydrostatic pressure  $P$  in a liquid. The pressure is a special case of stresses  $\sigma_{ij}$ . By integrating all stresses which are acting on the surface of the cube, we get the change of momentum by surface forces. For the total change of momentum we get

$$\frac{\partial}{\partial t} \left\{ \int_V \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dV \right\} + \int_A \rho \mathbf{v}(\mathbf{v} \cdot d\mathbf{A}) = \int_A \sigma_{ij} n_j \cdot dA. \quad (2.7)$$

Another type of forces are the body forces  $f_i$ , which are acting on all particles in the cube. Therefore the total body forces are estimated by an integration over the volume

$$\frac{\partial}{\partial t} \left\{ \int_V \rho(\mathbf{x}, t) \mathbf{v}(\mathbf{x}, t) dV \right\} + \int_A \rho \mathbf{v}(\mathbf{v} \cdot d\mathbf{A}) = \int_A \sigma_{ij} n_j \cdot dA + \int_V f_i dV. \quad (2.8)$$

By evaluating the LHS of Eq. (2.7), we get

$$\begin{aligned} \int_A \rho v_i(\mathbf{v} \cdot d\mathbf{A}) &= \int_V \nabla \cdot (\rho v_i \mathbf{v}) dV \\ &= \int_V \rho v_i \nabla \cdot \mathbf{v} dV + \int_V \mathbf{v} \cdot \nabla (\rho v_i) dV, \end{aligned} \quad (2.9)$$



## 2.1. EQUATIONS OF MOTION FOR AN ELASTIC MEDIUM

---

where Gauss's divergence theorem and the product rule of differential calculus are used. The surface integral on the RHS of Eq. (2.7) can also be transformed to a volume integral by applying Gauss's theorem.

$$\int_A \sigma_{ij} n_j \cdot dA = \int_V \frac{\partial \sigma_{ij}}{\partial x_j} dV. \quad (2.10)$$

Finally we get the conservation of momentum

$$\int_V \frac{\partial}{\partial t} \rho v_i dV + \int_V \rho v_i \nabla \cdot \mathbf{v} dV + \int_V \mathbf{v} \cdot \nabla (\rho v_i) dV = \int_V \frac{\partial \sigma_{ij}}{\partial x_i} dV + \int_V f_i dV \quad (2.11)$$

For the case  $\lim_{dV \rightarrow 0}$ , the intermediate value theorem of integral calculus can be applied to Eq. (2.11) which leads to the following system of 1st order linear differential equations:

$$\frac{\partial}{\partial t} (\rho v_i) + \rho v_i \nabla \cdot \mathbf{v} + \mathbf{v} \cdot \nabla (\rho v_i) = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \quad (2.12)$$

Eq. (2.12) can be simplified furthermore. After differentiating the terms on the LHS of Eq. (2.12), we get

$$v_i \frac{\partial \rho}{\partial t} + \rho \frac{\partial v_i}{\partial t} + \rho v_i \nabla \cdot \mathbf{v} + \mathbf{v} \rho \nabla (v_i) + \mathbf{v} v_i \nabla \rho = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \quad (2.13)$$

Collecting the terms with  $v_i$  leads to:

$$v_i \left\{ \frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} \right\} + \rho \frac{\partial v_i}{\partial t} + \mathbf{v} \rho \nabla v_i = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \quad (2.14)$$

Here the operator  $\frac{Dy}{Dt}$  denotes the substantial derivation of the variable  $y$ :

$$\frac{Dy}{Dt} = \frac{\partial y}{\partial t} + \mathbf{v} \nabla \cdot y \quad (2.15)$$

The variable  $y$  can be a scalar or a vector, like the density  $\rho$  or the velocity vector  $\mathbf{v}$ . The particles in the volume not only have a specific momentum, they also have a mass. Therefore the conservation of mass also has to be fulfilled. In analogy to the derivation of the momentum conservation we get the mass conservation in differential form as

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0. \quad (2.16)$$

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

The curly bracket on the LHS of Eq. (2.14) vanishes due to the conservation of mass. This finally leads to the following system of first order partial differential equations

$$\rho \frac{\partial v_i}{\partial t} + \mathbf{v} \rho \nabla v_i = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i. \quad (2.17)$$

Equation (2.17) denotes, that the momentum in the volume can be changed in four different ways:

1. We stand at a specific point in the volume and observe the temporal momentum change (1st term on the LHS of Eq. (2.17)).
2. We move through the volume and observe a spatial momentum variation (2nd term on the LHS of Eq. (2.17)).
3. We apply surface forces on the cube (1st term on the RHS of Eq. (2.17)).
4. We apply body forces on the cube (2nd term on the RHS of Eq. (2.17)).

These statements are true for every volume, which can be filled with water, gas or a rock. Until this point there is no description on how the medium is behaving when applying a certain amount of stress. In the following section, this relationship between stress and strain in a rock will be deduced. The derivation in sections 2.1.2 - 2.1.4 is based on Müller (2007).

### 2.1.2 Infinitesimal Deformation in the rock

#### Decomposition of the displacement vector

We investigate a body, which is deformed by external forces. In the undeformed state the point P has the position vector  $x_i$ , the point Q in the infinitesimal neighborhood the position vector  $x_i + y_i$ . After the deformation point P is moved to point P' by the displacement vector  $u_i$ , point Q is displaced to Q' by the vector

$$z_i = u_i + du_i = u_i + \frac{\partial u_i}{\partial x_j} y_j \quad (\text{Einstein notation (EN) !}).$$

The vector  $z_i$  describes (for moving Q in the neighborhood of P) the change, that the neighborhood of P experiences by the deformation. This change can be split into

## 2.1. EQUATIONS OF MOTION FOR AN ELASTIC MEDIUM

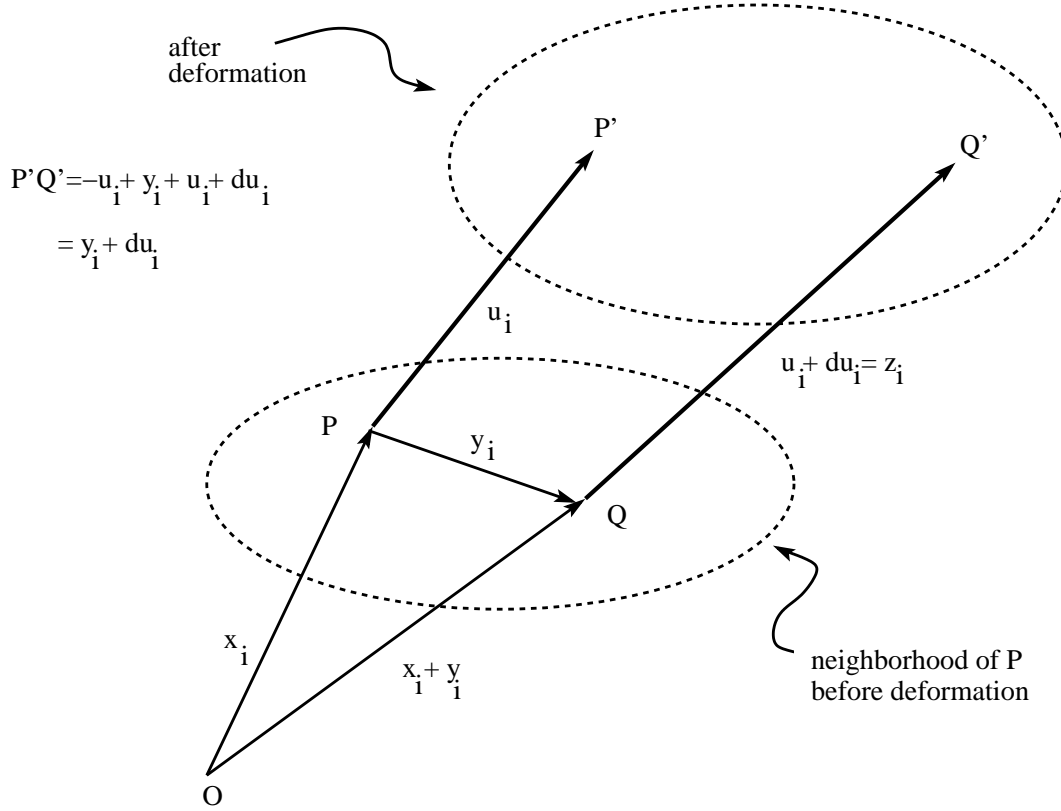


Figure 2.1: Definition of vector  $z_i$  (Müller (2007)).

two parts, a translation, a rotation around an axis through P and the deformation, where distances change their length:

$$z_i = u_i + du_i = \underbrace{u_i}_{\text{translation}} + \underbrace{\epsilon_{ij}y_j}_{\text{deformation}} + \underbrace{\zeta_{ij}y_j}_{\text{rotation}}$$

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad \zeta_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) \quad (2.18)$$

$$\epsilon_{ij} = \epsilon_{ji} \quad (2.19)$$

$$\zeta_{ij} = -\zeta_{ji} (\rightarrow \zeta_{11} = \zeta_{22} = \zeta_{33} = 0) \quad (2.20)$$

The matrices  $\epsilon_{ij}$  and  $\zeta_{ij}$  are 2nd order tensors.  $\epsilon_{ij}$  is symmetrical (2.19),  $\zeta_{ij}$  is anti-symmetrical (2.20).  $\epsilon_{ij}$  is called **deformation tensor**,  $\zeta_{ij}$  **rotation tensor**.

### 2.1.3 Analysis of the stress state

#### Stress

In the deformed state, two types of forces are acting on the body. **Body forces** (or volume forces) are acting on each particle in a volume element. Examples are gravity or inertia forces. **Surface forces** are acting through the surface of adjacent volume elements. In the following body forces are neglected. We investigate a deformed body which consists of two parts, part  $K_1$  with the surface  $S$ , which is surrounded by part  $K_2$  (Fig. 2.2). If  $K_2$  is removed,  $K_1$  will adapt a new equilibrium form. Therefore  $K_2$  leads to forces acting through  $S$  on  $K_1$ . To deform  $K_1$  back into the old state, forces  $\Sigma\Delta f$  have to act on each surface element  $\Delta f$  on  $S$ . The same forces are performed by  $K_2$ .  $\Sigma$  with dimensions force/surface is called **traction**. Its direction and magnitude depends on

1. the location of the surface element  $\Delta f$ ,
2. its normal direction  $\mathbf{n}$  (defined as the direction pointing outside of  $K_1$ ).

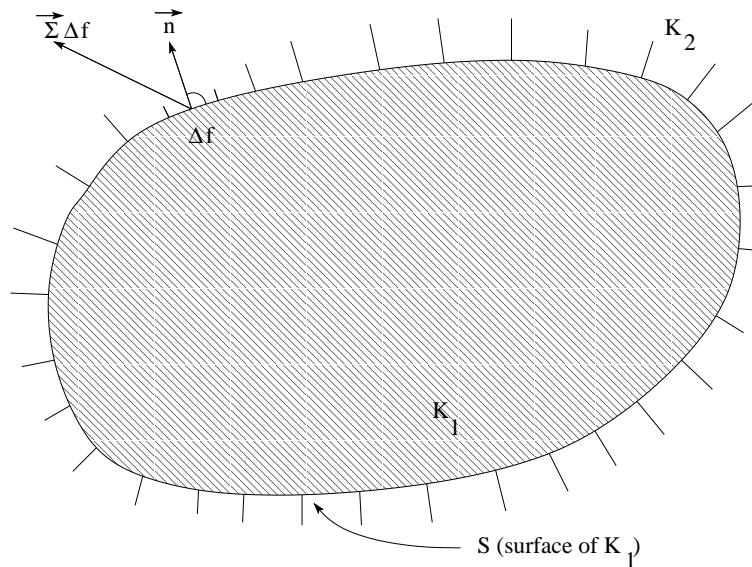


Figure 2.2: Definition of the traction vector  $\Sigma$  (Müller (2007)).

Components of  $\Sigma$  parallel to  $\mathbf{n}$ : normal stress

Components of  $\Sigma$  perpendicular to  $\mathbf{n}$ : shear stress

## 2.1. EQUATIONS OF MOTION FOR AN ELASTIC MEDIUM

---

If  $\Sigma$  is known in the whole body and for all directions  $\mathbf{n}$ , the stress state of the body is known. This can be described by 6 functions.

### Stress tensor $\sigma_{ij}$

We investigate an infinitesimal tetrahedron ABCD in the body (Fig. 2.3) and assume, that the traction on the 3 surfaces ABD, ABC and ACD are known. We want to calculate the traction  $\Sigma$  acting on the surface BCD.

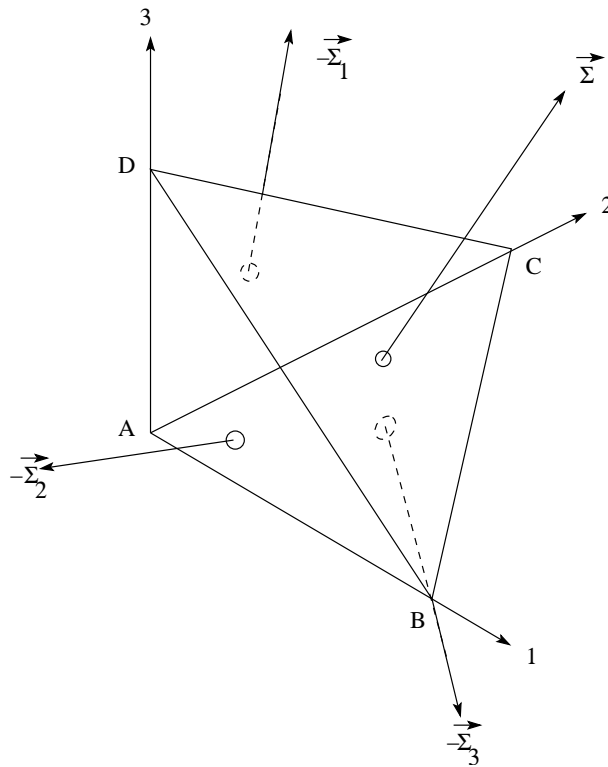


Figure 2.3: Calculation of the traction  $\Sigma$  (Müller (2007)).

Since the tetrahedron is assumed to be infinitesimal small, all traction vectors on the surface elements can be assumed to be constant. Normal directions and surfaces:

ABD : negative 2-direction,  $\Delta f_2$

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

ABC : negative 3-direction,  $\Delta f_3$   
 ACD : negative 1-direction,  $\Delta f_1$   
 BCD :  $\mathbf{n} = (n_1, n_2, n_3)$ ,  $\Delta f$

$$\Delta f_j = \Delta f n_j \quad (2.21)$$

We assume that forces and traction vectors acting on ABD, ABC and ACD for the positive 2- resp. 3- resp. 1-direction are known:

ABD :  $\Sigma_2 \Delta f_2$ ,  $\Sigma_2 = (\sigma_{21}, \sigma_{22}, \sigma_{23})$   
 ABC :  $\Sigma_3 \Delta f_3$ ,  $\Sigma_3 = (\sigma_{31}, \sigma_{32}, \sigma_{33})$   
 ACD :  $\Sigma_1 \Delta f_1$ ,  $\Sigma_1 = (\sigma_{11}, \sigma_{12}, \sigma_{13})$

This means that the 9 functions  $\sigma_{ij}$  are known. If body forces are neglected (they converge in the case of a volume reduction of the tetrahedron faster to zero than the surface forces), the force balance on the tetrahedron leads to:

$$-\Sigma_j \Delta f_j + \Sigma \Delta f = 0$$

With Eq.(2.21) we get:

$$\Sigma = \Sigma_j n_j \quad (2.22)$$

This result shows that the traction of an arbitrary orientated surface element can be calculated if the traction on 3 perpendicular surface elements is known. Eq. (2.22) can be written as:

$$\Sigma_1 = \sigma_{11}n_1 + \sigma_{21}n_2 + \sigma_{31}n_3$$

General:  $\Sigma_j = \sigma_{ij}n_i$  (EN!)

The 9 functions  $\sigma_{ij}$  can be described by the **stress tensor**. The components  $\sigma_{i1}$ ,  $\sigma_{i2}$ ,  $\sigma_{i3}$  form the traction vector for one surface element, whose normal vector is pointing in the direction of the (positive) i-axis.  $\sigma_{ii}$  (i.e.  $\sigma_{11}$ ,  $\sigma_{22}$  or  $\sigma_{33}$ ) is the normal stress, the two other components are shear stresses. The stress tensor ist symmetrical:

$$\sigma_{ij} = \sigma_{ji}.$$

Which also implies that  $\Sigma_j = \sigma_{ji}n_i$  equals

$$\Sigma_i = \sigma_{ij}n_j. \quad (2.23)$$

The stress tensor has 6 independent components.

---

## 2.1. EQUATIONS OF MOTION FOR AN ELASTIC MEDIUM

### 2.1.4 Linear stress-strain relationship

#### Generalized Hookes law

To describe how a body deforms when a certain amount of stress is applied, a relationship between stress and strain is needed. This implies that at every point in the body a relationship between the stress and the strain tensor exists:

$$\sigma_{ij} = f_{ij}(\epsilon_{11}, \epsilon_{12}, \dots, \epsilon_{33}) \quad (2.24)$$

In seismology the simplest form of (2.24) is sufficient. In this case  $\sigma_{ij}$  at a given point only depends on the actual  $\epsilon_{kl}$  at this point. In this case the deformation vanishes instantly if the stress stops  $\epsilon_{kl} = 0$ ,  $\sigma_{ij} = 0$ . This can be written as:

$$\begin{aligned} \sigma_{ij} &= f_{ij}(\epsilon_{11}, \epsilon_{12}, \dots, \epsilon_{33}) \\ f_{ij}(0, 0, \dots, 0) &= 0 \end{aligned} \quad (2.25)$$

If such a relationship is valid, it's called **ideal elasticity**. For infinitesimal deformations this means that  $\sigma_{ij}$  is a linear function of all  $\epsilon_{kl}$ :

$$\begin{aligned} \sigma_{ij} &= c_{ijkl}\epsilon_{kl} \quad (\text{Einstein notation twice !}) \\ c_{ijkl} &= \text{stiffness tensor} \end{aligned} \quad (2.26)$$

The **linear theory of elasticity** investigates elastic behaviour in bodys under the following circumstances:

1. The deformations are infinitesimal
2. The stress-strain relationships are linear.

The famous **Hookes law**, which describes the extension of springs, is a special case of Eq. (2.26). Therefore (2.26) is also called generalized Hookes law. As an example the relationship shown in Fig. 2.4 represents the stretching of a wire.

Between A and B the relationship between traction per surface unit of the cross-section  $\sigma_{11}$  and the relative change of the length  $\epsilon_{11}$  is linear and therefore is equal to Eq. (2.26) ( $E = \text{Young's modulus}$ ). Between B and C the relationship is not linear anymore, but is still predictable by ideal elasticity, i.e. if  $\sigma_{11}$  equals zero, the deformation  $\epsilon_{11}$  is also zero. Beyond point C an irreversible deformation (plastic flow) can be observed. Finally the wire rips apart. The 4th order tensor  $c_{ijkl}$  has 81

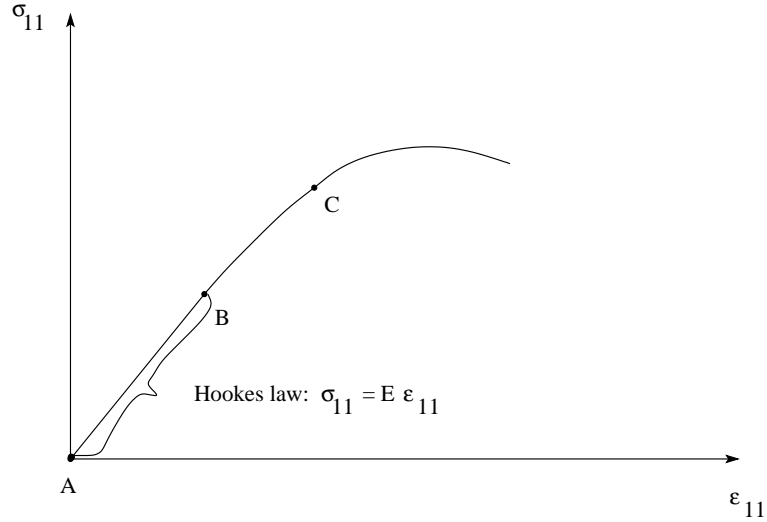


Figure 2.4: Scope of application for Hookes law.

(= 9 x 9) components. Due to the symmetry of the stress- and strain tensors only 36 (= 6 x 6) components are independent. Since the elastic deformation energy (= elastic energy per volume unit) is a state function, the number can be reduced to 21 (Aki and Richards (1980), p.21-23). This is the maximum number of independent elasticity constants, which an anisotropic body can have. For special forms of anisotropy and especially in the case of isotropy this number can be further reduced. In an isotropic body, which means uniformity in all directions, only 2 elastic constants are required and the stiffness tensor  $c_{ijkl}$  has the form (Aki and Richards (1980), p.23):

$$c_{ijkl} = \delta_{ij}\delta_{kl}\lambda + (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\mu. \quad (2.27)$$

In this case the **stress-strain relationship** (2.26) can be written as:

$$\sigma_{ij} = \lambda\theta\delta_{ij} + 2\mu\epsilon_{ij}, \quad (2.28)$$

where

$\lambda$  and  $\mu$  : Lamé parameters

$\theta = \epsilon_{11} + \epsilon_{22} + \epsilon_{33}$  : cubic dilatation

$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$  : Kronecker's delta



---

## 2.1. EQUATIONS OF MOTION FOR AN ELASTIC MEDIUM

### 2.1.5 The elastic wave equations

In the theory of infinitesimal deformations it can be assumed that all particle velocities are small. Therefore the nonlinear advection term on the LHS of (2.17) vanishes. Together with the linear stress-strain relationship (2.28) and the definition of the strain tensor (2.18) we get the following system of coupled, linear partial differential equations to describe the propagation of waves in an elastic medium:

$$\begin{aligned}
 \rho \frac{\partial v_i}{\partial t} &= \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \\
 \sigma_{ij} &= \lambda \theta \delta_{ij} + 2\mu \epsilon_{ij} \\
 \epsilon_{ij} &= \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)
 \end{aligned} \tag{2.29}$$

(2.29) is a system of first order partial differential equations. Using  $v_i = \frac{\partial u_i}{\partial t}$ , (2.29) can be transformed into a system of second order partial differential equations:

$$\begin{aligned}
 \rho \frac{\partial^2 u_i}{\partial t^2} &= \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \\
 \sigma_{ij} &= \lambda \theta \delta_{ij} + 2\mu \epsilon_{ij} \\
 \epsilon_{ij} &= \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)
 \end{aligned} \tag{2.30}$$

This expression is called **Stress-Displacement** formulation. Another common form of the elastic equations of motion can be deduced by taking the time derivative of the stress-strain relationship and the strain tensor in Eq. (2.29). Since the Lamé parameters  $\lambda$  and  $\mu$  do not depend on time, Eq. (2.29) can be written as:

$$\begin{aligned}
 \rho \frac{\partial v_i}{\partial t} &= \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \\
 \frac{\partial \sigma_{ij}}{\partial t} &= \lambda \frac{\partial \theta}{\partial t} \delta_{ij} + 2\mu \frac{\partial \epsilon_{ij}}{\partial t} \\
 \frac{\partial \epsilon_{ij}}{\partial t} &= \frac{1}{2} \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)
 \end{aligned} \tag{2.31}$$

This expression is called **Stress-Velocity** formulation. For simple cases (2.30) and (2.31) can be solved analytically. More complex problems require numerical solutions. One possible approach for a numerical solution is described in the next section.

## 2.2 Solution of the elastic wave equation by finite differences

---

### 2.2.1 Discretization of the wave equation

For the numerical solution of the elastic equations of motion, Eqs. (2.30) have to be discretized in time and space on a grid. The particle displacement  $\mathbf{u}$ , the stresses  $\sigma_{ij}$ , the Lamé parameters  $\lambda$  and  $\mu$  are calculated and defined at discrete Cartesian coordinates  $x = i \, dh$ ,  $y = j \, dh$  and discrete times  $t = n \, dt$ .  $dh$  denotes the spatial distance between two adjacent grid points and  $dt$  the difference between two successive time steps. Therefore every grid point is located in the interval  $i \in \mathbb{N}[[1, NX]$ ,  $j \in \mathbb{N}[[1, NY]$  and  $n \in \mathbb{N}[[1, NT]$ , where  $NX$ ,  $NY$  and  $NT$  are the number of discrete spatial grid points and time steps, respectively. Finally the partial derivatives are replaced by **finite-difference (FD)** operators. Two types of operators can be distinguished, forward and backward operators  $D^+$ ,  $D^-$ . The derivative of a function  $y$  with respect to a variable  $x$  can be approximated by the following operators:

$$\begin{aligned} D_x^+ y &= \frac{y[i+1] - y[i]}{dh} && \text{forward operator} \\ D_x^- y &= \frac{y[i] - y[i-1]}{dh} && \text{backward operator} \end{aligned} \tag{2.32}$$

To calculate the spatial derivatives of the wavefield variables at the correct positions, the variables are not placed on the same grid points, but staggered by half of the spatial grid point distance (Virieux (1986) and Levander (1988)). Fig. 2.5 shows the distribution of the material parameters and wavefield variables on the spatial grid. To guarantee the stability of the **standard staggered grid (SSG)** code, the Lamé parameter  $\mu$  and density  $\rho$  have to be averaged harmonically and arithmetically (Moczó et al. (2004), Bohlen and Saenger (2006)), respectively

$$\begin{aligned} \langle \mu \rangle [j + \frac{1}{2}][i + \frac{1}{2}] &= \left[ \frac{1}{4} \left( \mu^{-1}[j][i] + \mu^{-1}[j][i+1] + \mu^{-1}[j+1][i+1] + \mu^{-1}[j+1][i] \right) \right]^{-1} \\ \rho_x[j][i + \frac{1}{2}] &= \frac{1}{2} (\rho[j][i+1] + \rho[j][i]) \\ \rho_y[j + \frac{1}{2}][i] &= \frac{1}{2} (\rho[j+1][i] + \rho[j][i]) \end{aligned} \tag{2.33}$$

## 2.2. SOLUTION OF THE ELASTIC WAVE EQUATION BY FINITE DIFFERENCES

---

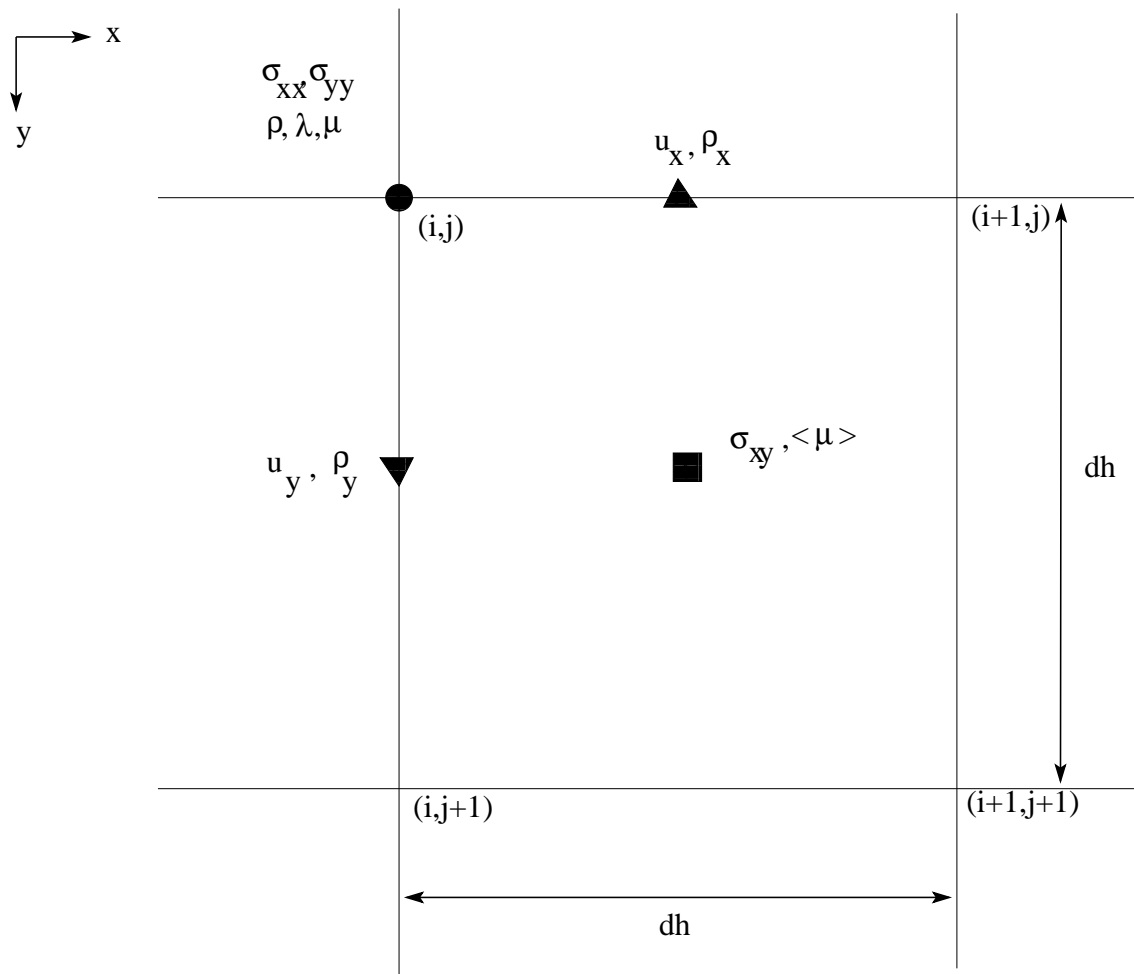


Figure 2.5: Grid geometry for a standard staggered grid (SSG) in Cartesian coordinates as suggested by Virieux (1986) and Levander (1988).

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

The discretization of the linear stress strain relationship in (2.30) at time step  $n$  leads to the following system of equations (for simplicity I skip the time index  $n$ ):

$$\begin{aligned}
 u_{xx}[j][i] &\approx \frac{u_x[j][i + \frac{1}{2}] - u_x[j][i - \frac{1}{2}]}{dh} \\
 u_{yy}[j][i] &\approx \frac{u_y[j + \frac{1}{2}][i] - u_y[j - \frac{1}{2}][i]}{dh} \\
 u_{yx}[j + \frac{1}{2}][i + \frac{1}{2}] &\approx \frac{u_y[j + \frac{1}{2}][i + 1] - u_y[j + \frac{1}{2}][i]}{dh} \\
 u_{xy}[j + \frac{1}{2}][i + \frac{1}{2}] &\approx \frac{u_x[j + 1][i + \frac{1}{2}] - u_x[j][i + \frac{1}{2}]}{dh} \\
 \sigma_{xy}[j + \frac{1}{2}][i + \frac{1}{2}] &= \langle \mu \rangle [j + \frac{1}{2}][i + \frac{1}{2}] \left( u_{xy}[j + \frac{1}{2}][i + \frac{1}{2}] + u_{yx}[j + \frac{1}{2}][i + \frac{1}{2}] \right) \\
 \sigma_{xx}[j][i] &= \lambda[j][i] * \left( u_{xx}[j][i] + u_{yy}[j][i] \right) + 2 * \langle \mu \rangle [j][i] * u_{xx}[j][i] \\
 \sigma_{yy}[j][i] &= \lambda[j][i] * \left( u_{xx}[j][i] + u_{yy}[j][i] \right) + 2 * \langle \mu \rangle [j][i] * u_{yy}[j][i]
 \end{aligned} \tag{2.34}$$

The discretization of the momentum equation in (2.30) leads to the following system of equations:

$$\begin{aligned}
 utt_x^n[j][i + \frac{1}{2}] &= \left( \sigma_{xx}[j][i + 1] - \sigma_{xx}[j][i] + \sigma_{xy}[j + \frac{1}{2}][i] - \sigma_{xy}[j - \frac{1}{2}][i] \right) \\
 utt_y^n[j + \frac{1}{2}][i] &= \left( \sigma_{xy}[j][i + \frac{1}{2}] - \sigma_{xy}[j][i - \frac{1}{2}] + \sigma_{yy}[j + 1][i] - \sigma_{yy}[j][i] \right) \\
 u_x^{n+1}[j][i + \frac{1}{2}] &= 2 * u_x^n[j][i + \frac{1}{2}] - u_x^{n-1}[j][i + \frac{1}{2}] + \frac{dt^2}{dh * \rho_x[j][i + \frac{1}{2}]} * utt_x^n[j][i + \frac{1}{2}] \\
 u_y^{n+1}[j + \frac{1}{2}][i] &= 2 * u_y^n[j + \frac{1}{2}][i] - u_y^{n-1}[j + \frac{1}{2}][i] + \frac{dt^2}{dh * \rho_y[j + \frac{1}{2}][i]} * utt_y^n[j + \frac{1}{2}][i]
 \end{aligned} \tag{2.35}$$

## 2.2. SOLUTION OF THE ELASTIC WAVE EQUATION BY FINITE DIFFERENCES

---

### 2.2.2 Accuracy of FD operators

The derivation of the FD operators in the last section was a simple replacement of the partial derivatives by finite differences. In the following more systematic approach, the first derivative of a variable  $f$  at a grid point  $i$  is estimated by a Taylor series expansion (Jastram (1992)):

$$(2k-1) \frac{\partial f}{\partial x} \Big|_i = \frac{1}{dh} (f_{i+(k-1/2)} - f_{i-(k-1/2)}) + \frac{1}{dh} \sum_{l=2}^N \frac{((k-\frac{1}{2})dh)^{2l-1}}{(2l-1)!} \frac{\partial^{(2l-1)} f}{\partial x^{(2l-1)}} \Big|_i + \mathcal{O}(dh)^{2N}$$

For an operator with length  $2N$ ,  $N$  equations are added with a weight  $\beta_k$ :

$$\left[ \sum_{k=1}^N \beta_k (2k-1) \right] \frac{\partial f}{\partial x} \Big|_i = \frac{1}{dh} \sum_{k=1}^N \beta_k (f_{i+(k-1/2)} - f_{i-(k-1/2)}) + \frac{1}{dh} \sum_{k=1}^N \sum_{l=2}^N \beta_k \frac{((k-\frac{1}{2})dh)^{2l-1}}{(2l-1)!} \frac{\partial^{(2l-1)} f}{\partial x^{(2l-1)}} \Big|_i + \mathcal{O}(dh)^{2N} \quad (2.36)$$

The case  $N=1$  leads to the FD operator derived in the last section, which has a length of  $2N=2$ . The Taylor series is truncated after the first term ( $\mathcal{O}(dh)^2$ ). Therefore this operator is called **2nd order FD operator** which refers to the truncation error of the Taylor series and not to the order of the approximated derivative. To understand equation (2.36) better, we estimate a **4th order FD operator**. This operator has the length  $2N = 4$  or  $N=2$ . The sums in Eq. (2.36) lead to:

$$(\beta_1 + 3\beta_2) \frac{\partial f}{\partial x} \Big|_i = \frac{1}{dh} (\beta_1 (f_{i+1/2} - f_{i-1/2}) + \beta_2 (f_{i+3/2} - f_{i-3/2})) + \frac{dh^3}{dh} \left[ \beta_1 \frac{1}{8 \cdot 3!} + \beta_2 \frac{27}{8 \cdot 3!} \right] \frac{\partial^3 f}{\partial x^3} \Big|_i \quad (2.37)$$

The weights  $\beta_k$  can be calculated by the following approach: The factor in front of the partial derivative on the LHS of Eq. (2.37) should equal 1, therefore

$$(\beta_1 + 3\beta_2) = 1.$$

The coefficients in front of  $\frac{\partial^3 f}{\partial x^3} \Big|_i$  on the RHS of Eq. (2.37) should vanish:

$$(\beta_1 + 27\beta_2) = 0.$$

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

The weights  $\beta_k$  can be estimated by solving the matrix equation:

$$\begin{pmatrix} 1 & 3 \\ 1 & 27 \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The resulting coefficients are  $\beta_1 = 9/8$  and  $\beta_2 = -1/24$ . Therefore the 4th order backward- and forward operators are:

$$\begin{aligned} \left. \frac{\partial f}{\partial x} \right|_{i+1/2} &= \frac{1}{dh} [\beta_1(f_{i+1} - f_i) + \beta_2(f_{i+2} - f_{i-1})] && \text{forward operator} \\ \left. \frac{\partial f}{\partial x} \right|_{i-1/2} &= \frac{1}{dh} [\beta_1(f_i - f_{i-1}) + \beta_2(f_{i+1} - f_{i-2})] && \text{backward operator} \end{aligned} \quad (2.38)$$

The coefficients  $\beta_i$  in the FD operator are called **Taylor coefficients**. The accuracy of higher order FD operators can be improved by seeking for FD coefficients  $\beta_k$  that approximate the first derivative in a certain frequency range (Holberg (1987)). These numerically optimized coefficients are called **Holberg coefficients**.

### 2.2.3 Initial and Boundary Conditions

To find a unique solution of the problem, initial and boundary conditions have to be defined. The initial conditions for the elastic forward problem are:

$$\begin{aligned} u_i(\mathbf{x}, t) &= 0 \\ \frac{\partial u_i(\mathbf{x}, t)}{\partial t} &= 0 \end{aligned} \quad (2.39)$$

for all  $\mathbf{x} \in V$  at  $t = 0$ .

For the geophysical application two types of boundary conditions are very important:

1. **Horizontal Free Surface:** The interface between the elastic medium and air at the surface is very important when trying to model surface waves or multiple reflections in a marine environment. Since all stresses in the normal direction at this interface vanish

$$\sigma_{xy} = \sigma_{yy} = 0.0 \quad (2.40)$$

this boundary condition is called (stress) **free surface**. Two types of implementations are common. In the implicit definition of the free surface, a

## 2.2. SOLUTION OF THE ELASTIC WAVE EQUATION BY FINITE DIFFERENCES

---

small layer with the acoustic parameters of air ( $V_p = 300$  m/s,  $V_s = 0.0$  m/s,  $\rho = 1.25$  kg/m<sup>3</sup>) is placed on top of the model. One advantage of the implicit definition of the free surface is the easy implementation of topography on the FD grid, however to get accurate results for surface waves or multiples, this approach requires a fine spatial sampling of the FD grid near the free surface. An explicit free surface can be implemented by using the mirroring technique by Levander, which leads to stable and accurate solutions for plain interfaces (Levander (1988), Robertsson et al. (1995)). If the planar free surface is located at grid point  $j = h$ , the stress at this point is set to zero and the stresses below the free surface are mirrored with an inverse sign:

$$\begin{aligned}
 \sigma_{yy}(h, i) &= 0 \\
 \sigma_{yy}(h - 1, i) &= -\sigma_{yy}(h + 1, i) \\
 \sigma_{xy}(h - \frac{1}{2}, i + \frac{1}{2}) &= -\sigma_{xy}(h + \frac{1}{2}, i + \frac{1}{2}) \\
 \sigma_{xy}(h - \frac{3}{2}, i + \frac{1}{2}) &= -\sigma_{xy}(h + \frac{3}{2}, i + \frac{1}{2})
 \end{aligned} \tag{2.41}$$

When updating the stress component  $\sigma_{xx} = (\lambda + 2\mu)u_{xx} + \lambda u_{yy}$  at the free surface, only horizontal particle displacements should be used because vertical derivatives over the free surface lead to instabilities (Levander (1988)). The vertical derivative of the y-displacement  $u_{yy}$  can be replaced by using the boundary condition at the free surface:

$$\begin{aligned}
 \sigma_{yy} &= (\lambda + 2\mu)u_{yy} + \lambda u_{xx} = 0 \\
 u_{yy} &= -\frac{\lambda}{(\lambda + 2\mu)}u_{xx}
 \end{aligned} \tag{2.42}$$

Therefore the stress  $\sigma_{xx}$  can be written as

$$\sigma_{xx} = \frac{4(\lambda\mu + \mu^2)}{\lambda + 2\mu}u_{xx} \tag{2.43}$$

2. Absorbing Boundary Conditions: Due to limited computational resources, the FD grid has to be as small as possible. To model problems with an infinite extension in different directions, e.g. a full or half-space problem, an artificial absorbing boundary condition has to be applied. A commonly used implementation is described in Cerjan et al. (1985): The numerical grid is enlarged by a few grid points "FW" (typically FW = 30 gridpoints) in each direction. The

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

values of the stress and particle velocity in this boundary frame are multiplied by a factor "damp":

$$\text{damp} = \exp(-a^2x^2) \quad (2.44)$$

with  $a = \sqrt{-\log(\text{amp})}/(\text{FW})$  and  $\text{amp} = 0.92$ . The seismic waves are damped inside the boundary frame and cannot be reflected back into the model. This type of absorbing boundary is not able to damp the reflected waves from the boundary completely.

A more effective way to damp the waves near the boundaries are **Perfectly Matched Layers (PMLs)**. This can be achieved by a coordinate stretch of the wave equations in the frequency domain (Komatitsch and Martin (2007)). The coordinate stretch creates exponentially decaying plane wave solutions in the absorbing boundary frame. The PML's are only reflectionless if the exact wave equation is solved. As soon as the problem is discretized (for example using finite differences) you are solving an approximate wave equation and the analytical perfection of the PML is no longer valid. To overcome this shortcoming the wavefield is damped by the damping function

$$c = -V_{\text{pml}} * \frac{\log(\alpha)}{L} \quad (2.45)$$

where  $V_{\text{pml}}$  denotes the typical P-wave velocity of the medium in the absorbing boundary frame,  $\alpha = 1 \times 10^{-4}$  and  $L$  is the thickness of the absorbing boundary layer. A comparison between the exponential damping and the PML boundary is shown in Fig.2.6. The PMLs are damping the seismic waves by a factor 5-10 more effective than the absorbing boundary frame.



## 2.2. SOLUTION OF THE ELASTIC WAVE EQUATION BY FINITE DIFFERENCES

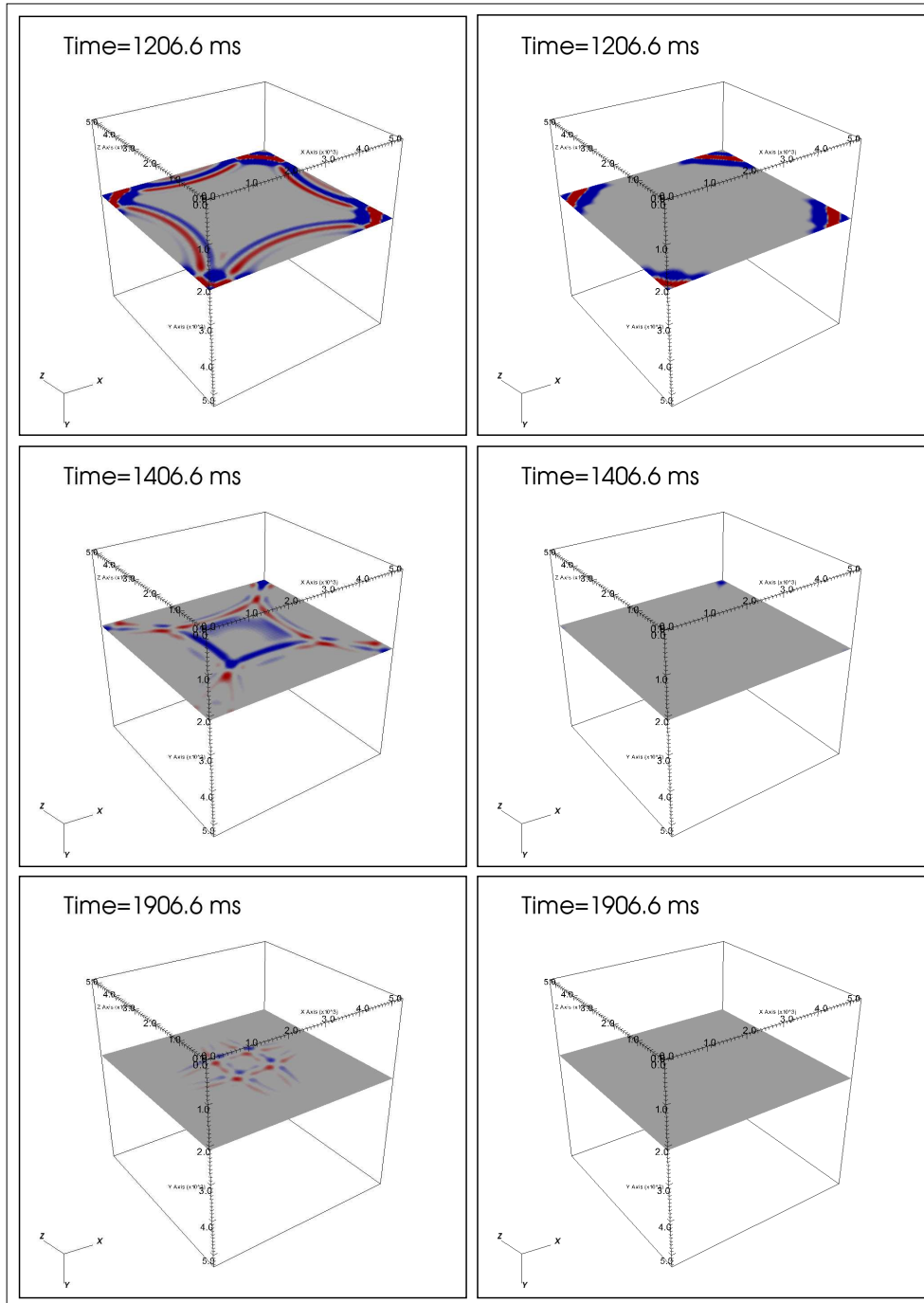


Figure 2.6: Comparison between exponential damping (left column) and PML (right column) absorbing boundary conditions for a homogeneous full space model.

## 2.3 Numerical Artefacts and Instabilities

---

To avoid numerical artefacts and instabilities during a FD modelling run, spatial and temporal sampling conditions for the wavefield have to be satisfied. These will be discussed in the following two sections.

### 2.3.1 Grid Dispersion

The first question when building a FD model is: What is the maximum spatial grid point distance  $dh$ , for a correct sampling of the wavefield? To answer this question we take a look at this simple example: The particle displacement in x-direction is defined by a sine function:

$$u_x = \sin\left(2\pi \frac{x}{\lambda}\right), \quad (2.46)$$

where  $\lambda$  denotes the wavelength. When calculating the derivation of this function analytically at  $x = 0$  and setting  $\lambda = 1$  m we get:

$$\left. \frac{du_x}{dx} \right|_{x=0} = \left. \frac{2\pi}{\lambda} \cos\left(2\pi \frac{x}{\lambda}\right) \right|_{x=0} = 2\pi. \quad (2.47)$$

In the next step the derivation is approximated numerically by a staggered 2nd order finite-difference operator:

$$\left. \frac{du_x}{dx} \right|_{x=0} \approx \left. \frac{u_x(x + \frac{1}{2}\Delta x) - u_x(x - \frac{1}{2}\Delta x)}{\Delta x} \right|_{x=0} = \frac{\sin\left(\frac{2\pi(x + \frac{1}{2}\Delta x)}{\lambda}\right) - \sin\left(\frac{2\pi(x - \frac{1}{2}\Delta x)}{\lambda}\right)}{\Delta x}. \quad (2.48)$$

Using the Nyquist-Shannon sampling theorem it should be sufficient to sample the wavefield with  $\Delta x = \lambda/2$ . In table 2.1 the numerical solutions of eq. (2.48) and the analytical solution (2.47) are compared for different sample intervals  $\Delta x = \lambda/n$ , where  $n$  is the number of gridpoints per wavelength. For the case  $n=2$ , which corresponds to the Nyquist-Shannon theorem, the numerical solution is  $\left. \frac{du_x}{dx} \right|_{x=0} = 4.0$ , which is not equal with the analytical solution  $2\pi$ . A refinement of the spatial sampling of the wavefield results in an improvement of the finite difference solution. For  $n = 16$  the numerical solution is accurate to the second decimal place. The effect of a sparsely sampled pressure field is illustrated in Fig. 2.7 for a homogeneous block model with stress free surfaces. The dimensions of the FD grid are fixed and the central frequency of the source signal is increased systematically. When using a spatial

### 2.3. NUMERICAL ARTEFACTS AND INSTABILITIES

---

sampling of 16 grid points per minimum wavelength (Fig. 2.7, top) the wavefronts are sharply defined. For  $n = 4$  grid points a slight numerical dispersion of the wave occurs (Fig. 2.7, center). This effect is obvious when using the Nyquist criterion ( $n = 2$ ) (Fig. 2.7, bottom). Since the numerical calculated wavefield seem to be dispersive this numerical artefact is called **grid dispersion**. To avoid the occurrence of grid dispersion the following criteria for the spatial grid spacing  $dh$  has to be satisfied:

$$dh \leq \frac{\lambda_{\min}}{n} = \frac{V_{\min}}{n f_{\max}}. \quad (2.49)$$

Here  $\lambda_{\min}$  denotes the minimum wavelength,  $V_{\min}$  the minimum velocity in the model and  $f_{\max}$  is the maximum frequency of the source signal. Depending on the accuracy of the used FD operator the parameter  $n$  is different. In table 2.2  $n$  is listed for different FD operator lengths and types (Taylor and Holberg operators). The Holberg coefficients are calculated for a minimum dispersion error of 0.1% at  $3f_{\max}$ . For short operators  $n$  should be chosen relatively large, so the spatial grid spacing is small, while for longer FD operators  $n$  is smaller and the grid spacing can be larger.

$n$	$\Delta x$ [m]	$\frac{dy_x}{dx} _{x=0}$ []
analytical	-	$2\pi \approx 6.283$
2	$\lambda/2$	4.0
4	$\lambda/4$	5.657
8	$\lambda/8$	6.123
16	$\lambda/16$	6.2429
32	$\lambda/32$	6.2731

Table 2.1: Comparison of the analytical solution Eq. (2.47) with the numerical solution Eq. (2.48) for different grid spacings  $\Delta x = \lambda/n$ .

## CHAPTER 2. THE SEISMIC FORWARD PROBLEM

---

FDORDER	n (Taylor)	n (Holberg)
2nd	12	12
4th	8	8.32
6th	6	4.77
8th	5	3.69
10th	5	3.19
12th	4	2.91

Table 2.2: The number of grid points per minimum wavelength  $n$  for different orders (2nd-12th) and types (Taylor and Holberg) of FD operators. For the Holberg coefficients  $n$  is calculated for a minimum dispersion error of 0.1% at  $3f_{\max}$ .

### 2.3. NUMERICAL ARTEFACTS AND INSTABILITIES

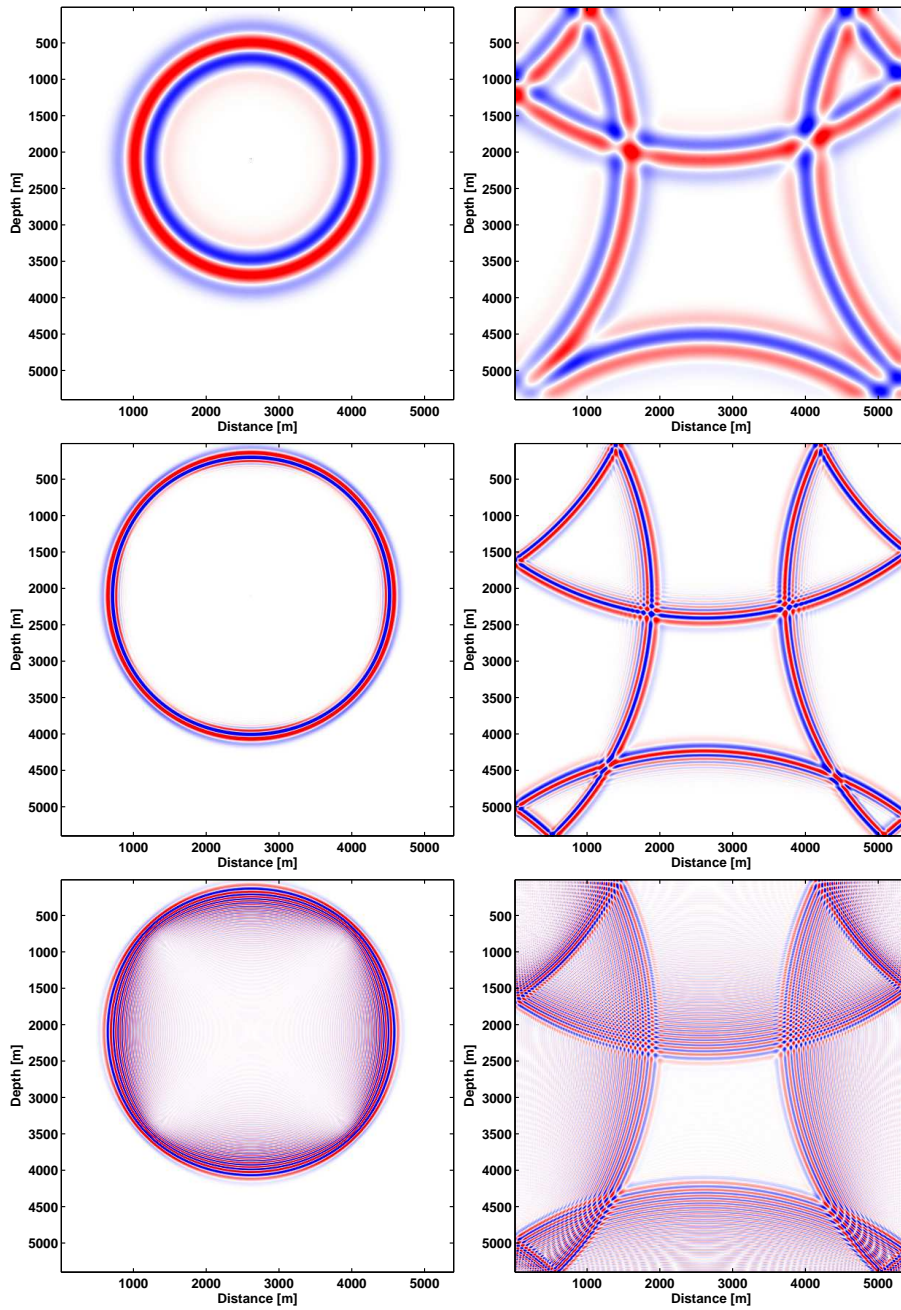


Figure 2.7: The influence of grid dispersion in FD modeling: Spatial sampling of the wavefield using  $n=16$  (top),  $n=4$  (center) and  $n=2$  gridpoints (bottom) per minimum wavelength  $\lambda_{\min}$ .

### 2.3.2 The Courant Instability

Beside the spatial, the temporal discretization has to satisfy a sampling criterion to ensure the stability of the FD code. If a wave is propagating on a discrete grid, then the timestep  $dt$  has to be less than the time for the wave to travel between two adjacent grid points with grid spacing  $dh$ . For an elastic 2D grid this means mathematically:

$$dt \leq \frac{dh}{h\sqrt{2}V_{\max}}, \quad (2.50)$$

where  $V_{\max}$  is the maximum velocity in the model. The factor  $h$  depends on the order of the FD operator and can easily be calculated by summing over the weighting coefficients  $\beta_i$

$$h = \sum_i \beta_i. \quad (2.51)$$

In table 2.3  $h$  is listed for different FD operator lengths and types (Taylor and Holberg operators). Criterion (2.50) is called **Courant-Friedrichs-Lewy criterion** (Courant et al. (1928), Courant et al. (1967)). Fig. 2.8 shows the evolution of the pressure field when the Courant criterion is violated. After a few time steps the amplitudes are growing to infinity and the calculation becomes unstable.

FDORDER	h (Taylor)	h (Holberg)
2nd	1.0	1.0
4th	7/6	1.184614
6th	149/120	1.283482
8th	2161/1680	1.345927
10th	53089/40320	1.387660
12th	1187803/887040	1.417065

Table 2.3: The factor  $h$  in the Courant criterion for different orders (2nd-12th) and types (Taylor and Holberg) of FD operators.

### 2.3. NUMERICAL ARTEFACTS AND INSTABILITIES

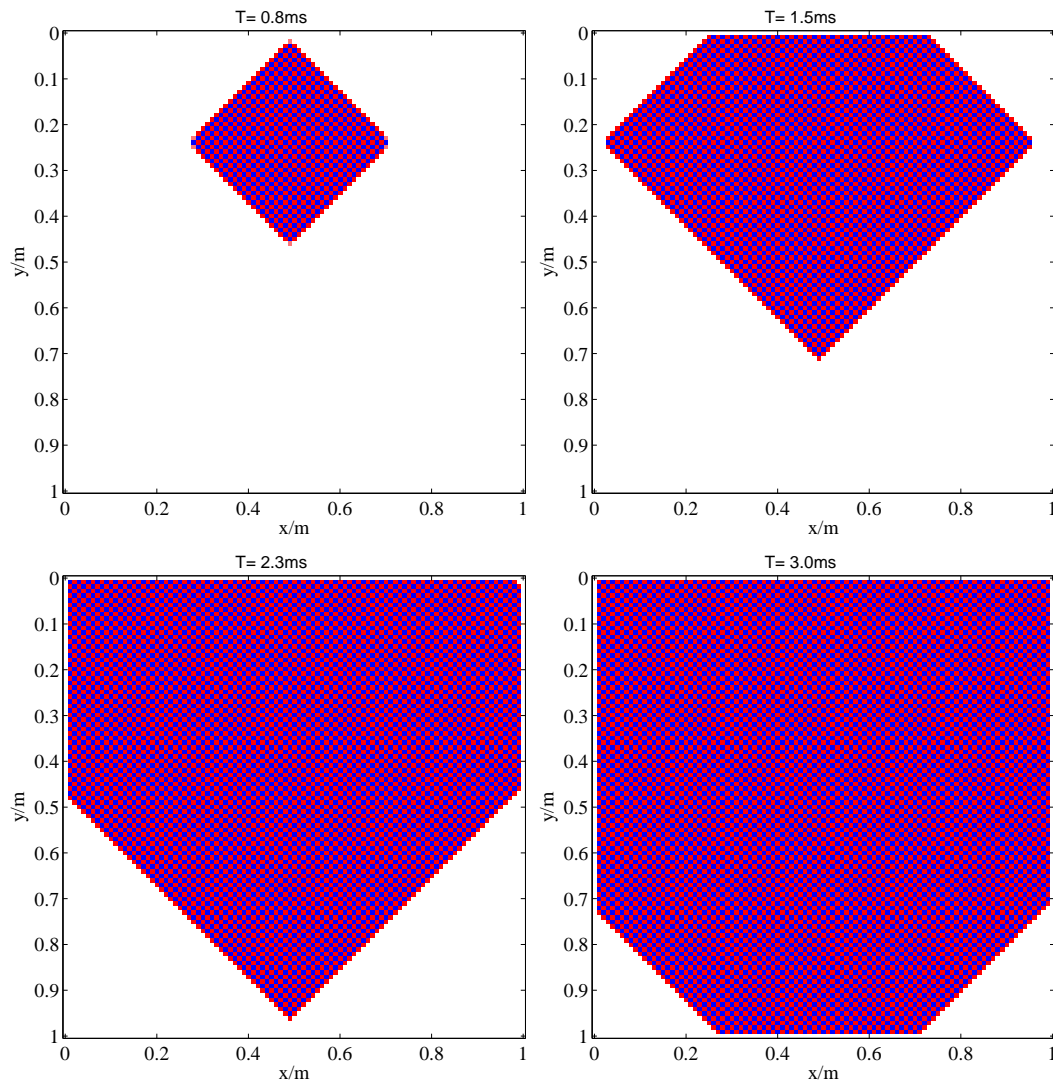


Figure 2.8: Temporal evolution of the Courant instability. In the colored areas the wave amplitudes are extremely large.

### 2.4 Parallelization of the 2D elastic FD Code

---

For small to moderate sized problems Eqs. (2.34) and (2.35) can be solved on modern day CPUs very fast. Larger problems, like sediment models with low S-wave velocities require a dense spatial sampling and therefore large FD grids. This leads to large computation times on single CPUs. As a result the code has to be parallelized. Two strategies are very common. In case of the shot parallelization (see e.g. Shipp and Singh (2002), Fig. 2.9, top) the computation of individual shots is distributed along different CPUs. The advantage of this approach is that no communication between the CPUs is required, so the implementation is fairly simple and straightforward. The second parallelization strategy is based on domain decomposition (see e.g. Bohlen (2002), Fig. 2.9, bottom). The spatial computational domain is subdivided into approximately equally sized parts and distributed along different CPUs. At the boundaries of the domains the wavefield variables, e.g. the displacements  $u_x$  and stresses  $\sigma_{xx}$  have to be exchanged at each time step to model a correct propagation of the seismic waves over the whole FD grid. Therefore the boundary wavefield variables of each CPU are stored in a padding layer and exchanged between the CPUs (denoted by the arrows in Fig. 2.9, bottom). The amount of data exchanged between the CPUs depends on the order of the spatial FD operators. The thickness of the padding layer equals half of the order of the FD operator. Even for larger 2D problems the computation time decreases linear with the number of CPUs until the computational domains are so small, that the communication time exceed the computation time. The forward code in this thesis is parallelized using domain decomposition. The communication between the CPUs is achieved with the Message Passing Interface (MPI). In recent years the application of a few GPUs to solve large numerical problems instead of large CPU Clusters has become a very interesting topic (De Nil (2008), personal communication). GPUs show a much better performance and are cheaper than CPUs. First benchmark tests of elastic wave propagation codes (Komatitsch et al. (2009)) on NVIDIA GPUs in combination with the programming language CUDA show a speedup of a factor 25 compared with standard CPUs. Unfortunately GPUs are not easy to program to work efficiently. The domain decomposition approach can not simply transferred to the GPUs.



## 2.4. PARALLELIZATION OF THE 2D ELASTIC FD CODE

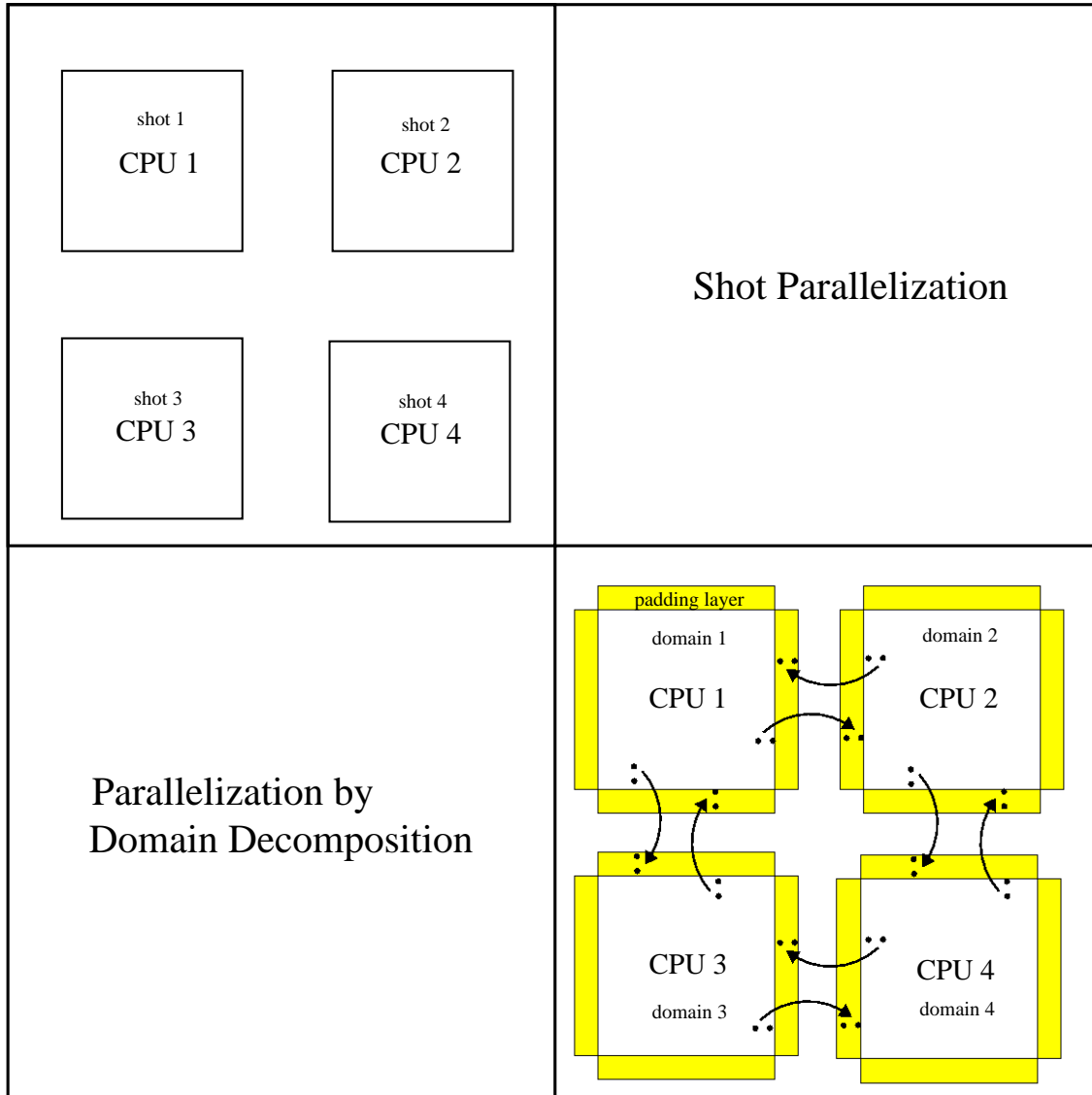


Figure 2.9: Parallelization strategies: Shot parallelization (top) and decomposition of the global grid into subgrids each computed by a different CPU (bottom). Arrows illustrate communication between subgrids at each timestep (modified from Bohlen (2002)).



# 3

## The adjoint problem

The aim of full waveform tomography is to find an "optimum" model which can explain the data very well. It should not only explain the first arrivals of specific phases of the seismic wavefield like refractions or reflections, but also the amplitudes which contain information on the distribution of the elastic material parameters in the underground. To achieve this goal three problems have to be solved:

1. What is an "optimum" model ?
2. How can this model be found ?
3. Is this model unique or are other models existing, which could explain the data equally well ?

Problem 1 and 2 are discussed in this chapter, problem 3 in chapter 10.

### 3.1 What is an "optimum" model ?

---

In reflection seismics the  $i^{\text{th}}$  component of the elastic displacement field  $u_i(\mathbf{x}_s, \mathbf{x}_r, t)$  excited by sources located at  $\mathbf{x}_s$  will be recorded by receivers at  $\mathbf{x}_r$  at time  $t$ . For a given distribution of the material parameters the forward problem Eq. 2.30 can be solved by finite differences (section 2.2). The result is a model data set  $\mathbf{u}^{\text{mod}}$ . This modelled data can be compared with the field data  $\mathbf{u}^{\text{obs}}$ . If the misfit or data residuals  $\delta\mathbf{u} = \mathbf{u}^{\text{mod}} - \mathbf{u}^{\text{obs}}$  (Fig. 3.1) between the modelled and the field data is small the model can explain the data very well. If the residuals are large the model cannot explain the data. The misfit can be measured by a vector norm  $|\mathbf{L}|_p$  which is

## CHAPTER 3. THE ADJOINT PROBLEM

---

defined for  $p = 1, 2, \dots$  as

$$|L|_p = \left( \sum_i |\delta u_i|^p \right)^{1/p} \quad (3.1)$$

The special case  $|L|_\infty$  is defined as

$$|L|_\infty = \max_i |\delta u_i|^p \quad (3.2)$$

The L2-norm

$$E = |L|_2 = \frac{1}{2} \delta \mathbf{u}^T \delta \mathbf{u} \quad (3.3)$$

has a special physical meaning. It represents the residual elastic energy contained in the data residuals  $\delta \mathbf{u}$ . An optimum model can be found in a minimum of the residual energy. Therefore the optimum model is the solution of a nonlinear optimization problem.

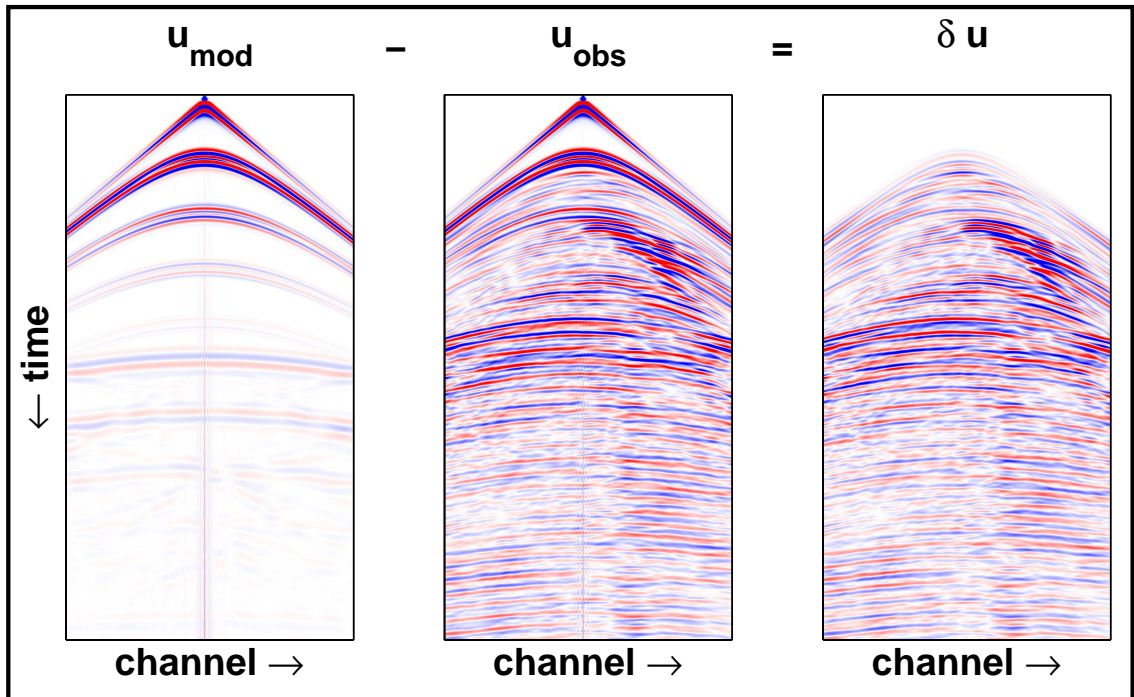


Figure 3.1: Definition of data residuals  $\delta \mathbf{u}$ .

## 3.2 How to find an optimum model

Figure 3.2 shows a schematic sketch of the residual energy at one point in space as a function of two model parameters  $\lambda$  and  $\mu$ . The colors represent different values of the residual energy. Red areas represent models with high residual energy which do not fit the data, while the blue parts are good fitting models with low residual energies. The aim is to find the minimum of the residual energy marked by the red cross. Starting at a point  $\mathbf{m}_1 = (\lambda_1(\mathbf{x}), \mu_1(\mathbf{x}), \rho_1(\mathbf{x}), \dots)$  in the parameter space we want to find the minimum by updating the material parameters in an iterative way

$$\mathbf{m}_2 = \mathbf{m}_1 + \mu_1 \delta \mathbf{m}_1, \quad (3.4)$$

along the search direction  $\delta \mathbf{m}_1$  with the step length  $\mu_1$ .

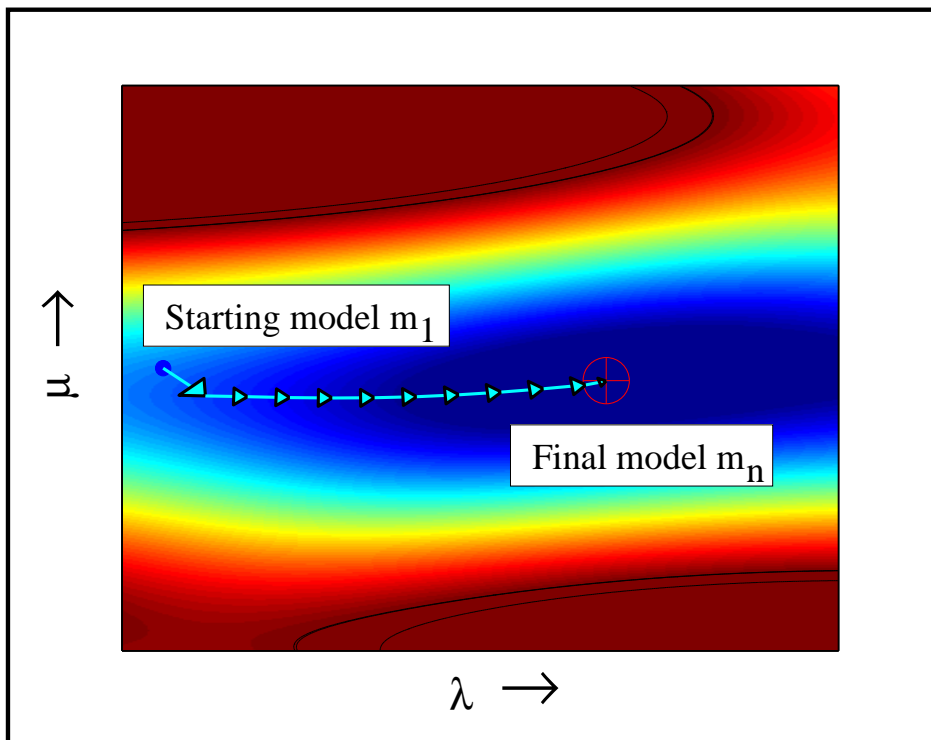


Figure 3.2: Schematic sketch of the residual energy at one point in space as a function of two model parameters  $m_1$  and  $m_2$ . The blue dot denotes the starting point in the parameter space, while the red cross marks a minimum of the objective function.

### CHAPTER 3. THE ADJOINT PROBLEM

---

To find the optimum search direction  $\delta\mathbf{m}_1$  we expand the residual energy  $E(\mathbf{m}_1 + \delta\mathbf{m}_1)$  near the starting point in a Taylor series:

$$E(\mathbf{m}_1 + \delta\mathbf{m}_1) \approx E(\mathbf{m}_1) + \delta\mathbf{m}_1 \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1 + \frac{1}{2} \delta\mathbf{m}_1 \left( \frac{\partial^2 E}{\partial \mathbf{m}^2} \right)_1 \delta\mathbf{m}_1^T \quad (3.5)$$

and set the derivative of Eq. 3.5 with respect to  $\delta\mathbf{m}_1$  zero

$$\frac{\partial E(\mathbf{m}_1 + \delta\mathbf{m}_1)}{\partial \delta\mathbf{m}_1} = \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1 + \delta\mathbf{m}_1 \left( \frac{\partial^2 E}{\partial \mathbf{m}^2} \right)_1 = 0 \quad (3.6)$$

Which finally leads to

$$\delta\mathbf{m}_1 = - \left( \frac{\partial^2 E}{\partial \mathbf{m}^2} \right)_1^{-1} \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1 = -\mathbf{H}_1^{-1} \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1 \quad (3.7)$$

where  $(\partial E/\partial \mathbf{m})_1$  denotes the steepest-descent direction of the objective function and  $\mathbf{H}_1^{-1}$  the inverse Hessian matrix. The inverse Hessian matrix for the elastic problem is often singular and can only be calculated with high computational costs. Therefore the inverse Hessian matrix is approximated by a preconditioning operator  $\mathbf{P}$ . There is no general rule for an optimum preconditioning operator, but two very simple operators are described in more detail in chapter 4 for a cross-well acquisition geometry and in chapter 6.1.3 for a reflection geometry.

$$\delta\mathbf{m}_1 \approx -\mathbf{P}_1 \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1. \quad (3.8)$$

By replacing  $\delta\mathbf{m}_1$  in Eq. 3.4 with Eq. 3.8 we get

$$\mathbf{m}_2 = \mathbf{m}_1 - \mu_1 \mathbf{P}_1 \left( \frac{\partial E}{\partial \mathbf{m}} \right)_1, \quad (3.9)$$

The optimum model parameters can be found along the negative gradient direction of the residual energy. The starting point  $\mathbf{m}_1$  is not a particular point, so the update function can be applied to every point in the parameter space  $\mathbf{m}_n$

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \mathbf{P}_n \left( \frac{\partial E}{\partial \mathbf{m}} \right)_n. \quad (3.10)$$

### 3.3 Calculation of the gradient direction $\frac{\partial E}{\partial \mathbf{m}}$

---

To estimate the gradient direction  $\partial E / \partial \mathbf{m}$  the residual energy is rewritten as:

$$E = \frac{1}{2} \delta \mathbf{u}^T \delta \mathbf{u} = \frac{1}{2} \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \delta \mathbf{u}^2(\mathbf{x}_r, \mathbf{x}_s, t) \quad (3.11)$$

After derivation with respect to a model parameter  $\mathbf{m}$  we get

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{m}} &= \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \frac{\partial \delta \mathbf{u}}{\partial \mathbf{m}} \delta \mathbf{u} \\ &= \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \frac{\partial (\mathbf{u}^{\text{mod}}(\mathbf{m}) - \mathbf{u}^{\text{obs}})}{\partial \mathbf{m}} \delta \mathbf{u} \\ &= \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \frac{\partial \mathbf{u}^{\text{mod}}(\mathbf{m})}{\partial \mathbf{m}} \delta \mathbf{u} \end{aligned} \quad (3.12)$$

Eq. (3.12) can be related to the mapping of small changes from the data to the model space and vice versa (Fig. 3.3). A small change in the model space  $\delta \mathbf{m}$ , e.g. one model parameter at one point in space, will result in a small perturbation of the data space  $\delta \tilde{\mathbf{u}}$ , e.g. one wiggle in the seismic section. If the Fréchet derivative  $\frac{\partial \mathbf{u}}{\partial \mathbf{m}}$  is known, all the small perturbations in model space can be integrated over the model volume  $V$  to calculate the total change in data space (Tarantola (2005)):

$$\delta \tilde{\mathbf{u}}(\mathbf{x}_s, \mathbf{x}_r, t) = \int_V dV \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \delta \mathbf{m}, \quad (3.13)$$

or by introducing the linear operator  $\hat{L}$

$$\delta \tilde{\mathbf{u}} = \hat{L} \delta \mathbf{m} := \int_V dV \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \delta \mathbf{m}.$$

In a similar way small changes in the data space  $\delta \tilde{\mathbf{u}}'$  can be integrated to calculate the total change in the model space  $\delta \mathbf{m}'$  (Tarantola (2005))

$$\delta \mathbf{m}' = \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \right]^* \delta \tilde{\mathbf{u}}', \quad (3.14)$$

or as operator equation

$$\delta \mathbf{m}' = \hat{L}^* \delta \tilde{\mathbf{u}}'.$$

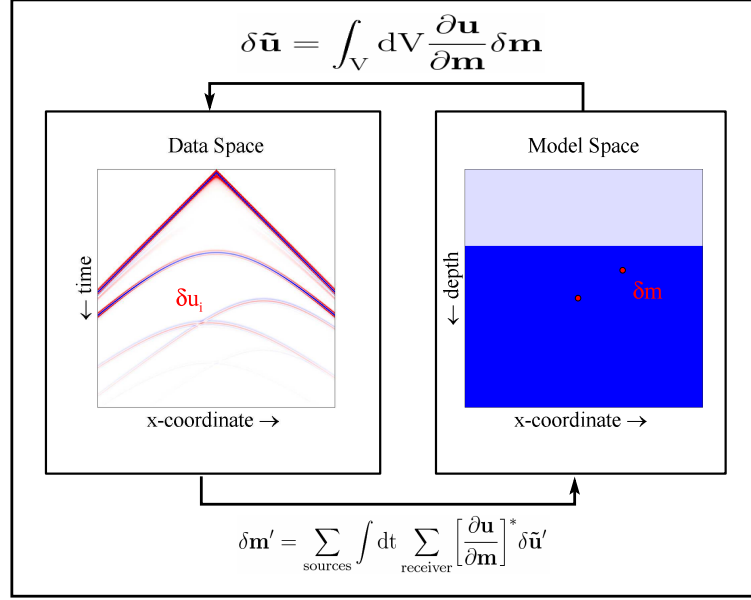


Figure 3.3: Mapping between model and data space and vice versa.

In this case the Frechét derivative  $\frac{\partial \mathbf{u}}{\partial \mathbf{m}}$  is replaced by its adjoint counterpart  $\frac{\partial \mathbf{u}}{\partial \mathbf{m}}^*$ . Note that  $\delta \tilde{\mathbf{u}} \neq \delta \tilde{\mathbf{u}}'$  and  $\delta \mathbf{m} \neq \delta \mathbf{m}'$ , so there is no unique way to map perturbations from the model to the data space or vice versa. Because the operator  $\hat{L}$  is linear, the kernel of  $\hat{L}$  and its adjoint counterpart  $\hat{L}^*$  are identical (see chapter 5.4.2 in Tarantola (2005))

$$\left[ \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \right]^* = \left[ \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \right]$$

Therefore the mapping from the data to the model space Eq. (3.14) is equal to the gradient of the residual energy Eq. (3.12):

$$\begin{aligned}
 \delta \mathbf{m}' &= \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial \mathbf{u}_i}{\partial \mathbf{m}} \right]^* \delta \tilde{\mathbf{u}}' \\
 &= \sum_{\text{sources}} \int dt \sum_{\text{receiver}} \left[ \frac{\partial \mathbf{u}_i}{\partial \mathbf{m}} \right] \delta \mathbf{u} \\
 &= \frac{\partial E}{\partial \mathbf{m}}
 \end{aligned} \tag{3.15}$$

if the perturbation of the data space  $\delta \tilde{\mathbf{u}}'$  is interpreted as data residuals  $\delta \mathbf{u}$ .



### 3.3. CALCULATION OF THE GRADIENT DIRECTION $\frac{\partial E}{\partial \mathbf{M}}$

---

So the approach to estimate the gradient direction  $\partial E/\partial \mathbf{m}$  can be split into 3 parts

1. Find a solution to the forward problem

$$\delta \mathbf{u} = \hat{\mathbf{L}} \delta \mathbf{m}.$$

2. Identify the Frechét kernels  $\partial \mathbf{u}/\partial \mathbf{m}$
3. Use the property, that a linear operator  $\hat{\mathbf{L}}$  and it's adjoint  $\hat{\mathbf{L}}^*$  have the same kernels and calculate the gradient direction by using:

$$\frac{\partial E}{\partial \mathbf{m}} = \delta \mathbf{m}' = \hat{\mathbf{L}}^* \delta \mathbf{u}'.$$

This is a very general approach. Now we apply this approach to the equations of motion for an elastic medium. The following derivation is much easier, when assuming a general elastic medium first and introduce the isotropy later on. Therefore the elastic forward problem Eqs. (2.30) can be written as

$$\begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \sigma_{ij} &= f_i, \\ \sigma_{ij} - c_{ijkl} \epsilon_{kl} &= T_{ij}, \\ \epsilon_{ij} &= \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \\ &+ \text{initial and boundary conditions,} \end{aligned} \tag{3.16}$$

where  $\rho$  denotes the density,  $\sigma_{ij}$  the stress tensor,  $\epsilon_{ij}$  the strain tensor,  $c_{ijkl}$  the stiffness tensor,  $f_i$ ,  $T_{ij}$  source terms for volume and surface forces, respectively. In the next step every parameter and variable in the elastic wave equation is perturbed by a first order perturbation as shown in Fig. 3.3:

$$\begin{aligned} u_i &\rightarrow u_i + \delta u_i, \\ \rho &\rightarrow \rho + \delta \rho, \\ \sigma_{ij} &\rightarrow \sigma_{ij} + \delta \sigma_{ij}, \\ c_{ijkl} &\rightarrow c_{ijkl} + \delta c_{ijkl}, \\ \epsilon_{ij} &\rightarrow \epsilon_{ij} + \delta \epsilon_{ij}, \end{aligned} \tag{3.17}$$

### CHAPTER 3. THE ADJOINT PROBLEM

---

These substitutions yield new equations of motion describing the displacement perturbations  $\delta u_i$  and stress perturbations  $\delta \sigma_{ij}$  as a function of new source terms  $\Delta f_i$  and  $\Delta T_{ij}$  (see appendix A)

$$\begin{aligned} \rho \frac{\partial^2 \delta u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \delta \sigma_{ij} &= \Delta f_i, \\ \delta \sigma_{ij} - c_{ijkl} \delta \epsilon_{kl} &= \Delta T_{ij}, \\ \delta \epsilon_{ij} &= \frac{1}{2} \left( \frac{\partial \delta u_i}{\partial x_j} + \frac{\partial \delta u_j}{\partial x_i} \right) \\ &+ \text{perturbated initial and boundary conditions} \end{aligned} \tag{3.18}$$

The new source terms are

$$\Delta f_i = -\delta \rho \frac{\partial^2 u_i}{\partial t^2} \tag{3.19}$$

and

$$\Delta T_{ij} = \delta c_{ijkl} \epsilon_{kl}. \tag{3.20}$$

Two points are important to notice:

1. Eq.(3.18) states that every change of a material parameter acts as a source (Eq.(3.19) and Eq.(3.20)), but the perturbated wavefield is propagating in the unperturbated medium.
2. The new wave equation (3.18) has mathematically the same form as the unperturbated elastic wave equation, and hence its solution can be obtained in terms of Green's functions  $G_{ij}$  of the elastic wave equation.

The solution of the perturbated elastic equations of motion (3.18) in terms of the elastic Green's function  $G_{ij}(\mathbf{x}, t; \mathbf{x}', t')$  (see Eq. B.4 in appendix B) can be written as:

$$\begin{aligned} \delta u_i(\mathbf{x}, t) &= \int_V dV \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') \Delta f_j(\mathbf{x}', t') \\ &\quad - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \Delta T_{jk}(\mathbf{x}', t'). \end{aligned} \tag{3.21}$$

### 3.3. CALCULATION OF THE GRADIENT DIRECTION $\frac{\partial E}{\partial \mathbf{M}}$

Substituting the force and traction terms given by Eqs.(3.19) and (3.20) into Eq.(3.21) yields after some rearranging

$$\begin{aligned} \delta u_i(\mathbf{x}, t) = & - \int_V dV \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}', t') \delta \rho \\ & - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \epsilon_{lm}(\mathbf{x}', t') \delta c_{ijklm} \end{aligned} \quad (3.22)$$

Introducing isotropy via Eq. (2.27) leads to:

$$\begin{aligned} \delta u_i(\mathbf{x}, t) = & - \int_V dV \left[ \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}', t') \right] \delta \rho \\ & - \int_V dV \left[ \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \epsilon_{lm}(\mathbf{x}', t') \delta_{jk} \delta_{lm} \right] \delta \lambda \\ & - \int_V dV \left[ \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \epsilon_{lm}(\mathbf{x}', t') (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \right] \delta \mu. \end{aligned} \quad (3.23)$$

Utilization of Eq.(3.23) to solve the forward problem is known as the Born approximation. In waveform tomography the Born approximation is not used to solve the forward problem. Instead the full elastic wave equation is solved. Equation (3.23) has the same form as the desired expression for the forward problem Eqs.(3.13):

$$\delta \mathbf{u} = \int_V dV \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \delta \mathbf{m}. \quad (3.24)$$

Therefore the Frechét kernels  $\frac{\partial u_i}{\partial \mathbf{m}(\mathbf{x})}$  for the individual material parameters can be identified as:

$$\begin{aligned} \frac{\partial u_i}{\partial \rho} &= - \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}', t') \\ \frac{\partial u_i}{\partial \lambda} &= - \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \epsilon_{lm}(\mathbf{x}', t') \delta_{jk} \delta_{lm} \\ \frac{\partial u_i}{\partial \mu} &= - \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') \epsilon_{lm}(\mathbf{x}', t') (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \end{aligned} \quad (3.25)$$

By definition the adjoint of the operator (3.24) can be written as

$$\delta \mathbf{m}'(\mathbf{x}) = \sum_{\text{sources}} \int_0^T dt \sum_{\alpha=1}^{N_{\text{rec}}} \left[ \frac{\partial u_i}{\partial \mathbf{m}} \right]^* \delta u'_i(\mathbf{x}_\alpha, t'), \quad (3.26)$$

### CHAPTER 3. THE ADJOINT PROBLEM

---

Because a linear operator and its transpose have the same kernels  $\partial u_i / \partial \mathbf{m}$ , the only difference arise in the variables of sum/integration, which are complementary. Inserting the integral kernels (3.25) in Eq.(3.26) yields

$$\begin{aligned}\delta\rho' &= - \sum_{\text{sources}} \int_0^T dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' G_{ij}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}, t) \delta u'_i(\mathbf{x}_\alpha, t'), \\ \delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' \frac{\partial G_{ij}}{\partial x_k}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \epsilon_{lm}(\mathbf{x}, t) \delta_{jk} \delta_{lm} \delta u'_i(\mathbf{x}_\alpha, t'), \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' \frac{\partial G_{ij}}{\partial x_k}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \epsilon_{lm}(\mathbf{x}, t) (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \delta u'_i(\mathbf{x}_\alpha, t').\end{aligned}$$

The terms only depending on time  $t$  and the positions  $\mathbf{x}$  can be moved in front of the sum over the receivers

$$\begin{aligned}\delta\rho' &= - \sum_{\text{sources}} \int_0^T dt \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}, t) \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' G_{ij}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \delta u'_i(\mathbf{x}_\alpha, t'), \\ \delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \epsilon_{lm}(\mathbf{x}, t) \delta_{jk} \delta_{lm} \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' \frac{\partial G_{ij}}{\partial x_k}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \delta u'_i(\mathbf{x}_\alpha, t'), \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \epsilon_{lm}(\mathbf{x}, t) (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' \frac{\partial G_{ij}}{\partial x_k}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \delta u'_i(\mathbf{x}_\alpha, t').\end{aligned}\tag{3.27}$$

Defining the wavefield

$$\Psi_j(\mathbf{x}, t) = \sum_{\alpha=1}^{N_{\text{rec}}} \int_0^T dt' G_{ij}(\mathbf{x}_\alpha, t'; \mathbf{x}, t) \delta u'_i(\mathbf{x}_\alpha, t'),\tag{3.28}$$

Eqs.(3.27) can be written as

$$\begin{aligned}\delta\rho' &= - \sum_{\text{sources}} \int_0^T dt \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}, t) \Psi_j, \\ \delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \epsilon_{lm}(\mathbf{x}, t) \delta_{jk} \delta_{lm} \frac{\partial \Psi_j}{\partial x_k}, \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \epsilon_{lm}(\mathbf{x}, t) (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \frac{\partial \Psi_j}{\partial x_k}.\end{aligned}\tag{3.29}$$

### 3.3. CALCULATION OF THE GRADIENT DIRECTION $\frac{\partial E}{\partial \mathbf{M}}$

The wavefield  $\Psi_j$  is generated by propagating the residual data  $\delta u'_i$  from the receiver positions backwards in time through the elastic medium. To obtain a more symmetric expression for the density gradient, let us integrate the density gradient in (3.29) by parts

$$\begin{aligned}\delta\rho' &= - \sum_{\text{sources}} \int_0^T dt \left( \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}, t) \right) \Psi_j(\mathbf{x}, t) \\ &= - \sum_{\text{sources}} \left\{ \left[ \frac{\partial u_j}{\partial t}(\mathbf{x}, T) \Psi_j(\mathbf{x}, T) \right]_0^T - \int_0^T dt \frac{\partial u_j}{\partial t}(\mathbf{x}, t) \frac{\partial \Psi_j}{\partial t}(\mathbf{x}, t) \right\}.\end{aligned}\tag{3.30}$$

According to Eqs. (2.39) the field  $u_j(\mathbf{x}, t)$  satisfies initial conditions of rest,  $u_j(\mathbf{x}, 0) = 0$  and  $\partial u_j(\mathbf{x}, 0)/\partial t = 0$ . The field  $\Psi_j(\mathbf{x}, t)$  satisfies final conditions of rest,  $\Psi_j(\mathbf{x}, T) = 0$ . Therefore

$$\delta\rho' = - \sum_{\text{sources}} \int_0^T dt \left( \frac{\partial^2 u_j}{\partial t^2}(\mathbf{x}, t) \right) \Psi_j(\mathbf{x}, t) = \sum_{\text{sources}} \int_0^T dt \frac{\partial u_j}{\partial t}(\mathbf{x}, t) \frac{\partial \Psi_j}{\partial t}(\mathbf{x}, t).\tag{3.31}$$

Writing out the implicit sums in the gradients of the Lamé parameters  $\delta\lambda'$  and  $\delta\mu'$  in Eqs. (3.29)

$$\begin{aligned}\delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \sum_l \sum_k \sum_j \sum_m \epsilon_{lm}(\mathbf{x}, t) \delta_{jk} \delta_{lm} \frac{\partial \Psi_j}{\partial x_k}, \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \sum_l \sum_k \sum_j \sum_m \epsilon_{lm}(\mathbf{x}, t) (\delta_{jl} \delta_{lm} + \delta_{jm} \delta_{kl}) \frac{\partial \Psi_j}{\partial x_k}.\end{aligned}\tag{3.32}$$

and neglecting all wavefield components and derivatives in z-direction leads to

$$\begin{aligned}\delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \left( \epsilon_{xx} + \epsilon_{yy} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right), \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \left[ \left( \epsilon_{xy} + \epsilon_{yx} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) \right] + 2 \left( \epsilon_{xx} \frac{\partial \Psi_x}{\partial x} + \epsilon_{yy} \frac{\partial \Psi_y}{\partial y} \right).\end{aligned}\tag{3.33}$$

Using the definition of the strain tensor  $\epsilon_{ij}$  we get

$$\begin{aligned}\delta\lambda' &= - \sum_{\text{sources}} \int_0^T dt \left( \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right), \\ \delta\mu' &= - \sum_{\text{sources}} \int_0^T dt \left[ \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) \right] + 2 \left( \frac{\partial u_x}{\partial x} \frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y} \frac{\partial \Psi_y}{\partial y} \right).\end{aligned}\tag{3.34}$$

## CHAPTER 3. THE ADJOINT PROBLEM

---

Finally the gradients for the Lamé parameters  $\lambda$ ,  $\mu$  and the density  $\rho$  can be written as

$$\begin{aligned}\delta\lambda' &= - \sum_{\text{sources}} \int dt \left( \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right) \\ \delta\mu' &= - \sum_{\text{sources}} \int dt \left( \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \left( \frac{\partial \Psi_x}{\partial y} + \frac{\partial \Psi_y}{\partial x} \right) + 2 \left( \frac{\partial u_x}{\partial x} \frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y} \frac{\partial \Psi_y}{\partial y} \right) \\ \delta\rho' &= \sum_{\text{sources}} \int dt \left( \frac{\partial u_x}{\partial t} \frac{\partial \Psi_x}{\partial t} + \frac{\partial u_y}{\partial t} \frac{\partial \Psi_y}{\partial t} \right)\end{aligned}\quad (3.35)$$

### 3.4 Estimation of an optimum step length $\mu_n$

---

The choice of the step length  $\mu_n$  in Eq. 3.10 is crucial for the convergence of the steepest gradient optimization method. I demonstrate this using a very familiar test problem for optimization routines, the Rosenbrock function (Rosenbrock (1960), Fig. 3.4)

$$f_r(x, y) = (1 - x)^2 + 100(y - x^2)^2 \quad (3.36)$$

The aim is to find the minimum of this function located at the point [1,1] which is surrounded by a very narrow valley. We start the search for the minimum at [-0.5,0.5]. An obvious first choice would be a constant step length. Fig. 3.4 (top) shows the convergence after 16000 iteration steps of the steepest descent method when choosing a step length  $\mu_n = 2e - 3$ . Note the large model update during the first iteration step, when the gradient of the Rosenbrock function is large. After reaching the narrow valley the gradient is much smaller and as a result the model updates are also decreasing. This leads to a very slow convergence speed. Especially near the minimum the model updates become very small. When choosing a larger step length ( $\mu_n = 2e - 3$ , Fig. 3.4 (bottom)) the model update is larger even when the gradient is small, but the code fails to converge at all. Instead it is trapped in a narrow part of the valley. To solve this problem a variable step length is introduced. For three test step lengths  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  three test models are calculated

$$\begin{aligned}\mathbf{m}_{\text{test1}} &= \mathbf{m}_n + \mu_1 \delta \mathbf{m}'_n \\ \mathbf{m}_{\text{test2}} &= \mathbf{m}_n + \mu_2 \delta \mathbf{m}'_n \\ \mathbf{m}_{\text{test3}} &= \mathbf{m}_n + \mu_3 \delta \mathbf{m}'_n\end{aligned}\quad (3.37)$$

and the corresponding L2-norms  $L_{2_1}$ ,  $L_{2_2}$  and  $L_{2_3}$  are estimated (Fig. 3.5). The true misfit function (yellow line) can be approximated by fitting a parabola through

### 3.4. ESTIMATION OF AN OPTIMUM STEP LENGTH $\mu_N$

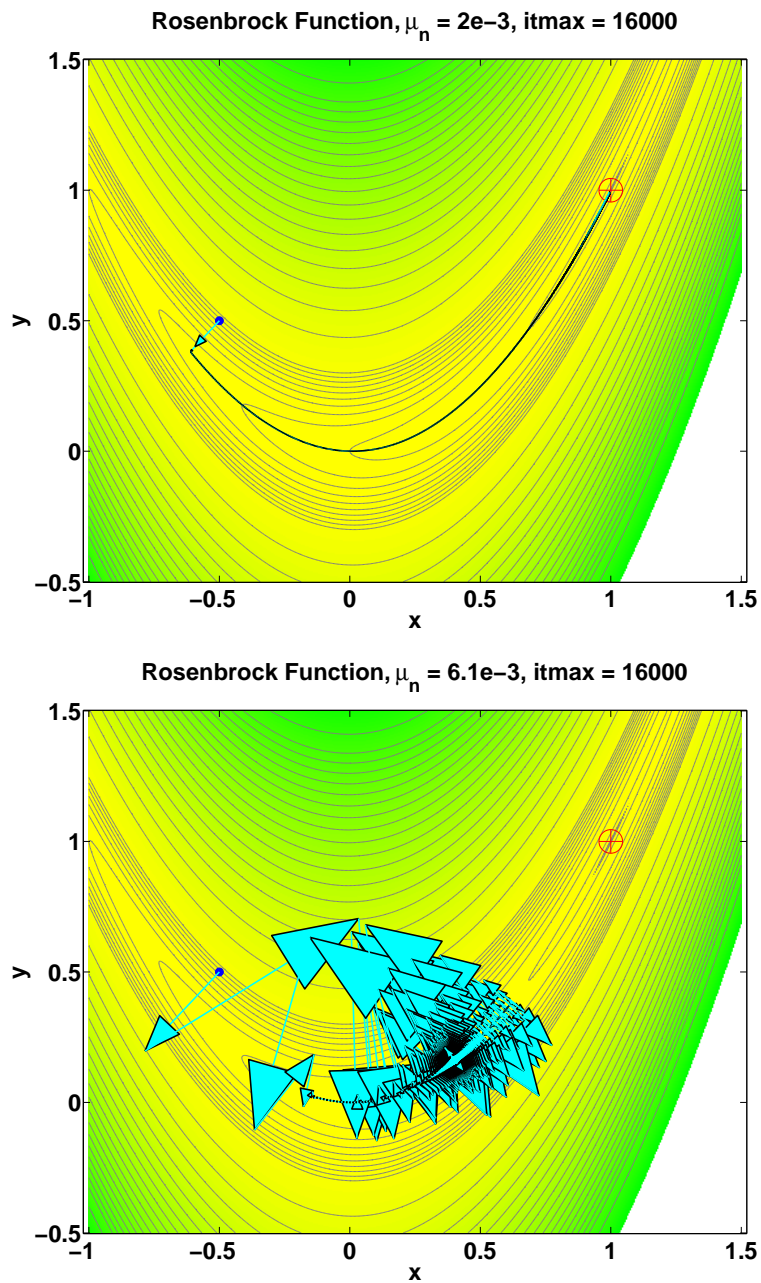


Figure 3.4: Results of the convergence test for the Rosenbrock function. The minimum is marked with a red cross, the starting point with a blue point. The maximum number of iterations is 16000. The step length  $\mu_n$  varies between  $2e-3$  (top) and  $6.1e-3$  (bottom).

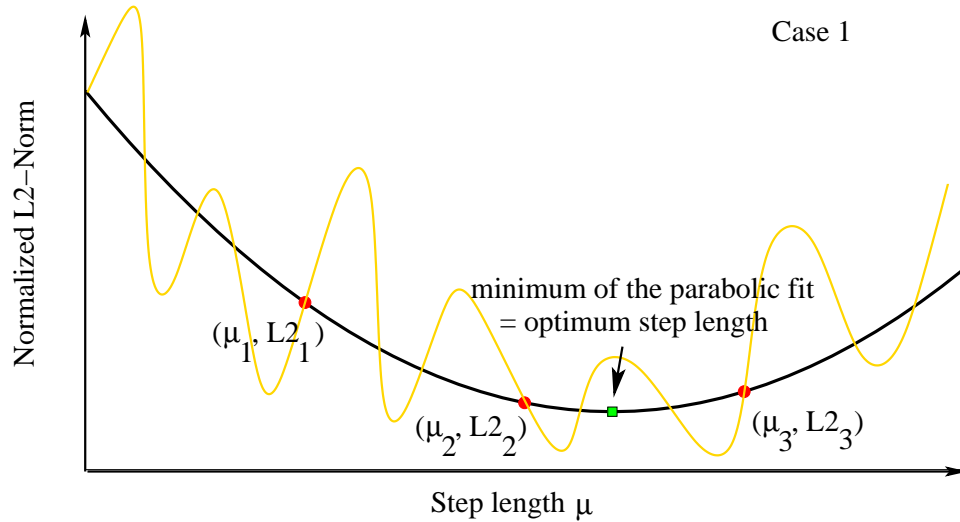


Figure 3.5: Line search algorithm to find the optimum step length  $\mu_{\text{opt}}$ : The true misfit function (yellow line) is approximated by a parabola fitted by 3 points.

the three points

$$L2_i = a\mu_i^2 + b\mu_i + c \quad (3.38)$$

where  $i \in \{1, 2, 3\}$  and  $a, b, c$  are the unknown coefficients. This system of equations can be written as matrix equation:

$$\begin{pmatrix} \mu_1^2 & \mu_1 & 1 \\ \mu_2^2 & \mu_2 & 1 \\ \mu_3^2 & \mu_3 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} L2_1 \\ L2_2 \\ L2_3 \end{pmatrix}$$

or

$$\mathbf{Ax} = \mathbf{b}. \quad (3.39)$$

The unknown coefficients of this matrix equation are formally defined by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \quad (3.40)$$

In the FWT code the solution vector  $\mathbf{x}$  is calculated by Gaussian elimination. In the following the step length at the extremum of the parabola is defined the extremum



### 3.4. ESTIMATION OF AN OPTIMUM STEP LENGTH $\mu_N$

---

step length  $\mu_{\text{ext}}$  (denoted as green square in Fig.3.5 - 3.7). This extremum step length is

$$\mu_{\text{ext}} = -\frac{b}{2a}. \quad (3.41)$$

The optimum step length  $\mu_{\text{opt}}$  (denoted as blue triangle in Fig.3.5 - 3.7) for the update of the model parameters can be estimated by distinguishing 6 different cases

1. The parabola is convex ( $2a > 0.0$ , Fig.3.5). The optimum step length is defined by

$$\mu_{\text{opt}} = \mu_{\text{ext}} = -\frac{b}{2a}. \quad (3.42)$$

2. The parabola is concave ( $2a < 0.0$ , Fig.3.6, top) and the maximum is located near the point  $[\mu_2, L2_2]$ . The optimum step length is defined as

$$\mu_{\text{opt}} = \mu_1. \quad (3.43)$$

3. The parabola is concave ( $2a < 0.0$ , Fig.3.6, bottom) and the maximum is located near the point  $[\mu_1, L2_1]$ . The optimum step length is defined as

$$\mu_{\text{opt}} = \mu_3. \quad (3.44)$$

4. The parabola is concave ( $2a < 0.0$ , Fig.3.7, top) and the maximum is located near the point  $[\mu_3, L2_3]$ . The optimum step length is defined as

$$\mu_{\text{opt}} = \mu_1. \quad (3.45)$$

5. The parabola is convex ( $2a > 0.0$ , Fig.3.7, bottom) and the minimum step length is negative. The optimum step length is defined as

$$\mu_{\text{opt}} = \mu_1. \quad (3.46)$$

6. Additionally if the parabola is convex ( $2a > 0.0$ ) and the minimum step length is 5 times larger than  $\mu_3$ . The optimum step length is defined as

$$\mu_{\text{opt}} = \mu_3. \quad (3.47)$$

## CHAPTER 3. THE ADJOINT PROBLEM

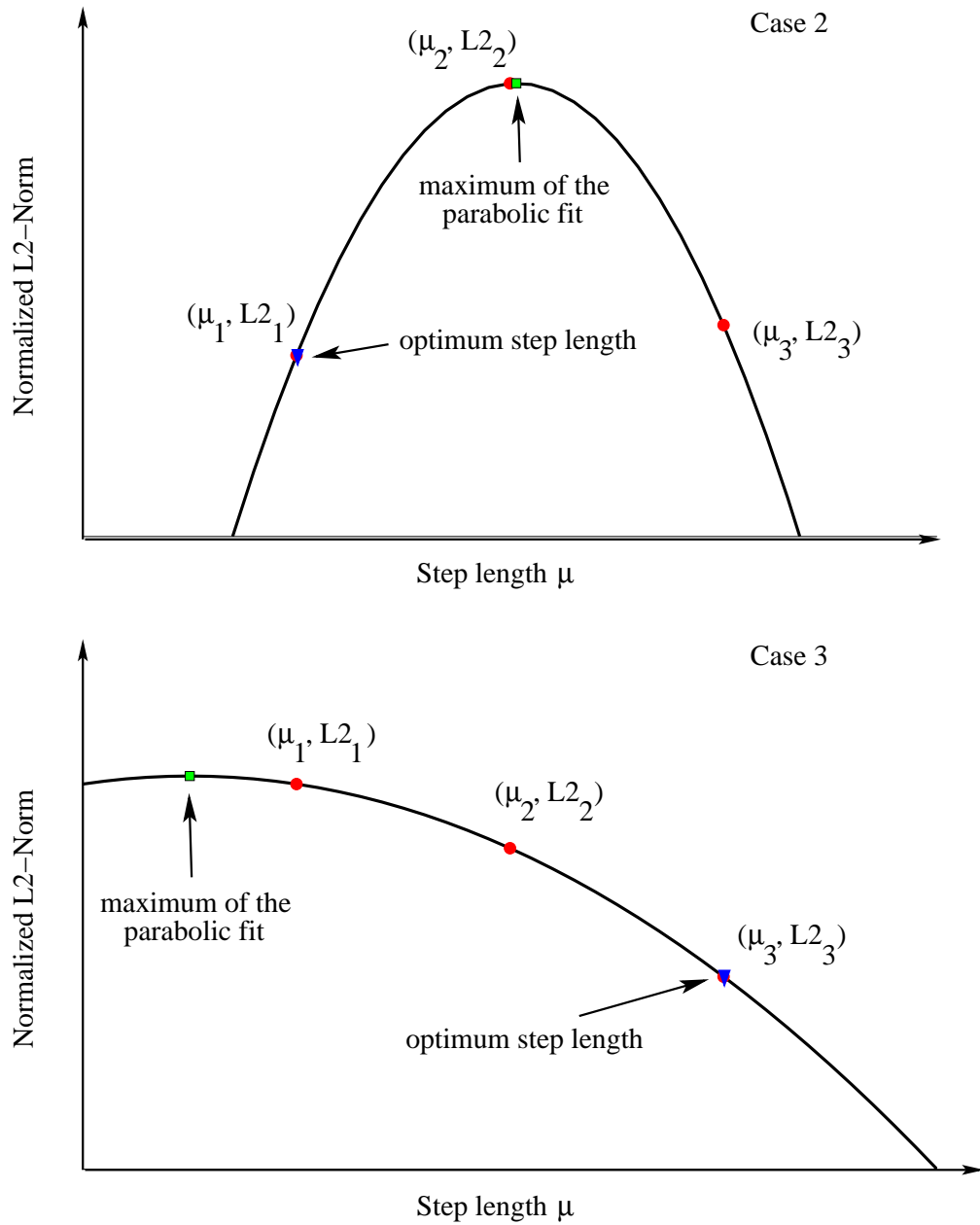


Figure 3.6: Optimum step length calculation by parabolic fitting: Case 2 (top) and 3 (bottom).

### 3.4. ESTIMATION OF AN OPTIMUM STEP LENGTH $\mu_N$

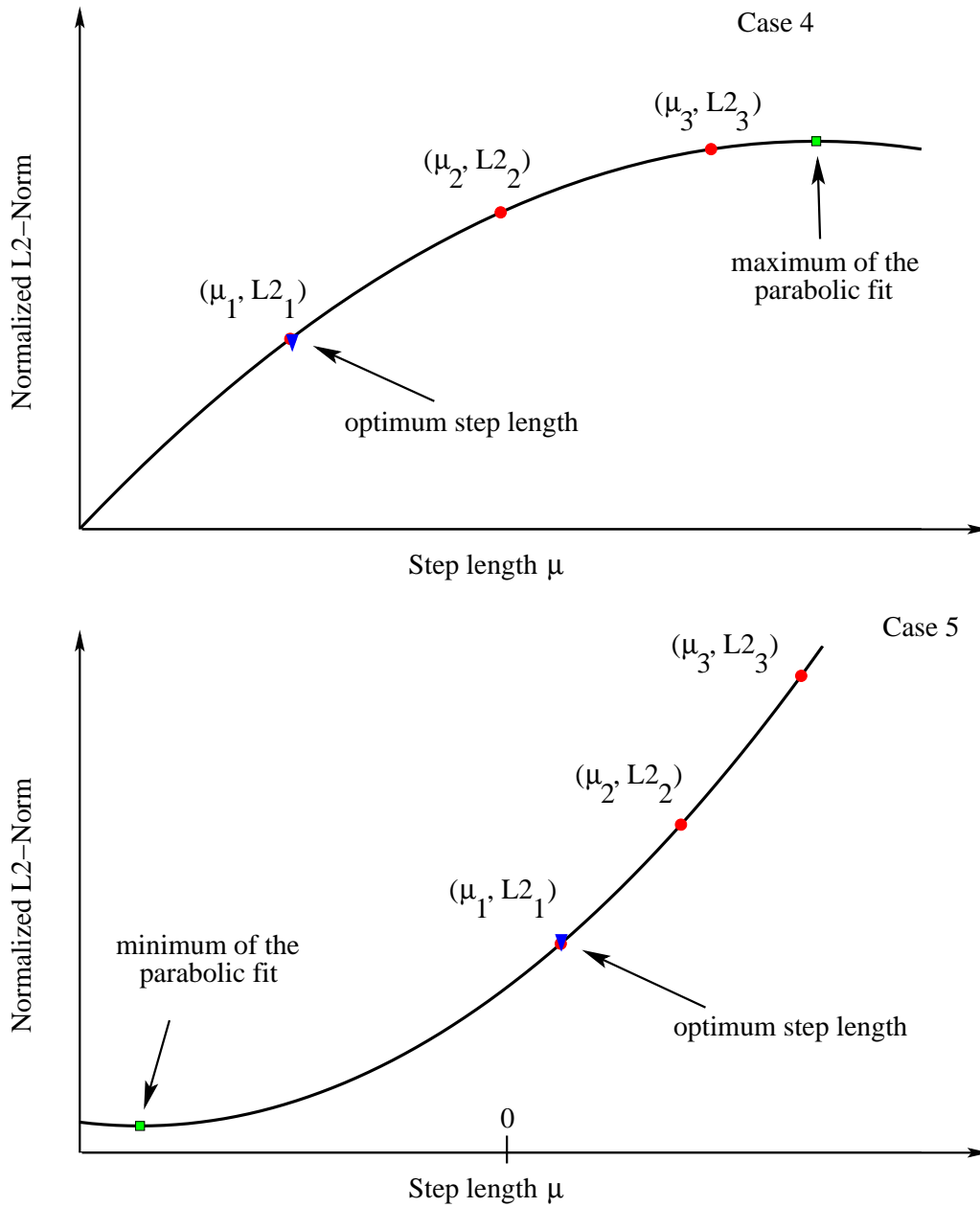


Figure 3.7: Optimum step length calculation by parabolic fitting: Case 4 (top) and 5 (bottom).

## CHAPTER 3. THE ADJOINT PROBLEM

---

The application of the variable step length calculation to the Rosenbrock test problem is shown in Fig. 3.8. The number of required iteration steps to reach the minimum is reduced by a factor 4 when compared with the constant step length gradient method. The only problem remaining is the slow convergence speed in the small valley of the Rosenbrock function, due to the fact that the update occurs along the gradient direction of the objective function resulting in a "criss-cross" pattern. This behaviour can be avoided by applying a nonlinear conjugate gradient method (chapter 3.5). In case of the FWT algorithm the three test step lengths for the individual material parameters are calculated by scaling the maximum of the gradient to the maximum of the actual models:

$$\begin{aligned}\mu_\lambda &= p \frac{\max(\lambda_n)}{\max(\delta\lambda_n)} \\ \mu_\mu &= p \frac{\max(\mu_n)}{\max(\delta\mu_n)} \\ \mu_\rho &= p \frac{\max(\rho_n)}{\max(\delta\rho_n)}\end{aligned}\tag{3.48}$$

For most tests in the following chapters  $p_1 = 0.0025$ ,  $p_2 = 0.005$ ,  $p_3 = 0.01$ , which corresponds to maximum model changes of 1/4, 1/2 and 1 %, worked very well for the optimum step length estimation. All material parameters are updated at the same time. To save computational time the corresponding  $L_2$ -norms are calculated for a few representative shots (in most cases 3). For the acoustic case the step length estimation by parabolic fitting works very well and leads to a smooth decrease of the misfit function during the FWT (Kurzmann (2007), personal communication, Kurzmann et al. (2008)). For the multiparameter elastic FWT the misfit function consists of more local minima and therefore the decrease of the objective function is not as smooth as in the acoustic case. Brossier (2009) proposed a more intensive bracketing stage before applying the parabolic fit. For  $p_1 = 0.0$  the test step lengths  $p_2$  and  $p_3$  are calculated to satisfy the following criteria:

$$\begin{aligned}L2_2(\mathbf{m}_{\text{test2}} = \mathbf{m}_n + \mu_2\delta\mathbf{m}'_n) &< L2_1(\mathbf{m}_{\text{test1}} = \mathbf{m}_n) \\ L2_3(\mathbf{m}_{\text{test3}} = \mathbf{m}_n + \mu_3\delta\mathbf{m}'_n) &> L2_2(\mathbf{m}_{\text{test2}} = \mathbf{m}_n + \mu_2\delta\mathbf{m}'_n)\end{aligned}\tag{3.49}$$

This approach leads to a smoother decrease of the objective function, but also increases the number of required forward models.

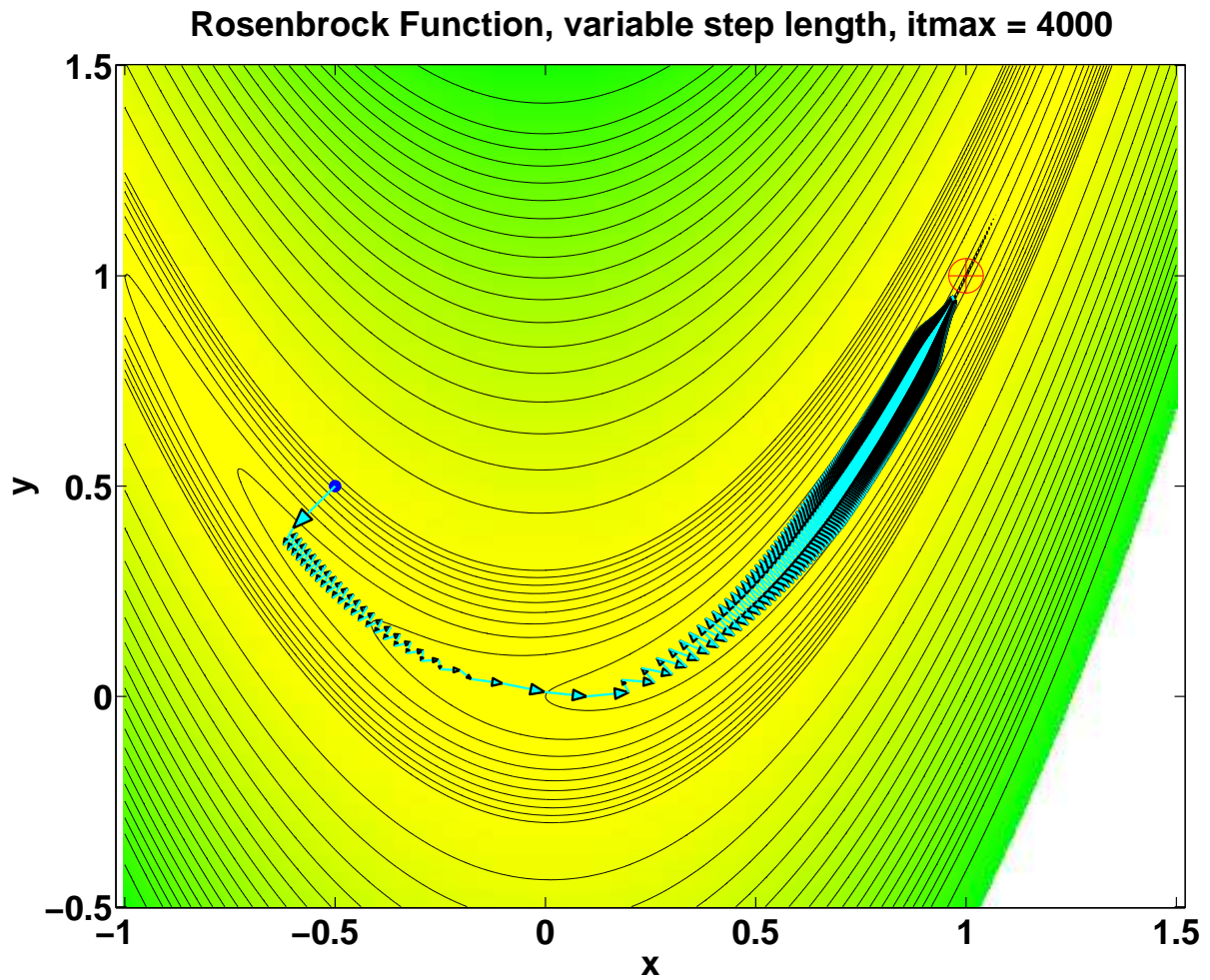


Figure 3.8: Results of the convergence test for the Rosenbrock function. The minimum is marked by a red cross, the starting point by a blue point. The maximum number of iterations is 4000. The optimum step length is calculated at each iteration by the parabola fitting algorithm. Note the criss-cross pattern of the updates in the narrow valley near the minimum.

### 3.5 Nonlinear Conjugate Gradient Method

---

To increase the convergence speed in narrow valleys it would be better to update the model at iteration step  $n$  not exactly along the gradient direction  $\delta\mathbf{m}_n$ , but along the conjugate direction  $\delta\mathbf{c}_n$

$$\delta\mathbf{c}_n = \delta\mathbf{m}_n + \beta_n\delta\mathbf{c}_{n-1}, \quad (3.50)$$

The first iteration step ( $n=1$ ) consists of a model update along the steepest descent direction:

$$\mathbf{m}_2 = \mathbf{m}_1 + \mu_1\delta\mathbf{m}_1, \quad (3.51)$$

For all subsequent iteration steps ( $n > 1$ ) the model is updated along the conjugate direction:

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \mu_n\delta\mathbf{c}_n, \quad (3.52)$$

where  $\delta\mathbf{c}_1 = \delta\mathbf{m}_1$ . The weighting factor  $\beta$  can be calculated in different ways:

1. Fletcher-Reeves:

$$\beta_n^{\text{FR}} = \frac{\delta\mathbf{m}_n^T\delta\mathbf{m}_n}{\delta\mathbf{m}_{n-1}^T\delta\mathbf{m}_{n-1}} \quad (3.53)$$

2. Polak-Ribière:

$$\beta_n^{\text{PR}} = \frac{\delta\mathbf{m}_n^T(\delta\mathbf{m}_n - \delta\mathbf{m}_{n-1})}{\delta\mathbf{m}_{n-1}^T\delta\mathbf{m}_{n-1}} \quad (3.54)$$

3. Hestenes-Stiefel:

$$\beta_n^{\text{HS}} = \frac{\delta\mathbf{m}_n^T(\delta\mathbf{m}_n - \delta\mathbf{m}_{n-1})}{\delta\mathbf{c}_{n-1}^T(\delta\mathbf{m}_n - \delta\mathbf{m}_{n-1})} \quad (3.55)$$

I use the very popular choice  $\beta_n = \max[0, \beta_n^{\text{PR}}]$  which provides an automatic direction reset. This is important because subsequent search directions lose conjugacy requiring the search direction to be reset to the steepest descent direction. Note that the conjugate gradient method doesn't require any additional computational time because only the gradient  $\delta\mathbf{m}_n$  at two subsequent iterations has to be known. The application of the nonlinear conjugate gradient method combined with the variable step length calculation to the Rosenbrock function is shown in Fig. 3.9. The criss-cross pattern of the steepest descent method has vanished. The conjugate gradient method converges already after 2000 iterations compared with 4000 iteration steps of the pure gradient method.

### 3.5. NONLINEAR CONJUGATE GRADIENT METHOD

Rosenbrock Function, conjugate gradient, variable step length, itmax = 2000

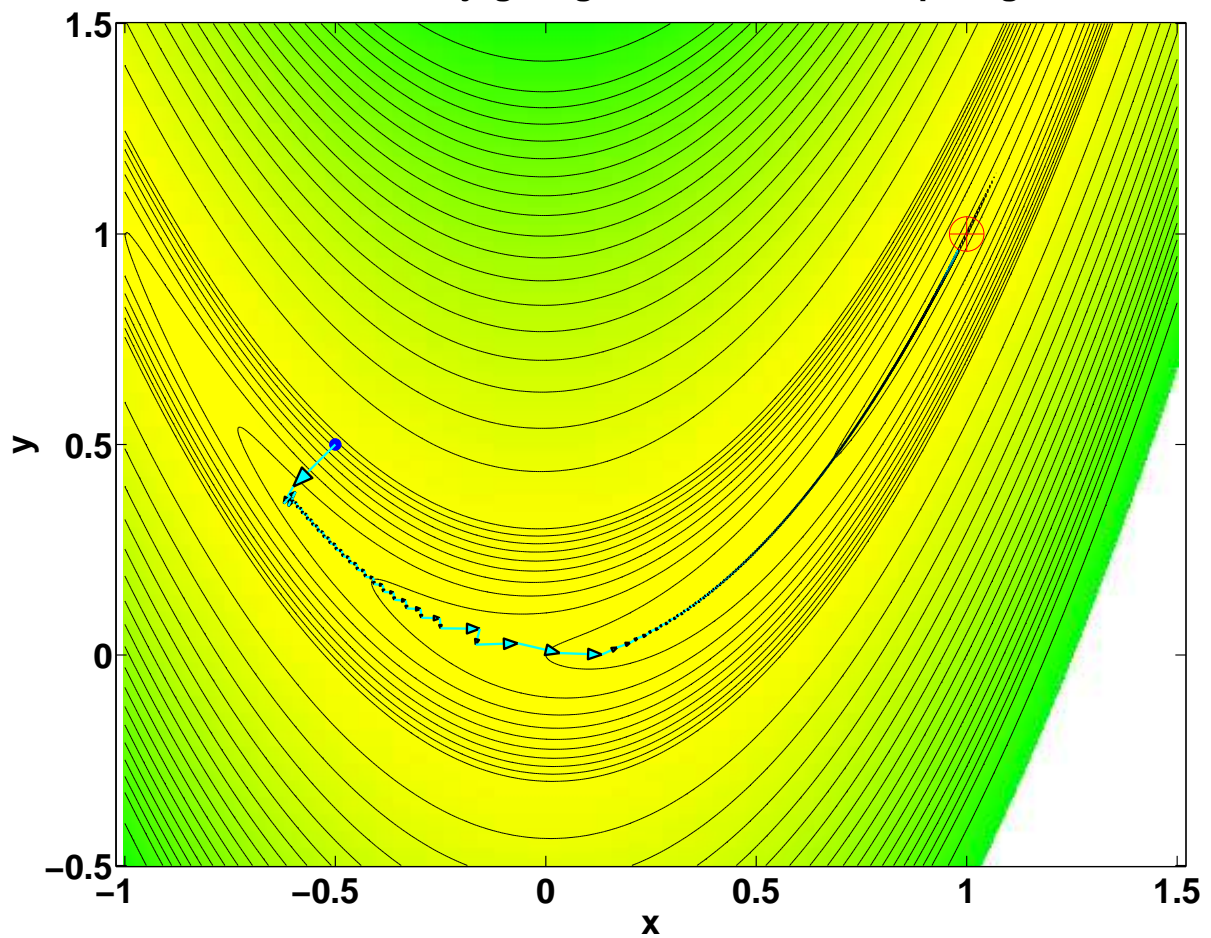


Figure 3.9: Results of the convergence test for the Rosenbrock function using the conjugate gradient method, where the optimum step length is calculated with the parabolic fitting algorithm. The minimum is marked by a red cross, the starting point by a blue point. The maximum number of iterations is 2000.

### 3.6 The elastic FWT algorithm

---

In summary the FWT algorithm consists of the following steps:

1. Define a starting model  $\mathbf{m}_1$  in the parameter space. This model should represent the long wavelength part of the underground very well, because the FWT code is only capable to reconstruct structures at or below the dominant seismic wavelength due to its slow convergence speed, the nonlinearity of the problem and the inherent use of the Born approximation to calculate the gradient direction.
2. At iteration step  $n$  do:
  - (a) For each shot solve the forward problem, stated in Eq.(3.16) for the actual model  $\mathbf{m}_n$  to generate a synthetic dataset  $\mathbf{u}^{\text{mod}}$  and the wavefield  $\mathbf{u}(\mathbf{x}, t)$ .
  - (b) Calculate the residual seismograms  $\delta\mathbf{u} = \mathbf{u}^{\text{mod}} - \mathbf{u}^{\text{obs}}$  for the x- and y-components of the seismic data.
  - (c) Generate the wavefield  $\Psi(\mathbf{x}, t)$  by backpropagating the residuals from the receiver postions.
  - (d) Calculate the gradients  $\delta\mathbf{m}_n$  of each material parameter according to Eqs.(3.35).
  - (e) To increase the convergence speed an appropriate preconditioning operator  $P$  is applied to the gradient  $\delta\mathbf{m}$

$$\delta\mathbf{m}_n^p = P\delta\mathbf{m}_n \quad (3.56)$$

Examples of simple preconditioning operator are given in chapter 4 for a cross-well acquisition geometry and in chapter 6.1.3 for a reflection geometry.

- (f) For a further increase of the convergence speed calculate the conjugate gradient direction for iteration steps  $n \geq 2$ :

$$\delta\mathbf{c}_n = \delta\mathbf{m}_n^p + \beta\delta\mathbf{c}_{n-1}, \text{ with } \delta\mathbf{c}_1 = \delta\mathbf{m}_1^p \quad (3.57)$$

where the weighting factor

$$\beta^{\text{PR}} = \delta\mathbf{m}_n^p \frac{\delta\mathbf{m}_n^p - \delta\mathbf{m}_{n-1}^p}{\delta\mathbf{m}_{n-1}^p \delta\mathbf{m}_{n-1}^p} \quad (3.58)$$

by Polak-Ribière is used. The convergence of the Polak-Ribière method is guaranteed by choosing  $\beta = \max[\beta^{\text{PR}}, 0]$ .



### 3.6. THE ELASTIC FWT ALGORITHM

---

- (g) Estimate the step length  $\mu_n$  by the line search algorithm described in chapter 3.4.
- (h) Update the material parameters using the gradient method

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \delta \mathbf{c}_n. \quad (3.59)$$

If the material parameters are not coupled by empirical relationships it is important to update all three elastic material parameters at the same time, otherwise strong artefacts may dominate the inversion result, especially in the case of very complex media.

3. If the residual energy  $E$  is smaller than a given value stop the iteration. Otherwise continue with the next iteration step.

This FWT algorithm is implemented in the DENISE (subwavelength **DE**tail resolving **Nonlinear Iterative SE**ismic inversion) 2D time domain code. The source code is written in C and uses the stress-displacement FD formulation (chapter 2.2) to solve the forward modeling and backpropagation of the data residuals (steps (a) and (c)). The code is parallelized with the domain decomposition approach using MPI (chapter 2.4).



# 4

## Simple example: A spherical low velocity anomaly

The different steps from building an FD model to solve the forward problem to the application of the FWT algorithm are illustrated by the following simple test problem. The test problem (Fig. 4.1) has the dimensions 160x260 m and consists of a homogeneous acoustic fullspace with a P-wave velocity  $V_p = 2000.0$  m/s, S-wave velocity  $V_s = 0.0$  m/s and a density  $\rho = 1000.0$  kg/m<sup>3</sup>, respectively. Embedded in the fullspace is a spherical inclusion with a diameter of 20 m, a P-wave velocity  $V_p = 1700.0$  m/s and the same density as the surrounding medium. For the cross well acquisition geometry 91 explosive sources (red stars in Fig. 4.1) are located in a depth between 40 m and 220 m. The source signature is a Ricker wavelet with a centre frequency of 125 Hz and a maximum frequency of 250 Hz. The acoustic wavefield is recorded by a line of 180 two component receivers (blue circles in Fig. 4.1) at the same depth as the sources but with a horizontal offset of 80 m. To implement a fullspace the model is surrounded by an exponential damping absorbing boundary layer with a thickness of 20 m. To calculate the dimensions of the FD grid the grid spacing  $dh$  is estimated by the grid dispersion criterion (2.49). For a 2nd order spatial FD operator with  $V_{\min} = 1700$ m/s,  $f_{\max} = 2f_c = 250$  Hz and  $n = 12$  (table 2.2) we get:

$$dh \leq \frac{V_{\min}}{n f_{\max}} = 0.5 \text{ m.} \quad (4.1)$$

Therefore the spatial model grid consists of 320x520 gridpoints. To avoid a violation of the Courant criterion the time step size  $dt$  is calculated according to Eq. (2.50) using  $V_{\max} = 2000.0$  m/s,  $h = 1.0$  and  $dh = 0.5$  m

$$dt \leq \frac{dh}{h\sqrt{2}V_{\max}} = 0.17 \text{ ms.} \quad (4.2)$$

## CHAPTER 4. SIMPLE EXAMPLE: A SPHERICAL LOW VELOCITY ANOMALY

---

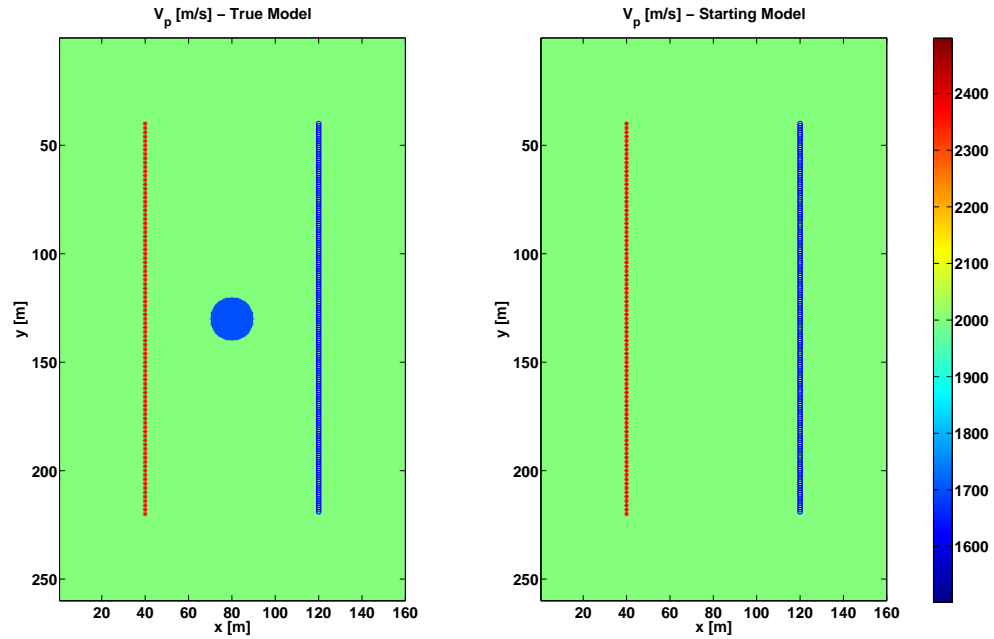


Figure 4.1: Simple acoustic test problem: A spherical low velocity anomaly in a homogeneous full space (left). The red stars denote the source positions, the blue circles the receiver positions. The starting model (right) consists of the homogeneous full space.

For the later application of the FWT algorithm it is important to apply a safety factor, to avoid a Courant instability when velocities larger than  $V_{\max} = 2000$  m/s occur during the model updates. In this case a time step of  $dt = 50 \mu\text{s}$  is used. The modeling should cover a time span of  $T = 65$  ms, so  $NT = 1300$  time steps are required. Fig. 4.2 shows the development of the pressure wavefield excited by shot 45 for the test problem. The wavefield is fairly simple. When the P-wave reaches the left half of the spherical inclusion a part of the energy is reflected and the wavefront is deformed due to the lower velocity in the inclusion (Fig. 4.2,  $T = 34$  ms). When the P-wave passes the transition between the sphere and the surrounding medium an even stronger reflection is excited at the right interface of the sphere (Fig. 4.2,  $T = 40$  ms).

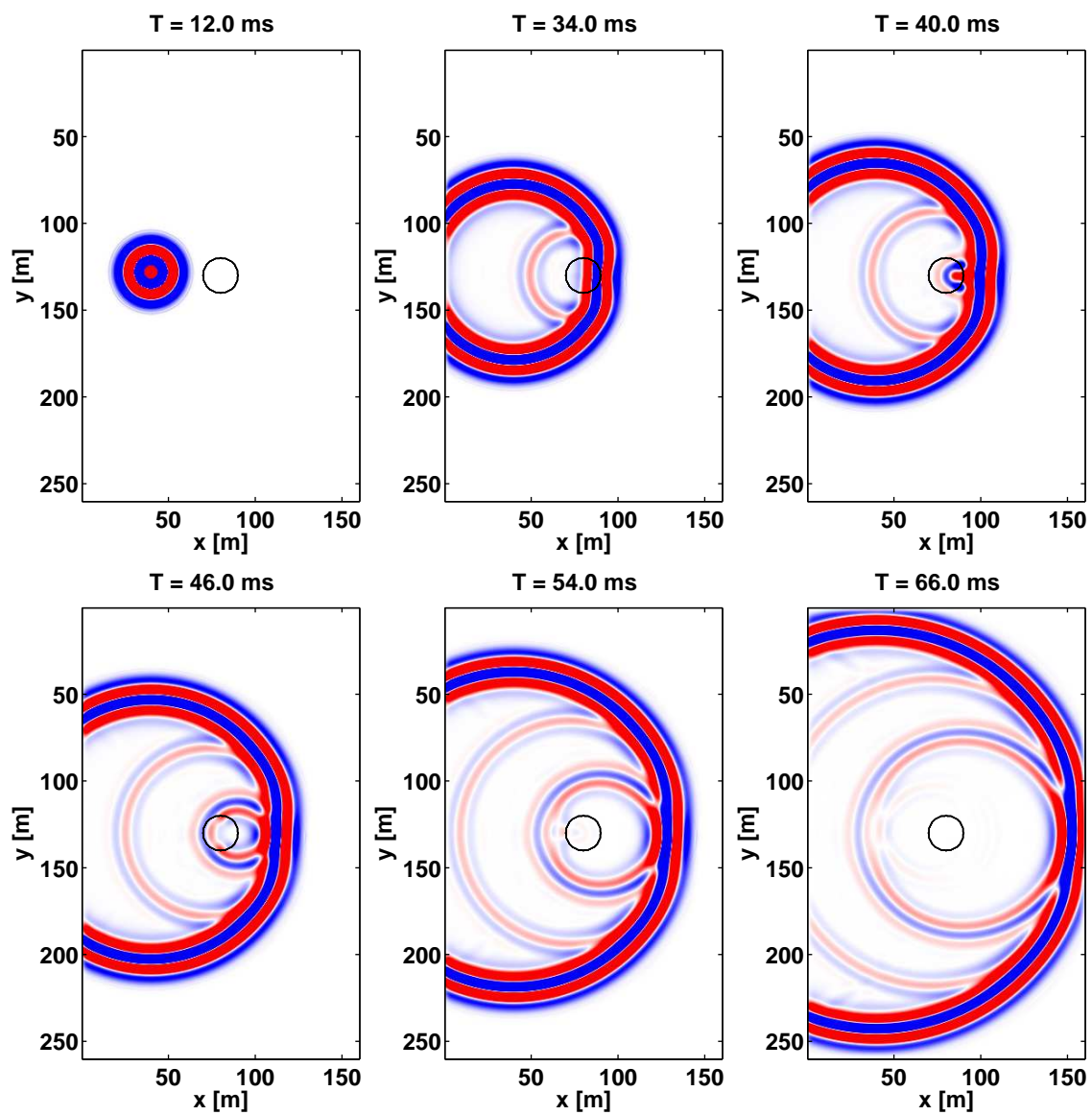


Figure 4.2: Pressure wavefield excited by shot 45 for the spherical low velocity anomaly model at 6 different time steps.

## CHAPTER 4. SIMPLE EXAMPLE: A SPHERICAL LOW VELOCITY ANOMALY

---

The starting model should represent the long wavelength part of the wavefield very well, otherwise the FWT code will not converge. For this simple example the starting model consists of the homogeneous background velocity model without the spherical inclusion (Fig. 4.1, right). The models for the S-wave velocity and density are known. Eventhough I always show the results in this chapter in terms of seismic P-wave velocity the inversion parameter is the Lamé parameter  $\lambda$ . After the definition of the starting point in the parameter space, the FWT algorithm can be applied.

1. At iteration step 1 do:
  - (a) For each shot solve the forward problem, stated in Eq.(3.16) for the actual model  $\mathbf{m}_0$  to generate a synthetic dataset  $\mathbf{u}^{\text{mod}}$  and the wavefield  $\mathbf{u}(\mathbf{x}, t)$ .
  - (b) Calculate the residual seismograms  $\delta\mathbf{u} = \mathbf{u}^{\text{mod}} - \mathbf{u}^{\text{obs}}$  for the x- and y-displacement components of the seismic data. This is demonstrated in Fig. 4.3 for the y-component of shot 45. The seismic section of the starting model  $\mathbf{u}_y^{\text{mod}}$  is shown on the left. You can easily see the undisturbed direct P-wave which propagates through the homogeneous starting model. The seismic section of the true model  $\mathbf{u}_y^{\text{obs}}$  is shown in the center. The wavefront of the direct wave is disturbed by the low velocity anomaly. The reflection from the right part of the sphere is also very evident. The initial data residuals  $\delta\mathbf{u}_y = \mathbf{u}_y^{\text{mod}} - \mathbf{u}_y^{\text{obs}}$  on the right show a quite large misfit.

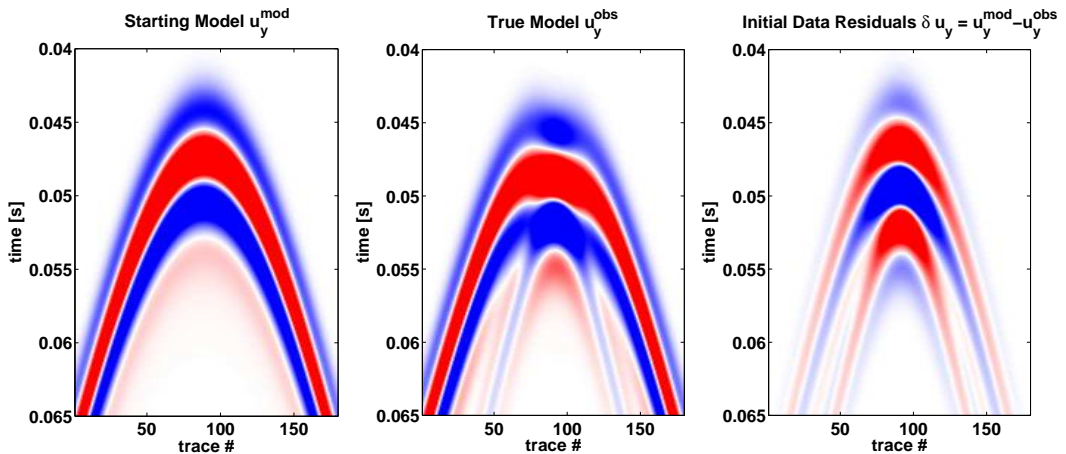


Figure 4.3: Seismic sections of the y-component for the simple test problem: The starting model (left), the true model (center) and the data residuals (right).

- 
- (c) Generate the wavefield  $\Psi(\mathbf{x}, t)$  by backpropagating the residuals from the receiver positions.
- (d) Calculate the gradients  $\delta\lambda'_0$  for the Lamé material parameter  $\lambda$  according to Eqs.(3.35). I simplified the wavefields in Eq. (3.35)

$$\delta\lambda' = - \sum_{\text{sources}} \int dt U \omega \quad (4.3)$$

by introducing the forward wavefield U

$$U = \left( \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \quad (4.4)$$

and the residual wavefield  $\omega$

$$\omega = \left( \frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y} \right) \quad (4.5)$$

The temporal evolution of the wavefields U and  $\omega$  for shot 45 is shown in Fig. 4.4 at corresponding times. The forward wavefield U consists of the direct P-wave, which simply propagates from the source point through the homogeneous starting model to the receivers (Fig. 4.4, top row). The residual wavefield  $\omega$  consists of the data residuals, backpropagated in time from the receiver positions into the homogeneous starting model (Fig. 4.4, bottom row). Note the focusing of the residual wavefield in the vicinity of the spherical anomaly. The gradient direction for each shot can be calculated using

$$\delta\lambda'(\text{shot}_i) = - \int dt U \omega \quad (4.6)$$

The results for the shots 2, 22, 40 and 82 are shown in Fig. 4.5. The gradients for the individual shots are shaped like Banana-Doughnut sensitivity kernels. The total gradient for the Lamé parameter  $\lambda$  is a superposition of the gradients for the individual shots

$$\delta\lambda' = \sum_{i=1}^{n_{\text{sources}}} \delta\lambda'(\text{shot}_i) \quad (4.7)$$

The gradient  $\delta\lambda'$  already shows the shape of the spherical inclusion (Fig. 4.5).

## CHAPTER 4. SIMPLE EXAMPLE: A SPHERICAL LOW VELOCITY ANOMALY

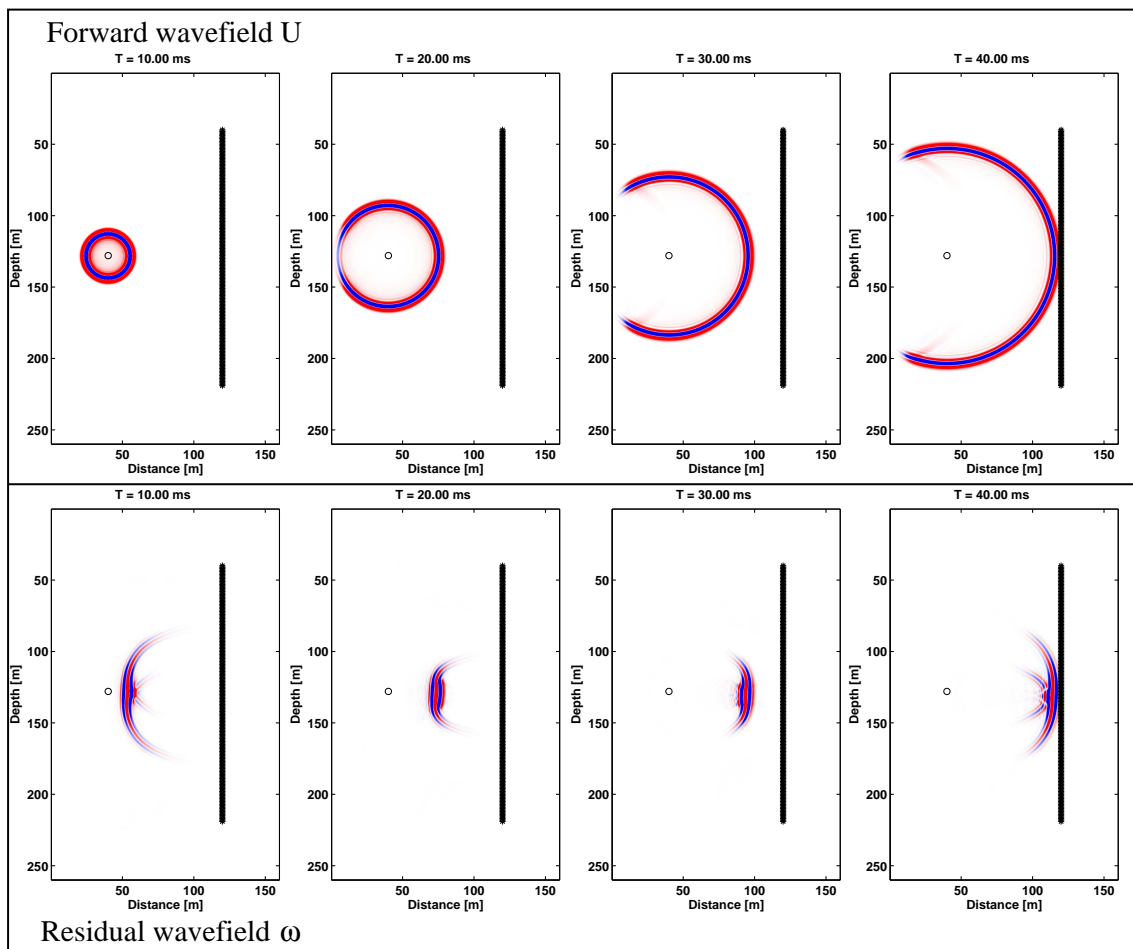


Figure 4.4: Development of the forward wavefield  $U$  (top row) and residual wavefield  $\omega$  (bottom row) at corresponding times.



$\delta\lambda$  – Gradient for ...

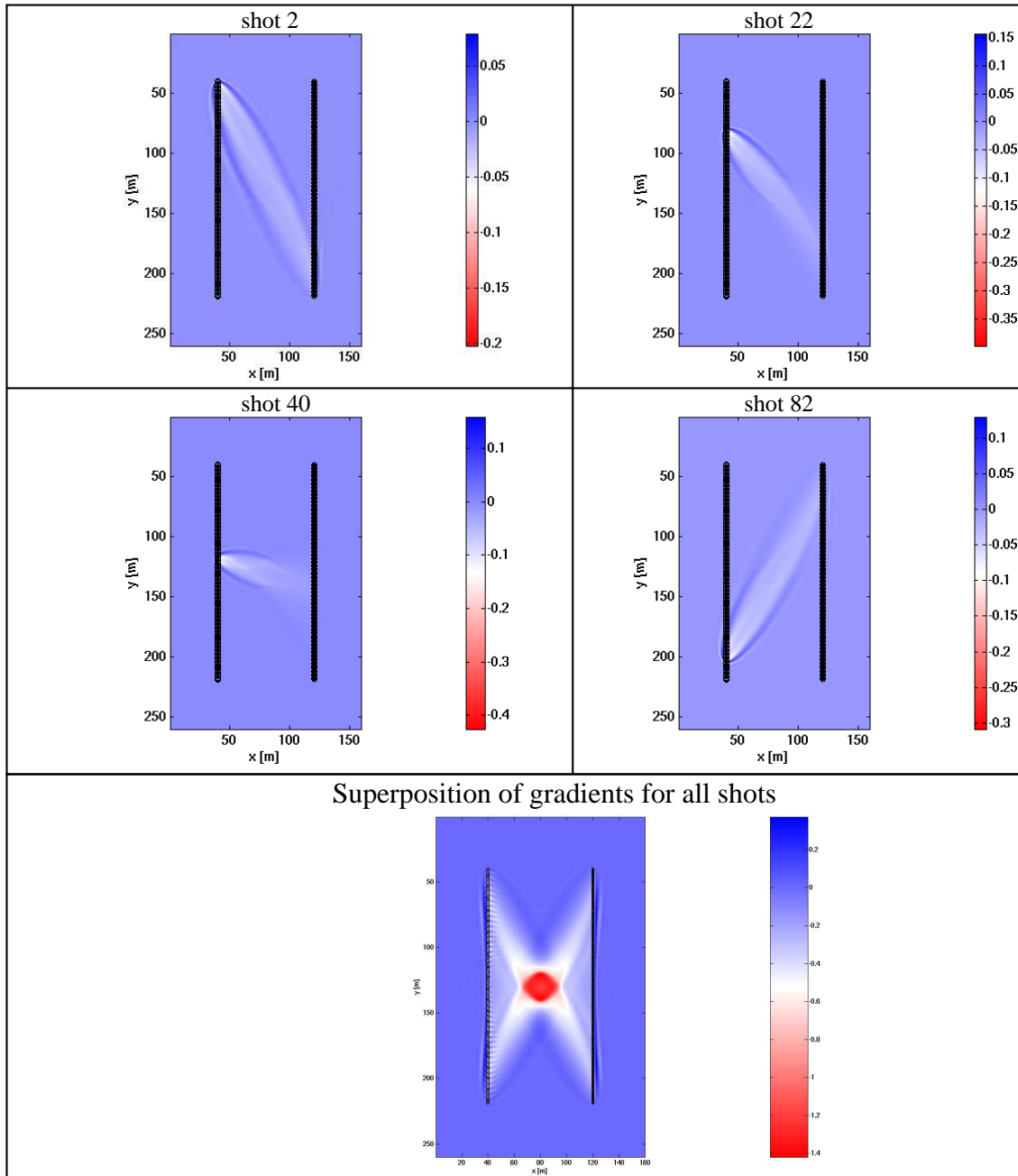


Figure 4.5: Gradients of the Lamé parameter  $\lambda$  for the shots 2, 22, 40 and 82. The total gradient is calculated by the superposition of all shots.

## CHAPTER 4. SIMPLE EXAMPLE: A SPHERICAL LOW VELOCITY ANOMALY

---

- (e) To increase the convergence speed an appropriate preconditioning operator  $P$  is applied to the gradient  $\delta\lambda_1$

$$\delta\lambda_1^P = P\delta\lambda_1 \quad (4.8)$$

Due to the low coverage and the absorbing boundary layer outside the cross-well geometry strong artefacts are present in this area (Fig. 4.7, left and center). Additionally the amplitude of the forward wavefield is very large near the source positions and is decreasing fast with larger offsets. Therefore the gradient is very large near the source positions and much weaker at larger offset. To suppress these effects a spatial Preconditioning operator is applied on the gradient, which is defined as (Fig. 4.6)

$$P = \begin{cases} 0 & \text{if } x_0 = 0.0 \text{ m} \leq x \leq x_{\text{gradt1}} = 32.5 \text{ m} \\ \exp\left(-\frac{1}{2}\left(a\frac{x-x_{\text{gradt1}}-\Delta l}{\Delta l/2}\right)^2\right) & \text{if } x_{\text{gradt1}} = 32.5 \text{ m} \leq x \leq x_{\text{gradt2}} = 42.5 \text{ m} \\ 1.0 & \text{if } x_{\text{gradt2}} = 42.5 \text{ m} \leq x \leq x_{\text{gradt3}} = 117.5 \text{ m} \\ \exp\left(-\frac{1}{2}\left(a\frac{x-x_{\text{gradt3}}}{\Delta l/2}\right)^2\right) & \text{if } x_{\text{gradt3}} = 117.5 \text{ m} \leq x \leq x_{\text{gradt4}} = 127.5 \text{ m} \\ 0 & \text{if } x_{\text{gradt4}} = 127.5 \text{ m} \leq x \leq x_n = 160 \text{ m} \end{cases}$$

with  $a = 3.0$ ,  $\Delta l = 10.0$  m and  $x_{\text{fw}} = x_{\text{gradt4}} - x_{\text{gradt1}}$ . The preconditioning operator does not change the gradient in the area of large ray coverage between sources and receivers, while it sets the gradient to zero in areas where the ray coverage is low. In the transition zone a Gaussian taper is applied.

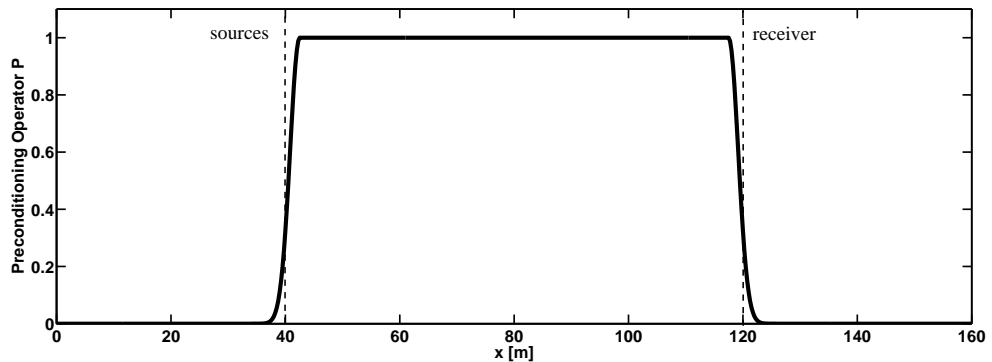


Figure 4.6: The preconditioning operator for the cross-well acquisition geometry.

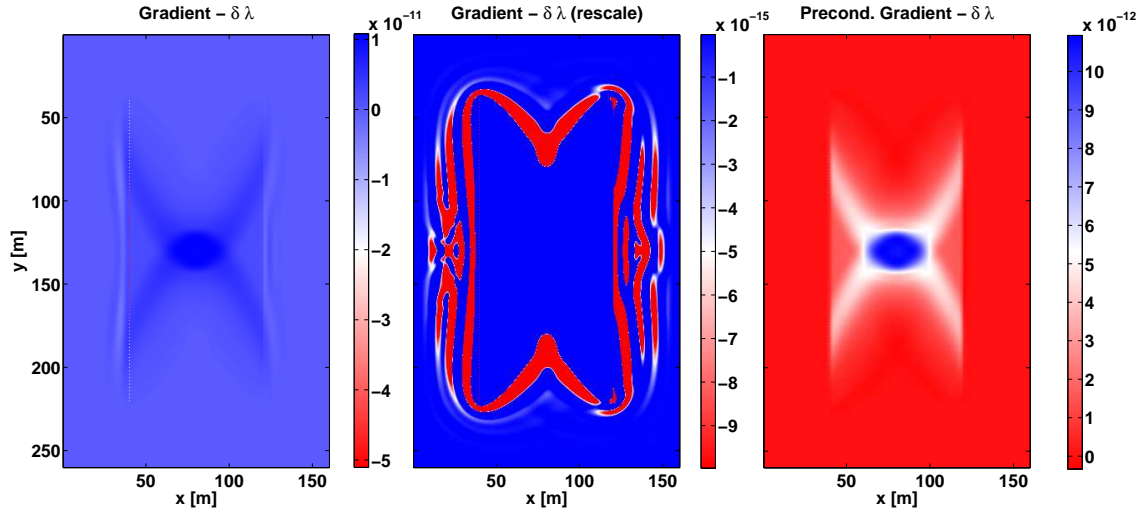


Figure 4.7: The effect of the preconditioning operator  $P$  as defined in the main text. The Gradient  $\delta\lambda$  before (left) and after the application of the preconditioning operator (right). The artefacts due to the low ray coverage and the absorbing boundary outside the cross-well acquisition geometry are shown in a rescaled image of the unpreconditioned gradient (center).

The effect of the preconditioning operator is shown in Fig. 4.7 before (left) and after its application (right). The artefacts are strongly reduced and the spherical low velocity anomaly is much more dominant. Notice that there is still a strong X shaped artefact surrounding the spherical inclusion, which can be explained by the strong ray coverage with similar azimuth or ray directions at the outer receiver/source positions (De Nil (2008), personal communication). During the FWT iterations these artefacts are reduced, but a more sophisticated preconditioning operator may suppress the artefacts and increase the convergence speed.

- (f) Estimate the step length  $\mu_1$  by the line search method as described in chapter 3.4 .
- (g) Update the material parameters using the gradient method

$$\mathbf{m}_2 = \mathbf{m}_1 - \mu_1 \delta\lambda_1^P. \quad (4.9)$$

For all iterations  $n > 1$  apply the conjugate gradient algorithm (chapter 3.5).

## CHAPTER 4. SIMPLE EXAMPLE: A SPHERICAL LOW VELOCITY ANOMALY

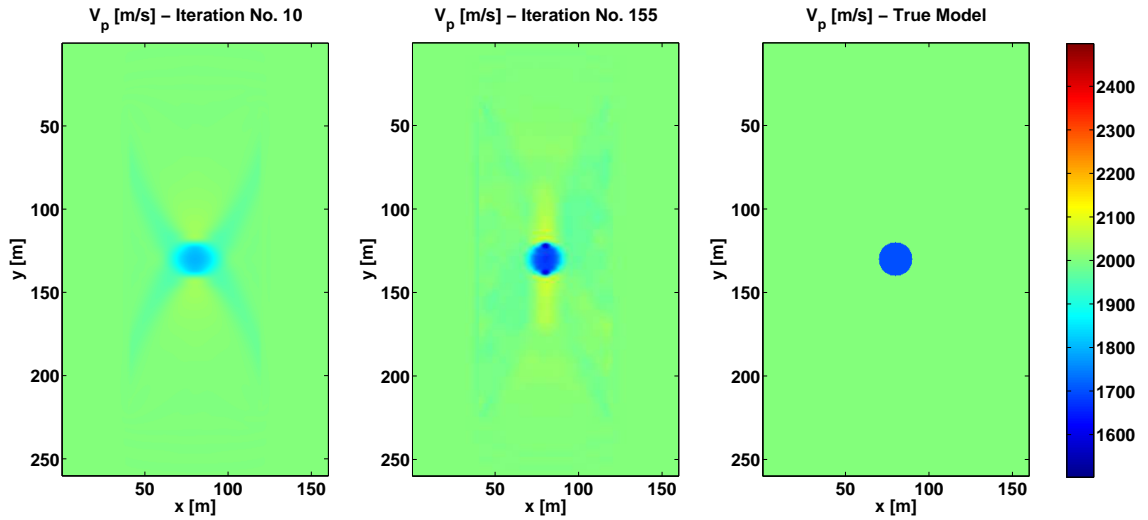


Figure 4.8: Inversion results for the P-wave velocity model of the spherical low velocity anomaly after 10 (left) and 155 FWT iterations (center) compared with the true model (right).

The inversion results for the P-wave velocity after 10 and 155 iterations are compared in Fig. 4.8 (left and center) with the true P-wave velocity model (right). The position and material parameters of the spherical low velocity anomaly could be recovered very well. The X shaped artefact surrounding the sphere is very obvious after ten iterations, but strongly reduced after 155 iteration steps. Eventhough some diffuse artefacts are still present. The seismic section ( $y$ -component) for the inversion result is shown in Fig. 4.9 together with the seismic section of the true model and the final data residuals. The deformed wavefront of the direct P-wave and the diffracted P-wave could be fitted very well. The data residuals are strongly reduced compared with the initial data residuals (Fig. 4.2). This can be verified by the development of the  $L_2$  norm of the data residuals as a function of the iteration step (Fig. 4.10). Up to iteration step 20 the norm is decreasing very fast to a value of 2.45 % of the initial data residuals. Afterwards the  $L_2$  norm shows strong fluctuations. Even though these fluctuations are strong in terms of the  $L_2$  the model changes are quite small. This demonstrates the strong nonlinear character of the elastic inversion problem, even if the problem seem to be quite simple.

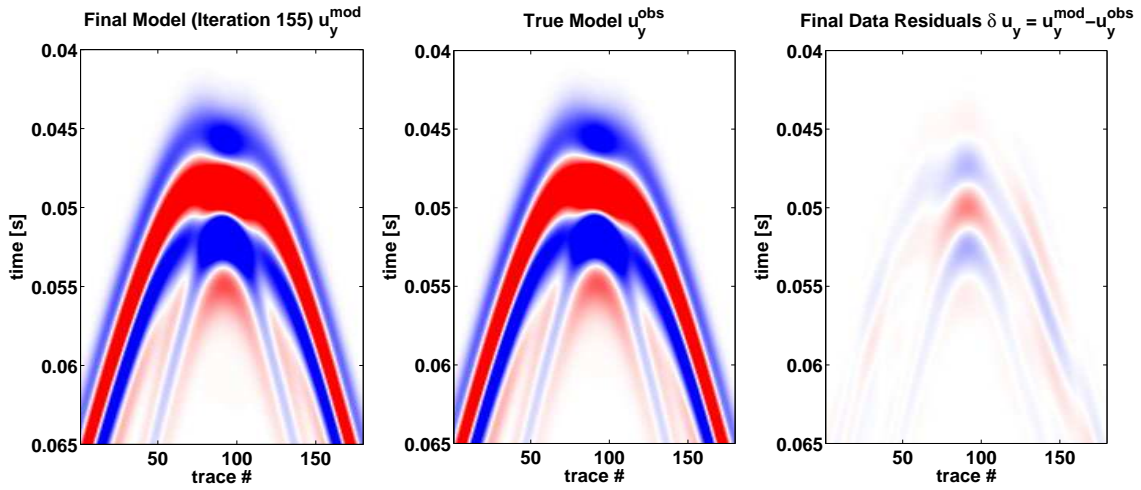


Figure 4.9: Seismic sections (y-component) for the inversion result (left), the true model (center) and the data residuals (right).

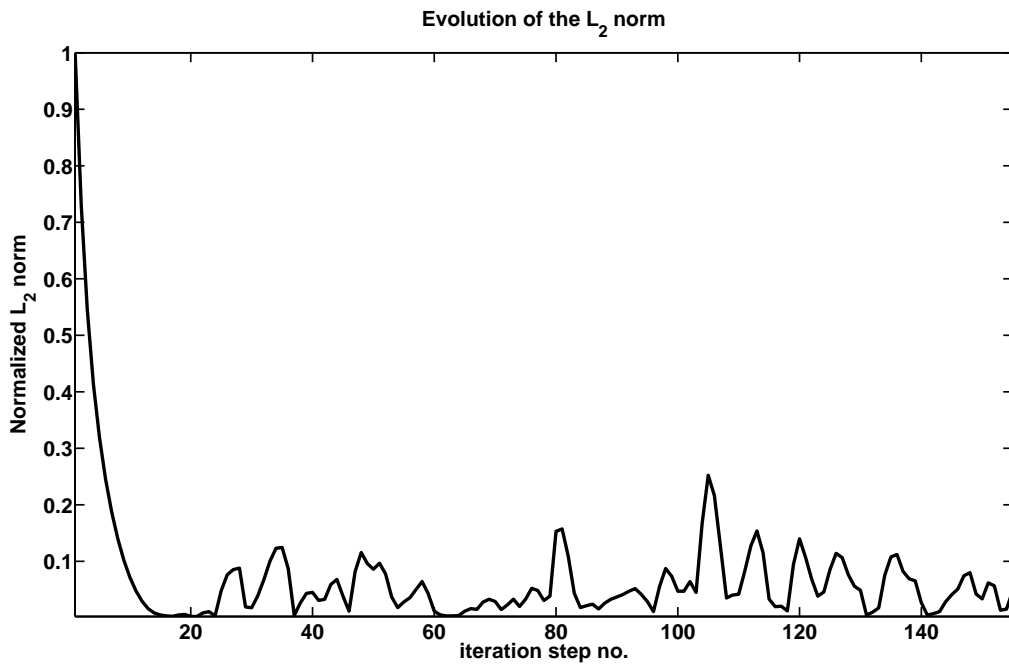


Figure 4.10: Evolution of the  $L_2$ -norm of the data residuals.



# 5

## The influence of model parametrization

As noted by many authors, e.g. Tarantola (1986) and Igel et al. (1993) the model parametrization has a major impact on the resolution and ambiguity of the FWT result. In this chapter the influence of different model parametrizations is studied in more detail.

### 5.1 The gradient direction $\frac{\partial \mathbf{E}}{\partial \mathbf{m}}$ in terms of different model parametrizations

---

The gradients in terms of other material parameters  $\mathbf{m}_{\text{new}}$  can be calculated by applying the chain rule on the Frechét kernel in the adjoint problem (Eq. (3.26)):

$$\delta \mathbf{m}_{\text{new}} = \sum_{\text{sources}} \int dt \sum_{\mathbf{R}} \left[ \frac{\partial \mathbf{u}}{\partial \mathbf{m}} \frac{\partial \mathbf{m}}{\partial \mathbf{m}_{\text{new}}} \right]^* \delta \mathbf{u} \quad (5.1)$$

Using the relationships between P-wave velocity  $V_p$ , S-wave velocity  $V_s$ , the Lamé parameters  $\lambda$ ,  $\mu$  and density  $\rho$ :

$$V_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \quad V_s = \sqrt{\frac{\mu}{\rho}} \quad (5.2)$$

or

$$\lambda = \rho V_p^2 - 2\rho V_s^2, \quad \mu = \rho V_s^2. \quad (5.3)$$

## CHAPTER 5. THE INFLUENCE OF MODEL PARAMETRIZATION

---

The gradient for  $V_p$  can be written as:

$$\begin{aligned}
 \delta V_p &= \sum_{\text{sources}} \int dt \sum_R \left[ \frac{\partial \mathbf{u}}{\partial \lambda} \frac{\partial \lambda}{\partial V_p} + \frac{\partial \mathbf{u}}{\partial \mu} \frac{\partial \mu}{\partial V_p} + \frac{\partial \mathbf{u}}{\partial \rho} \frac{\partial \rho}{\partial V_p} \right]^* \delta u_i \\
 &= \sum_{\text{sources}} \int dt \sum_R \left[ \frac{\partial \mathbf{u}}{\partial \lambda} 2\rho V_p \right]^* \delta u_i \\
 &= 2\rho V_p \sum_{\text{sources}} \int dt \sum_R \left[ \frac{\partial \mathbf{u}}{\partial \lambda} \right]^* \delta u_i \\
 &= 2\rho V_p \delta \lambda
 \end{aligned} \tag{5.4}$$

The gradients for  $V_s$  and  $\rho$  are calculated in a similar way, so the gradients in terms of seismic velocities can be written as:

$$\begin{aligned}
 \delta V_p &= 2\rho V_p \delta \lambda, \\
 \delta V_s &= -4\rho V_s \delta \lambda + 2\rho V_s \delta \mu, \\
 \delta \rho_{\text{vel}} &= (V_p^2 - 2V_s^2) \delta \lambda + V_s^2 \delta \mu + \delta \rho
 \end{aligned} \tag{5.5}$$

### 5.2 The CTS test problem

---

A simple model parameter analysis by Tarantola was based on the diffraction pattern of point-like diffractors (Tarantola (1986)). Here I present a more complex elastic test problem which was developed during an extensive and fruitful discussion with De Nil (2009). To investigate the influence of different model parametrizations we build two elastic models for the parameter sets  $\mathbf{m}_1 = [\lambda, \mu, \rho]$  and  $\mathbf{m}_2 = [V_p, V_s, \rho]$  (Fig. 5.1). The models consist of a free surface at the top, an elastic layer and a half space. Seismic body waves are traveling from the sources at the surface and are reflected back at the interface between the layer and half space. Embedded in the layer are different geometrical bodies, which are disturbing the wavefield of the reflected waves. These geometrical bodies consist of

1. 7 crosses indicating perturbations of the Lamé parameter  $\lambda$  or the P-wave velocity  $V_p$ .
2. 8 triangles indicating perturbations of the Lamé parameter  $\mu$  or the S-wave velocity  $V_s$ .



## 5.2. THE CTS TEST PROBLEM

---

3. 6 squares indicating perturbations of the density  $\rho$ .

Due to the different geometrical structures we call this model **Cross-Triangle-Squares** (CTS) model. The geometrical bodies are located at different non overlapping places. This does not represent a realistic geological situation, but it is an effective way to demonstrate the resolution and ambiguity of the FWT result when using different elastic parametrizations. The S-wave velocity  $V_s$  and density  $\rho$  for the different geometrical structures are calculated from the P-wave velocity  $V_p$  of the crosses using the following relationships

$$\begin{aligned} V_s &= V_p / \sqrt{3}, \\ \rho &= 0.31 * 1000.0 * V_p^{1/4}. \end{aligned} \tag{5.6}$$

The corresponding models for the Lamé parameters are calculated using the relationships in Eq. (5.3), but without mixing the structural models. Therefore the resulting models for seismic velocities and Lamé parameters are not equivalent and the resulting wavefields are different, which can be seen easily by comparing the seismic sections for the different parametrizations in Fig. 5.3 (c). The acquisition geometry consists of 100 explosive sources 40 m below the free surface. The source signature is a Ricker wavelet with a centre frequency of 5 Hz and a maximum frequency of 10 Hz. The elastic wavefield is recorded by 400 two component receivers in 40 m depth. Using an 8th order spatial FD operator for the forward modelling and backpropagation of the residual wavefield the model can be discretized with  $500 \times 150$  gridpoints in x- and y-direction with a spatial gridpoint distance of 20.0 m. The time is discretized using  $DT = 2.7$  ms, thus for a recording time of  $T = 6.0$  s 2222 time steps are required. Synthetic multicomponent datasets are calculated for the CTS model and inverted using a starting model with the correct elastic material parameters for the layer and the half space but without the geometrical structures. In Fig. 5.2 the inversion results are shown using the Lamé parameters and the seismic velocities as elastic inversion parameters. In both cases the elastic parameters could be reconstructed very well without any ambiguity. The Lamé parameter  $\lambda$  shows a little bit more artefacts than the P-wave velocity  $V_p$ , but the results are still quite similar in terms of resolution. Surprisingly the choice of parameters has a huge impact on the density results. Using Lamé parameters the squares of the density model could be recovered very well, but they are disturbed by extremely large triangular shaped  $\mu$  artefacts which would hide the true density result in a geological more realistic setting. When using seismic velocities as model parameters a stronger ambiguity is present, the crosses of the  $V_p$  model and the triangles of the  $V_s$  model are partly interpreted as density information, but they have the same magnitude

## CHAPTER 5. THE INFLUENCE OF MODEL PARAMETRIZATION

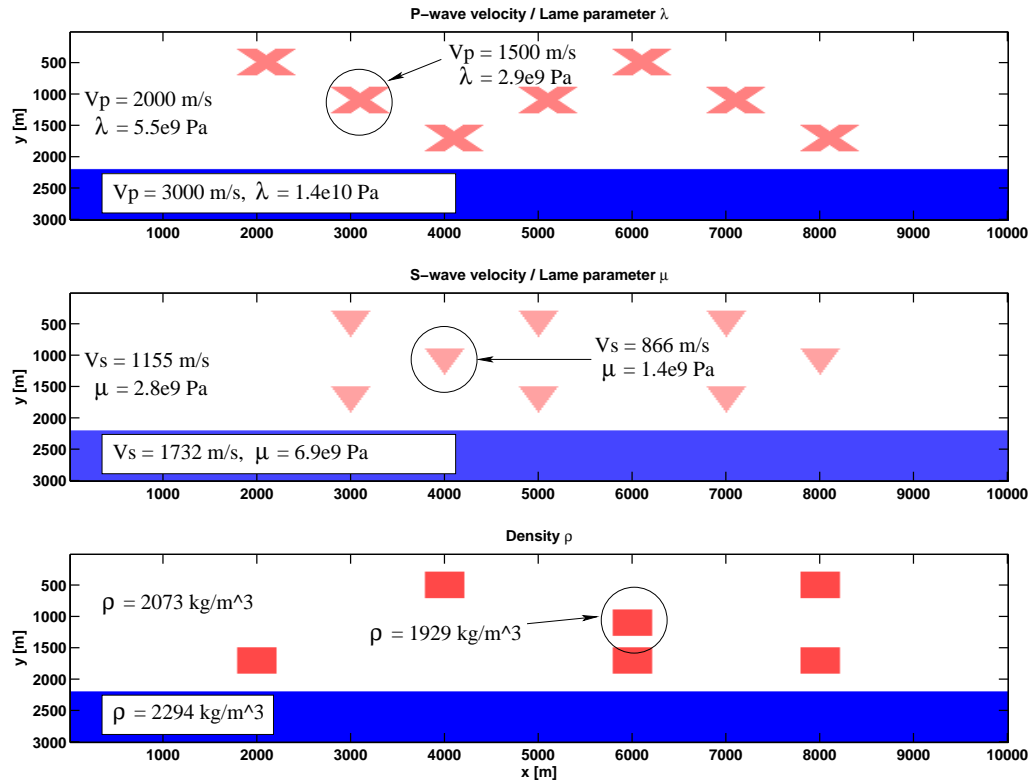


Figure 5.1: The Cross-Triangle-Squares test problem for the parameter sets  $\mathbf{m}_1 = [\lambda, \mu, \rho]$  and  $\mathbf{m}_2 = [V_p, V_s, \rho]$ .

as the true density model. In Fig. 5.3 the seismic sections of shot 50 are plotted for the starting model (a), the inversion result (b) the true model (c), the initial residuals (d), the final residuals (f) and the evolution of the residual energy (e) using the different parametrizations. The fit of the phases and amplitudes is very good in both cases. Even though the final residuals are a bit larger in case of the seismic velocities (Fig. 5.3 (e)).

## 5.2. THE CTS TEST PROBLEM

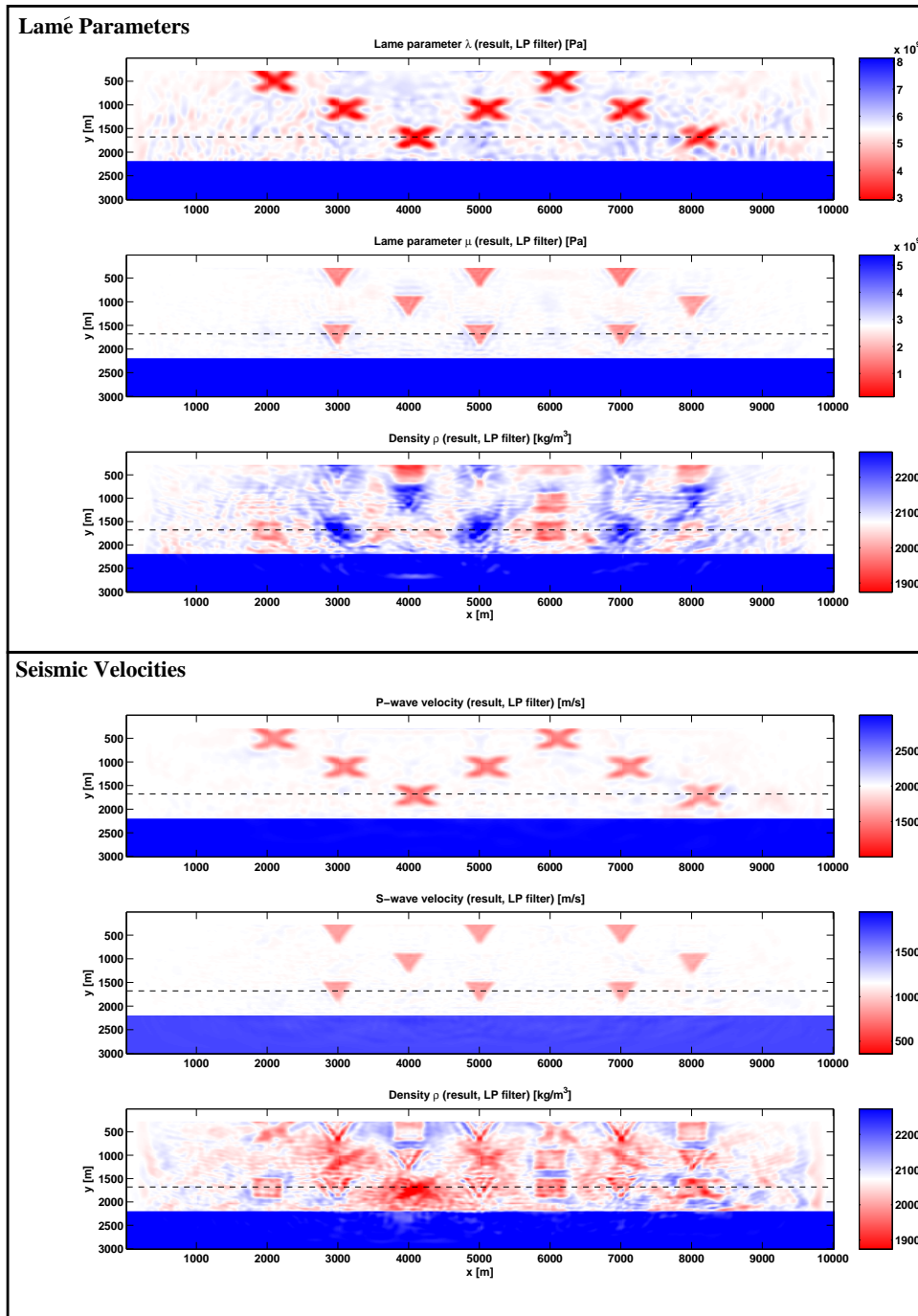


Figure 5.2: Results of the FWT for the Cross-Triangle-Squares model using Lamé parameters (top) and seismic velocities (bottom) as inversion parameters.

## CHAPTER 5. THE INFLUENCE OF MODEL PARAMETRIZATION

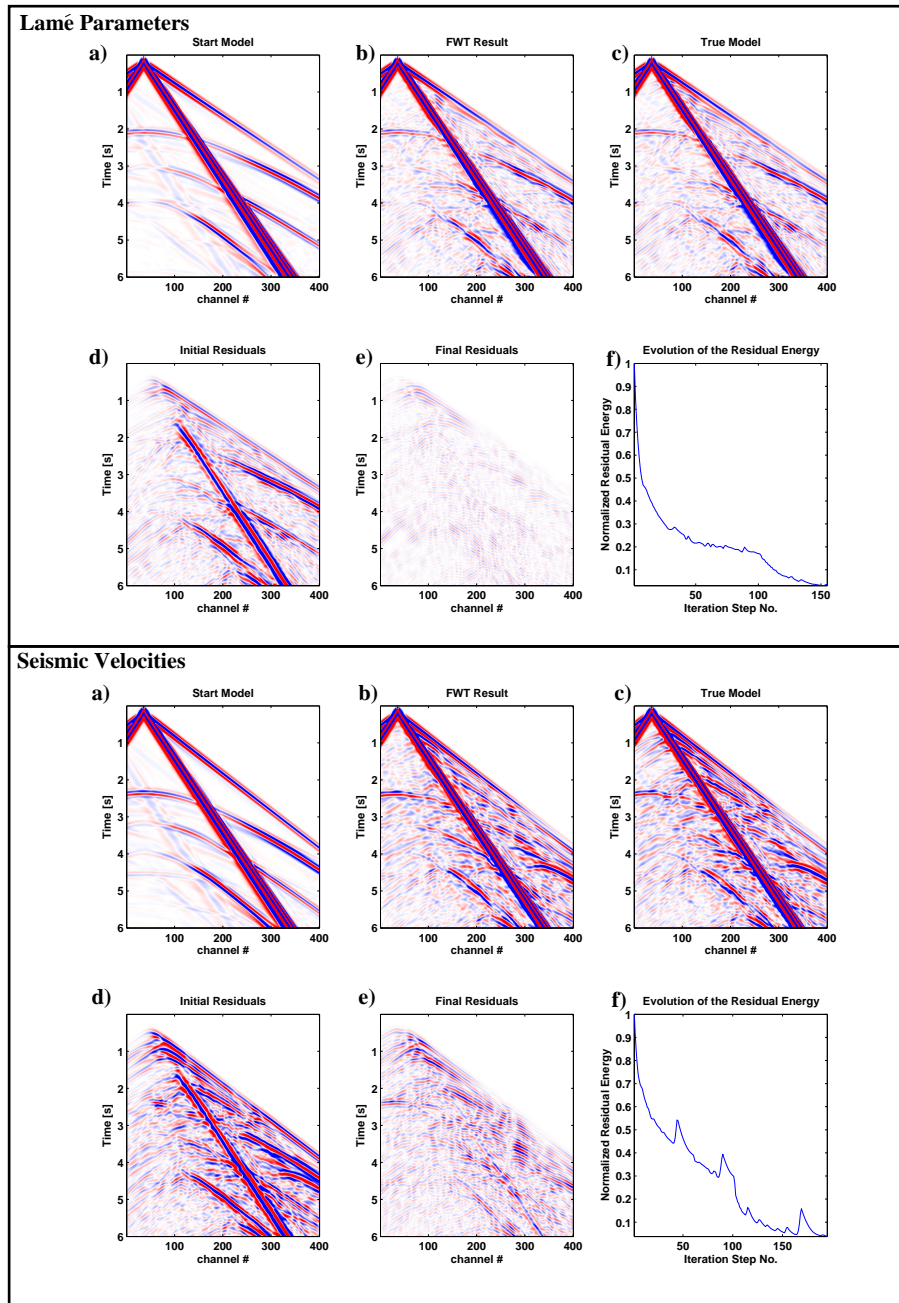


Figure 5.3: Seismic sections for the Cross-Triangle-Squares model using Lamé parameters (top) and seismic velocities (bottom). (a) starting model, (b) FWT result, (c) true model, (d) initial residuals, (e) final residuals and (f) evolution of the residual energy.

# 6

## A geological test problem - the elastic Marmousi2 model

Developed in the 1990s by the French Petroleum Institute (IFP) (Versteeg (1994)) the Marmousi model is a widely used test problem for seismic imaging techniques. Beside the original acoustic version of the model an elastic version was developed by Martin et al. (2006). This model consists of parts with simple (approximately 1D) and complex geological situations. In the following two sections the performance of the FWT code will be tested for the complex and the simple part of the Marmousi2 model, respectively.

### 6.1 The complex Marmousi2 model

---

The Marmousi2 model (Fig. 6.1) consists of a 500 m thick water layer above an elastic subseafloor model. The sediment model is very simple near the left and right boundaries but rather complex in the centre. At both sides, the subseafloor is approximately horizontally layered, while steep thrust faults are disturbing the layers in the centre of the model. Embedded in the thrust fault system and layers are small hydrocarbon reservoirs (Fig. 6.1, Martin et al. (2006)).

- One shallow gas sand in a simple structural area (A).
- One relatively shallow oil sand in a structural simple area (B).
- Four faulted trap gas sands at varying depths (C1,C2,C3,C4).
- Two faulted trap oil sands at medium to deep depths (D1,D2).

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

---

- One deep oil and gas sand anticlinal trap (E1,E2).
- Water wet sand.

The deeper parts of the model consist of salt and reef structures. The thrust fault system and the reef structures are not easy to resolve by conventional first arrival tomography, so it is an ideal test model for the FWT. Due to computational restrictions the original Marmousi-II model could not be used, because the very low S-wave velocities in the sediments would require a too small spatial sampling of the model. Therefore new S-wave velocities are calculated using the empirical scaling relation (5.6) for hard rocks. Additionally the size of the Marmousi2 model is reduced from  $17 \times 3.5$  km to  $10 \times 3.48$  km (Fig. 6.2).

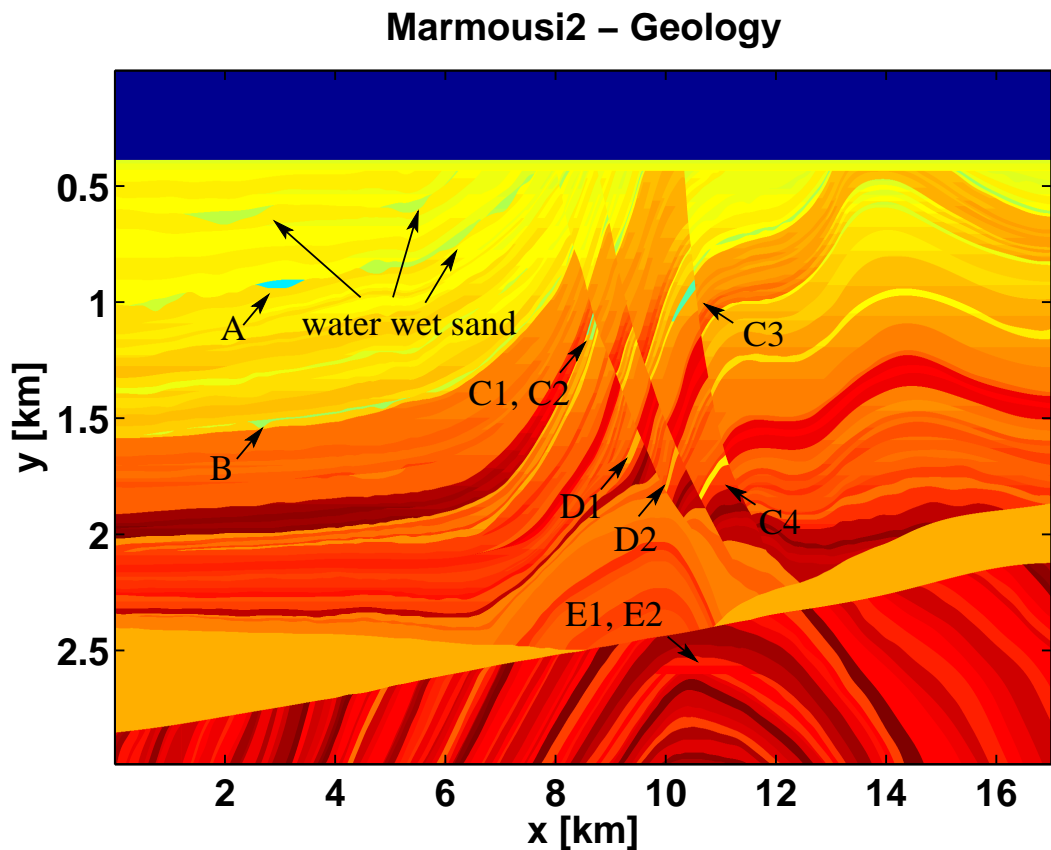


Figure 6.1: Marmousi2 model - geology.

## 6.1. THE COMPLEX MARMOUSI2 MODEL

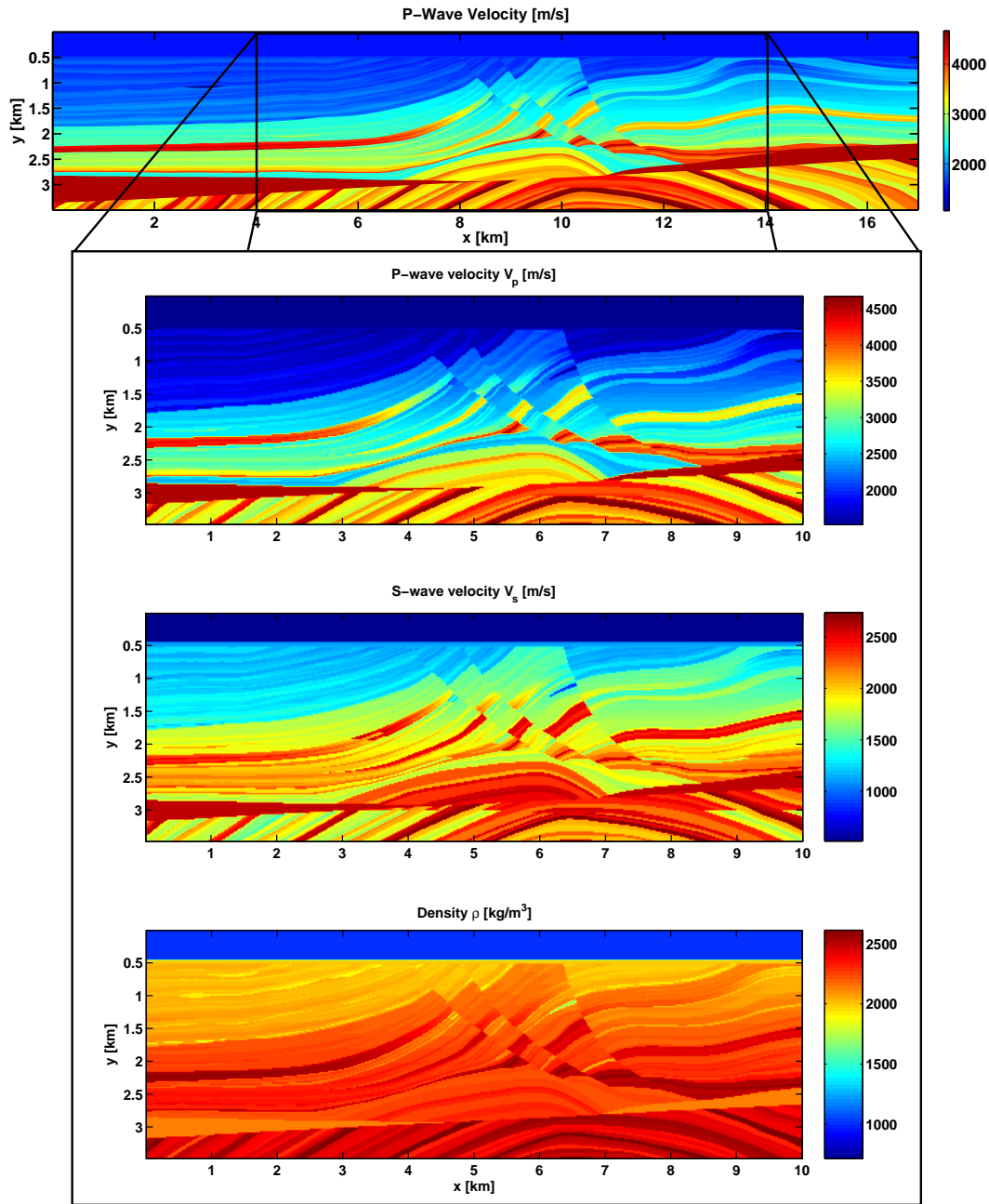


Figure 6.2: The reduced complex Marmousi2 model used for the elastic FWT.

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

---

### 6.1.1 Acquisition geometry and FD model

The acquisition geometry consists of a fixed streamer located 40 m below the free surface in the water layer. The streamer contains 400 two component geophones recording the displacements  $u_i$ . For the synthetic dataset 100 airgun shots are recorded. The sources are located at the same depth as the receivers. The source signature is a 10 Hz Ricker wavelet. The model has the dimensions  $10 \times 3.48$  km. Using an 8th order spatial FD operator the model can be discretized with  $500 \times 174$  gridpoints in x- and y-direction with a spatial gridpoint distance of 20.0 m. The time is discretized using  $DT = 2.7$  ms, thus for a recording time of  $T = 6.0$  s 2222 time steps are needed.

### 6.1.2 Elastic wave propagation in the complex Marmousi model

Fig. 6.3 shows the development of the pressure wavefield excited by shot 50 for the central part of the complex elastic Marmousi2 model at 6 different time steps. The P wave is traveling from the source through the water column ( $T=100.0$  ms) and is reflected at the seafloor ( $T=400.0$  ms). In the elastic subseafloor medium the wavefield becomes very complex. The layers in the steep thrust fault system produce numerous reflections and internal multiples ( $T=600.0$  ms). Additionally strong diffracted waves are generated at the sharp corners of the thrust faults between the disturbed high velocity sediment blocks within the thrust faults and the surrounding low velocity sediments. At the free surface strong multiple reflections occur ( $T=800.0$  ms). The wavefront of the direct wave is quite deformed due to strong velocity contrasts within the thrust fault system. After 1500 ms nearly all kinds of waves which can be found in the literature are present: Reflections, refractions, diffractions, (internal) multiples or interface waves. The trapped gas sand reservoirs C1, C2 and C3 produce strong reflections and mode conversions. This complexity is also visible in the seismic section, recorded by the streamer in the water column. As an example Fig. 6.9, (f) shows the seismic section of the y-component for shot 50. Beside the direct wave and a strong reflection from the seafloor numerous small reflection events from the thrust fault system are dominating the seismic section.



## 6.1. THE COMPLEX MARMOUSI2 MODEL

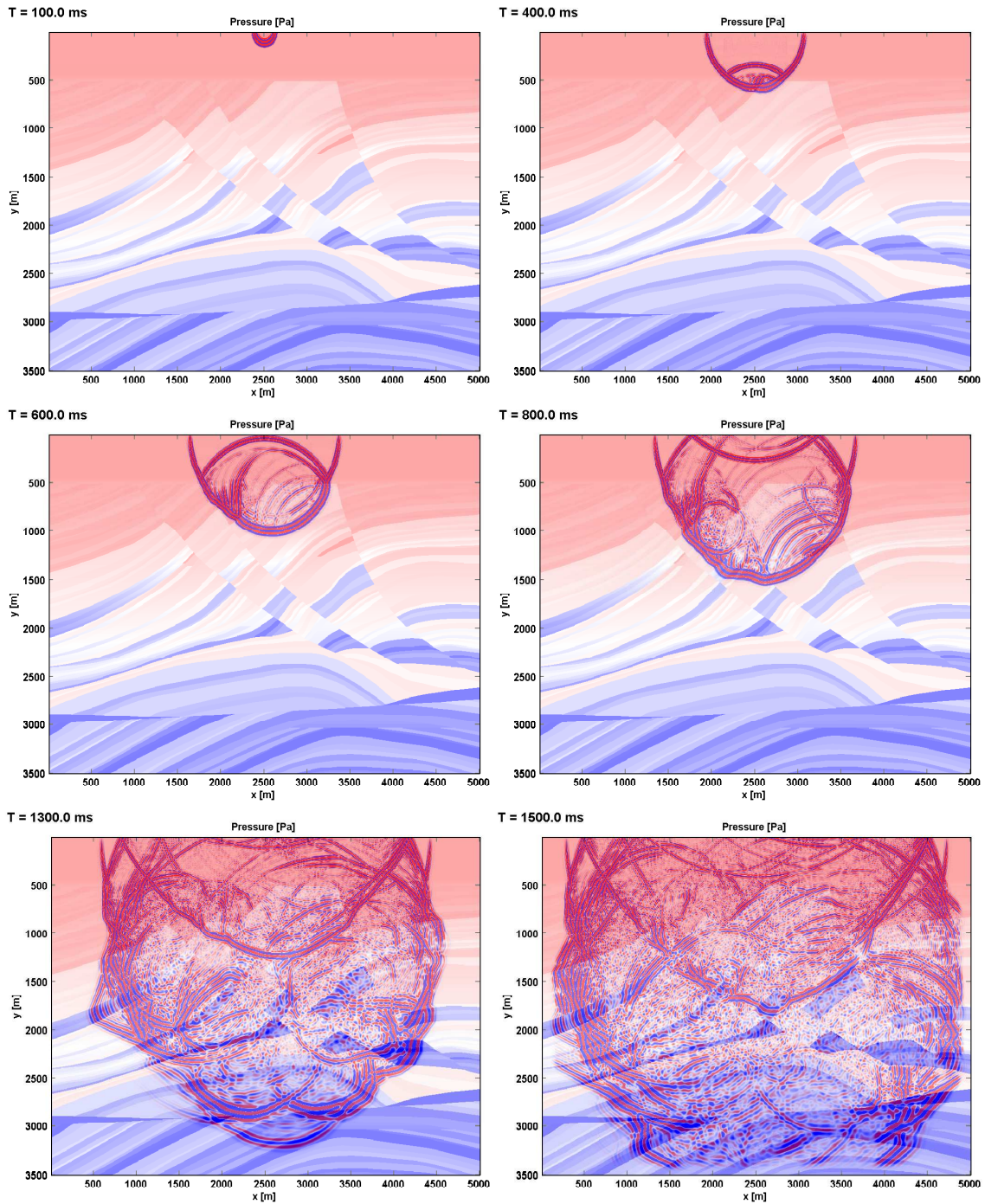


Figure 6.3: Pressure wavefield excited by shot 50 for the elastic Marmousi2 model at 6 different time steps .

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

### 6.1.3 FWT of the complex Marmousi model

Due to the results of the last section, I choose the seismic velocities as model parameters for the inversion. To generate a starting model which describes the long wavelength part of the material parameters correctly the true model  $\mathbf{m} = [V_p, V_s, \rho]$  was filtered using a spatial 2D-Gaussian filter

$$\mathbf{m}_{\text{smooth}}(x, y) = \frac{1}{2\pi\lambda_c^2} \int_{-\lambda_c}^{\lambda_c} \int_{-\lambda_c}^{\lambda_c} dx' dy' \mathbf{m}(x - x', y - y') \exp\left(-\frac{(x - x')^2 + (y - y')^2}{2\lambda_c^2}\right) \quad (6.1)$$

with a correlation length  $\lambda_c = 200.0$  m. As a result all the small scale structures vanished and only the large scale structures are present (Fig. 6.4).

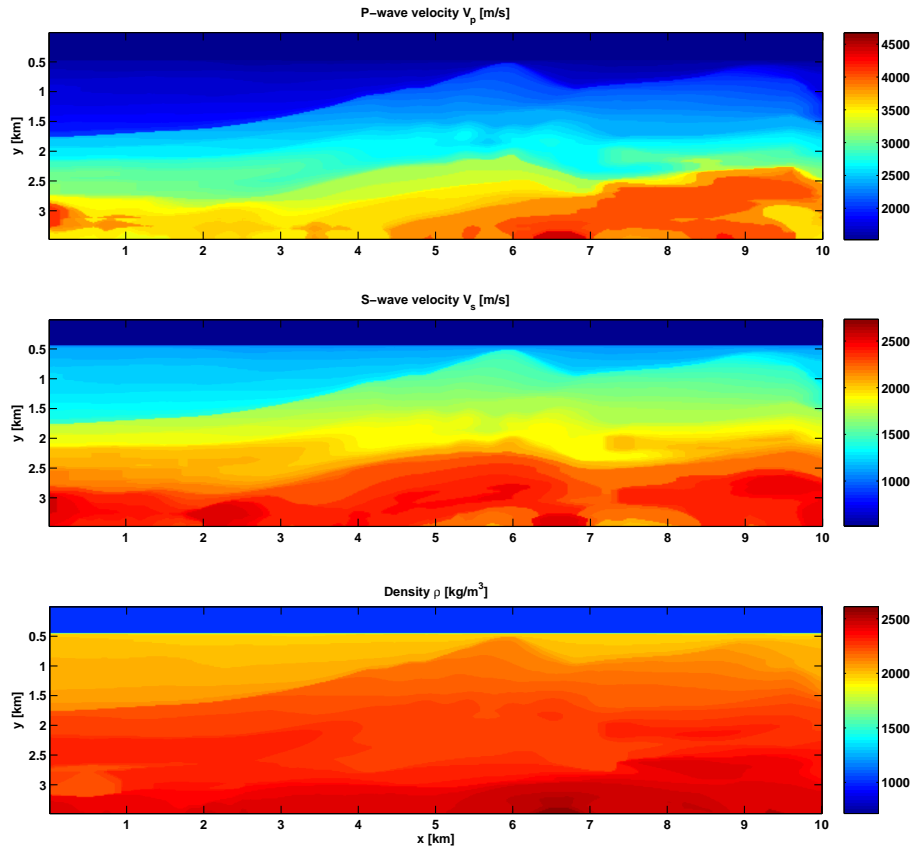


Figure 6.4: Starting models for the Marmousi-II model.

## 6.1. THE COMPLEX MARMOUSI2 MODEL

As in the case of the spherical low velocity anomaly the application of a preconditioning operator is vital to suppress the large gradient values near the source and receiver positions. Additionally strong artefacts are present near the free surface (Fig. 6.6, top) which are a few orders of magnitude larger than the gradient of the material parameters. This problem is also known in case of the acoustic inversion problem (Ben-Hadj Ali et al. (2008)). To suppress these effects a spatial preconditioning operator is applied on the gradient, which is defined as (Fig. 6.5)

$$P = \begin{cases} 0 & \text{if } y_0 = 0.0 \text{ m} \leq y \leq y_{\text{gradt1}} = 380.0 \text{ m} \\ \exp\left(-\frac{1}{2}\left(a\frac{y-y_{\text{gradt1}}-\Delta l}{\Delta l/2}\right)^2\right) & \text{if } y_{\text{gradt1}} = 380.0 \text{ m} \leq y \leq y_{\text{gradt2}} = 480.0 \text{ m} \\ \frac{y}{y_{\text{gradt2}}} & \text{if } y_{\text{gradt2}} = 480.0 \text{ m} < y \end{cases} \quad (6.2)$$

with  $a = 3.0$ ,  $\Delta l = 100.0$  m.

The preconditioning operator sets the gradient near the free surface and the

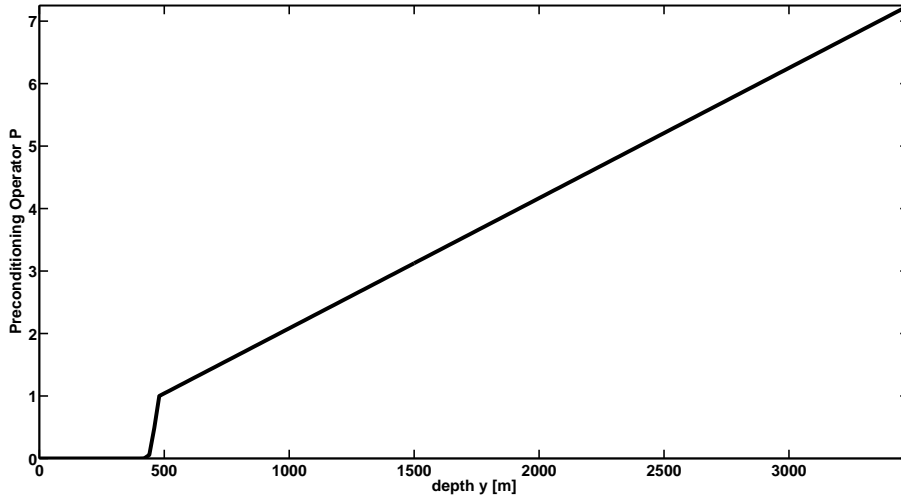


Figure 6.5: The preconditioning operator for the reflection geometry.

sources/receivers to zero. In a transition zone between 380.0 m and 480.0 m depth a Gaussian taper is applied. Beyond a depth of 480 m the operator scales the gradient linear with depth. This is a very crude correction for the amplitude loss in larger depths due to geometrical spreading and reflections in the upper parts of the model. Before the application of the preconditioning operator no subsurface structures are visible at all Fig. 6.6 (top), while strong reflectors are visible after its application.

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

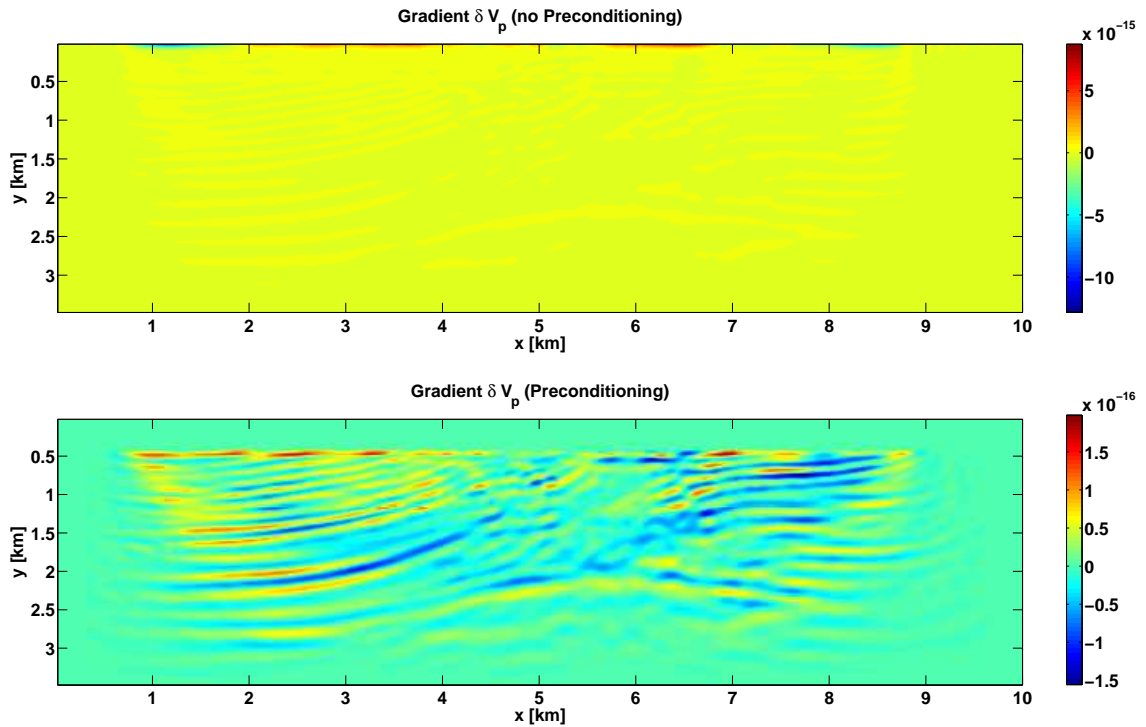


Figure 6.6: The influence of the preconditioning operator  $P$  for the elastic Marmousi2 model. The Gradient  $\delta V_p$  (iteration no. 1) before (top) and after the application of the preconditioning operator (bottom).

To achieve a smooth transition from the long wavelength starting model to the inversion result with short wavelength structures the application of a frequency filter with variable bandwidth on the data residuals  $\delta \mathbf{u}$  is vital, to avoid the convergence into a local minimum (see chapter 9). In this case the inversion is separated in two parts. In part I only frequencies below 10 Hz are inverted, while in part II the full spectral content up to 20 Hz is inverted. The inversion results after 350 iterations are shown in Fig. 6.7. Additionally depth profiles at  $x_{p1} = 3.5$  km and  $x_{p2} = 6.4$  km of the starting model and inversion result are compared with the true model in Fig. 6.8. The results contain a lot of small details. All fine layers which are completely absent in the starting model could be resolved. The thrust faults and the reef structures in the deeper part of the model are imaged also very well. Despite the deep hydro carbon reservoirs E1 and E2, the oil and gas reservoirs C1, C2, C3, C4, D1 and D2 are all clearly visible.

## 6.1. THE COMPLEX MARMOUSI2 MODEL

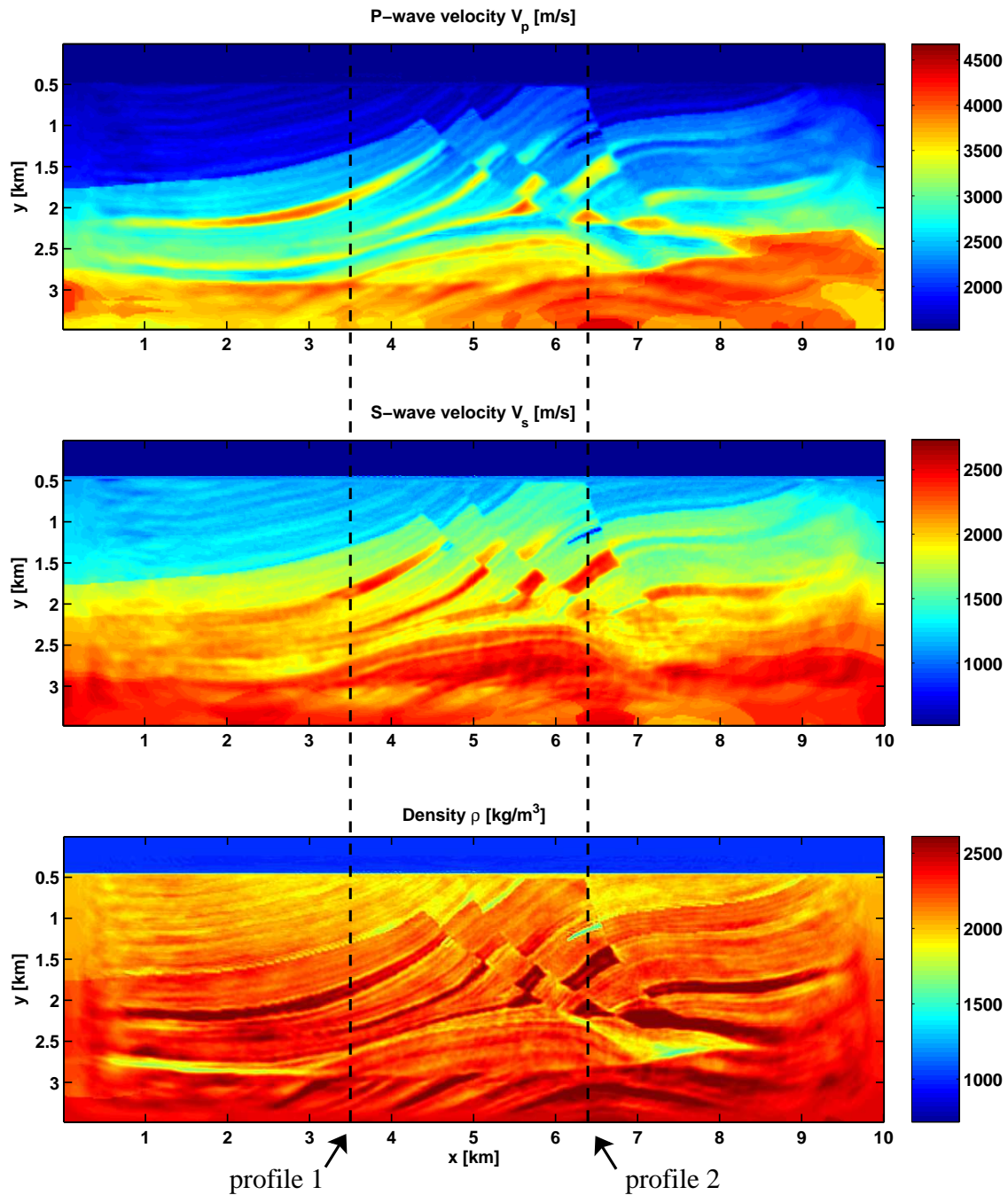


Figure 6.7: Results of the elastic FWT for the Marmousi-II model. The dashed lines denote the positions of the depth profiles 1 and 2 shown in Fig. 6.8.

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

---

It is quite surprising, that the shear wave velocity model could also be resolved very well, even though only streamer data and therefore mainly P-wave information is used. Even the density, a parameter which can be hardly estimated from seismic data, could be recovered from the seismic wavefield. Keep in mind though, that the density image is based not only density information, but contains also  $V_p$  and  $V_s$  information due to the ambiguity investigated by the CTS test problem (chapter 5.2). The depth profiles show the strong deviation of the starting model from the true Marmousi model, in some parts up to  $\pm 500 - 1000$  m/s for the velocity models and  $\pm 250 - 500$  kg/m<sup>3</sup> for the density models. Even though the FWT could reconstruct the velocity and density structures quantitatively fairly well. In Fig. 6.9 the seismic section of shot 50 is plotted for the starting model (a), the inversion result (b) the true model (c), the initial residuals (d), the final residuals (e) and the evolution of the residual energy (f). Notice the good fit of the first arrivals for the starting model, but the lack of small details beyond the first arrivals. Only the direct wave, the reflection from the ocean bottom and a few multiples are present. The inversion result fits most of the phases and amplitudes of the later small scale reflections from within the thrust fault system and therefore the data residuals are very small. The misfit function decreases smoothly for the first 20 Iteration steps, but exhibits strong fluctuations for the later iterations, due to the strong nonlinear character of the inversion problem for this complex geology. The performance of the FWT code was benchmarked on an Altix 4700 using the Marmousi2 model. Fig.6.10 shows the computation time (top) and the memory requirements for up to 50 CPUs (bottom). On a single CPU the elastic timedomain FWT is not efficient at the moment. For 150 iteration steps a total calculation time of 85 days and about 10 GB of RAM would be required. When using 50 CPUs the computation time is reduced to roughly 2 days. Note the linear speedup of the FWT code due to the optimized parallelization of the forward modelling FD code.

In conclusion the results look very impressive, but you should not forget that the starting model for the seismic velocities and density are unrealistically accurate and are not easy to derive from the seismic data for such a complex model. The influence of more realistic starting models on the FWT for a simpler test problem will be discussed in chapter 10. This simple test problem is subject of the next chapter.

## 6.1. THE COMPLEX MARMOUSI2 MODEL

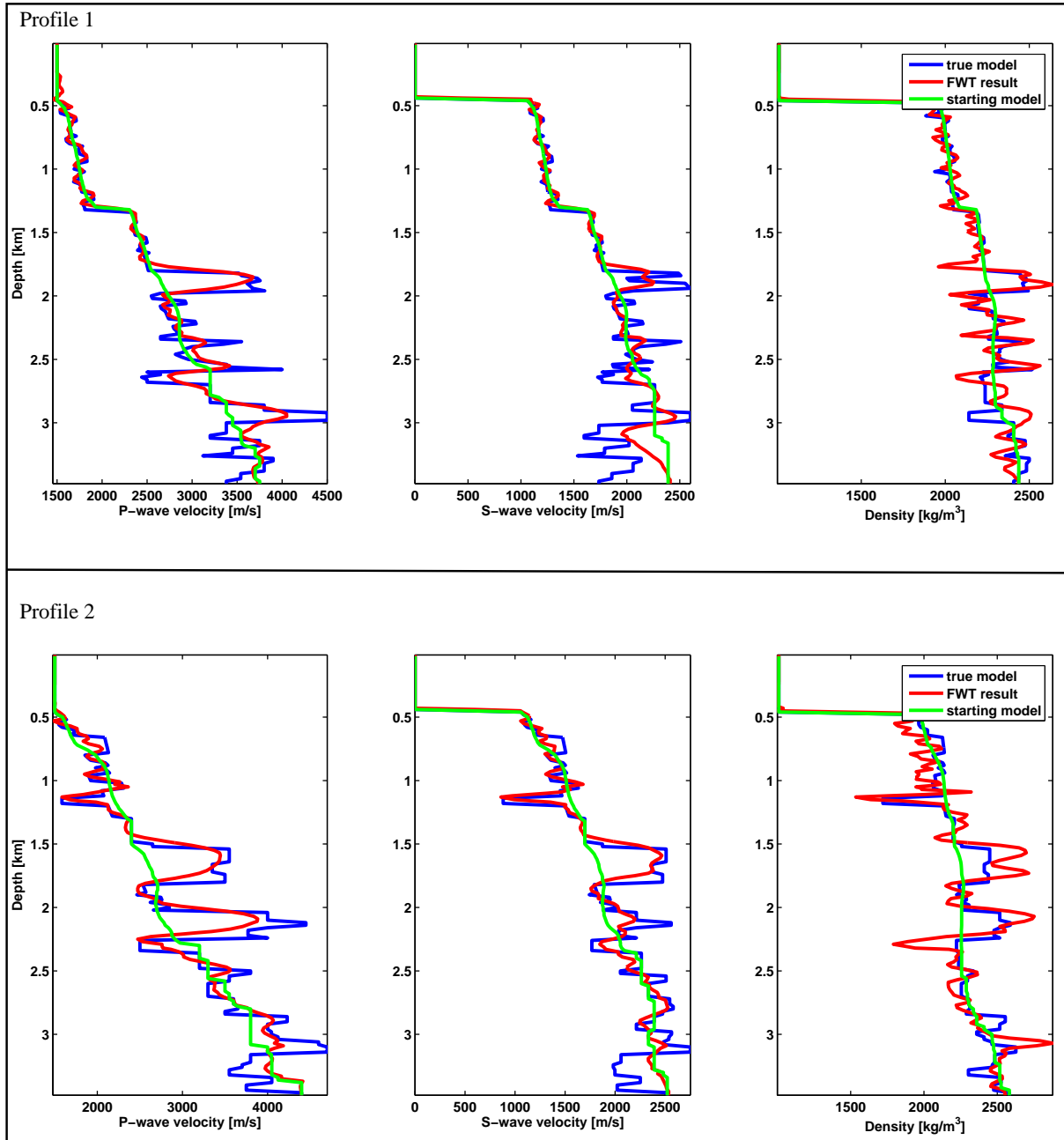


Figure 6.8: Depth profiles at  $x_{p1} = 3.5$  km (top) and  $x_{p2} = 6.4$  km (bottom) of the starting model and FWT result are compared with the true model for the Marmousi-II model: P-wave velocity (left), S-wave velocity (center) and density (right).



## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

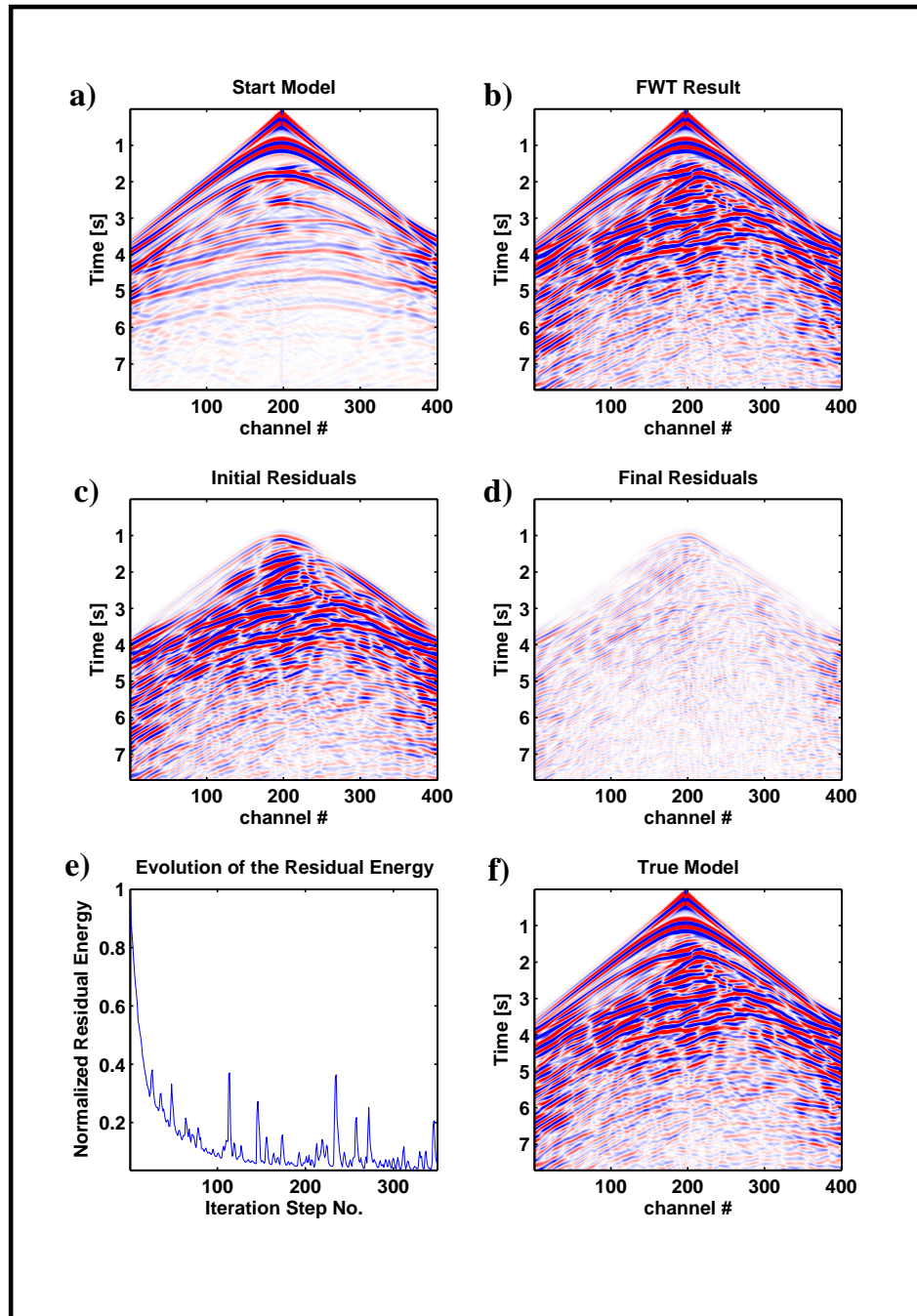


Figure 6.9: Seismic sections (shot 50, y-component) for the Marmousi-II model. (a) starting model, (b) FWT result, (c) initial residuals, (d) final residuals, (f) true model and (e) evolution of the residual energy.



## 6.1. THE COMPLEX MARMOUSI2 MODEL

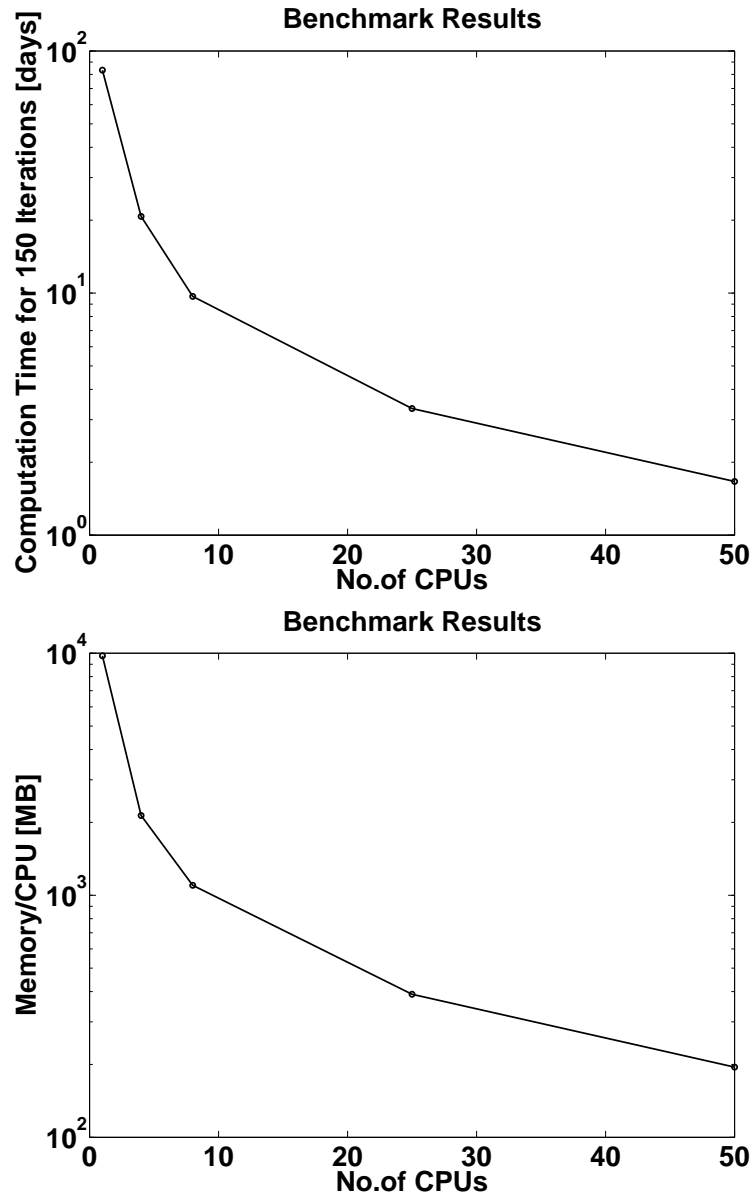


Figure 6.10: Benchmark results for the Marmousi2 model. The absolute calculation times (top) and the memory requirements (bottom) for up to 50 CPUs on an Altix 4700.

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

---

### 6.2 The simple Marmousi2 model

---

The "simple" part of the Marmousi2 model consists of a 500 m thick water layer above an elastic subseafloor model. The sediment model shows an approximately 1D geology down to a depth of 1 km with low seismic velocities and only small impedance contrasts between the fine layers with thicknesses of approximately 100 m. Between 1 and 1.7 km depth the sediment layers are tilted. Embedded in these layers are two hydro carbon reservoirs of different sizes, a small gas reservoir (marked by A in Fig. 6.1) and a small layer consisting of oil sand (marked by B in Fig. 6.1). Below 1.7 km the geology becomes more complex and consists of thick salt layers and rotated sediment blocks.

#### 6.2.1 Acquisition geometry and FD model

The reflection seismic acquisition geometry consists of 50 explosive sources 40 m below the free surface. The source signature is a Ricker wavelet with a centre frequency of 10 Hz. The elastic wavefield is recorded by a line of 400 two component receivers at the same depth as the sources. Despite the top of the model, where a free surface is present, the boundaries are surrounded by an exponential damping absorbing boundary layer with a thickness of 500 m. Using an 8th order spatial FD operator a grid point spacing of  $dh = 10.0$  m is used. The model has the dimensions  $5.0 \times 3.48$  km, therefore the spatial model grid consists of  $500 \times 348$  gridpoints. To avoid a violation of the Courant criterion a time step size  $dt = 0.55$  ms is used. For a recording time  $T = 6.0$  s about 11000 time steps are required.

#### 6.2.2 Elastic wave propagation in the simple Marmousi model

Fig. 6.12 shows the development of the pressure wavefield excited by shot 50 for the simple elastic Marmousi2 model at 6 different time steps. In contrast to the complex wavefield generated within the thrust fault system of the complex Marmousi2 model (Fig. 6.3) the wavefield in the simple 1D Marmousi model is dominated by simple low amplitude reflections from within the low velocity sediment layers down to a depth of 1.8 km ( $T=720$  ms -  $T=1329$  ms). Strong reflections and mode conversions are generated by the gas reservoir A. At the interface between the low velocity and the high velocity sediments and the salt layers at a depth of 1.8 km and 2.25 km stronger reflections occur ( $T=1080$  ms -  $T=2160$  ms).

## 6.2. THE SIMPLE MARMOUSI2 MODEL

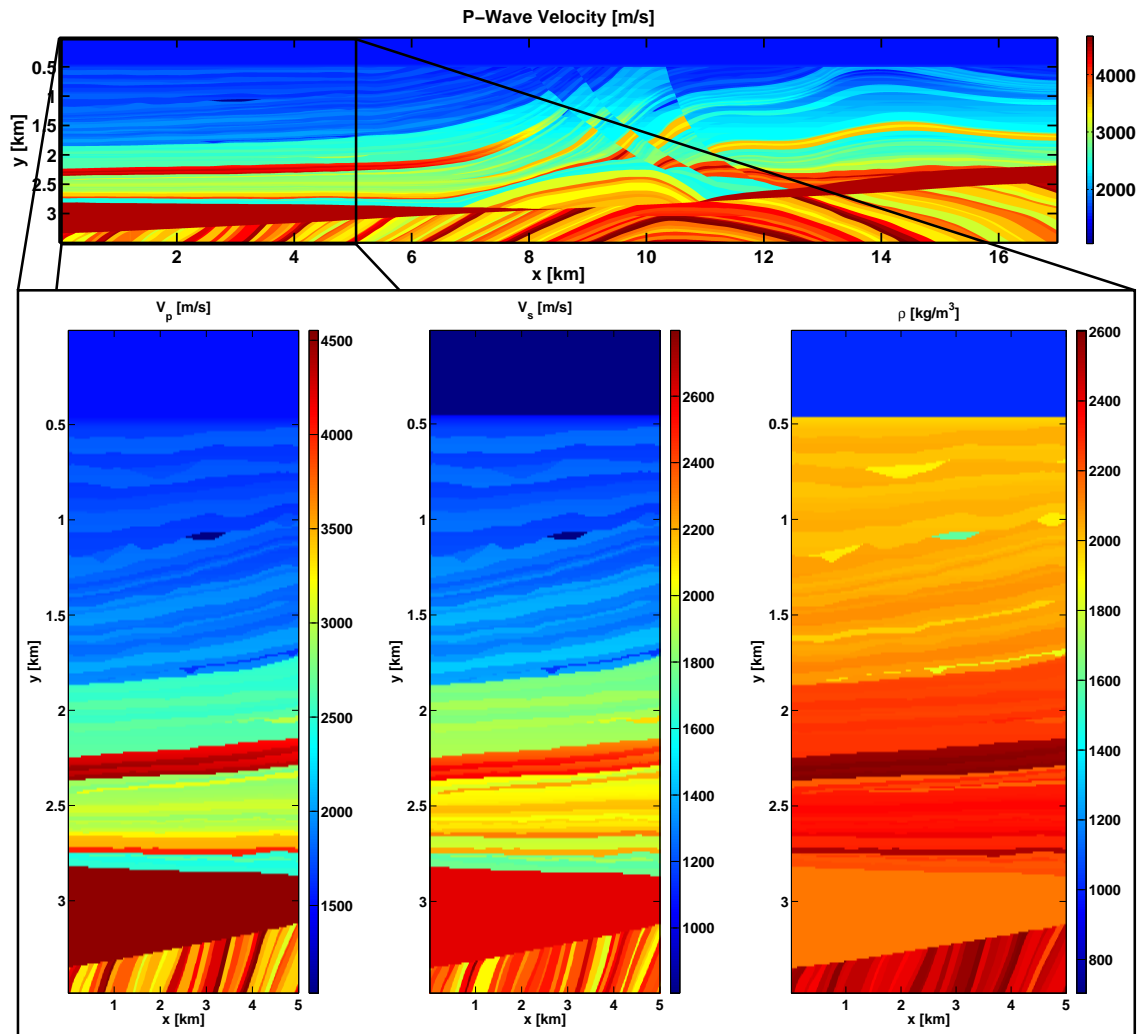


Figure 6.11: The part of the elastic Marmousi model with 1D geology used for the starting model study: P-wave velocity  $V_p$  (left), S-wave velocity  $V_s$  (center) and Density  $\rho$  (right).

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

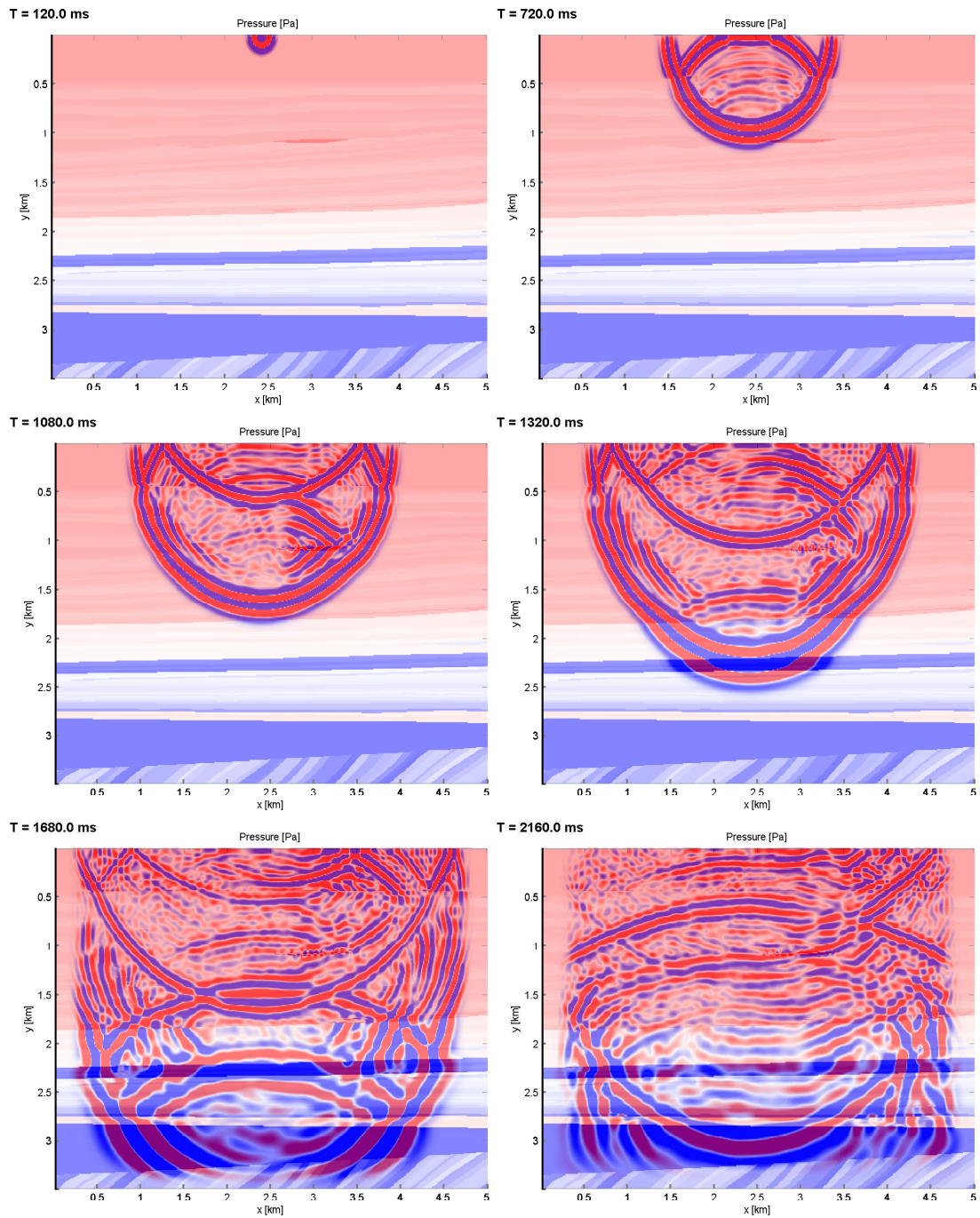


Figure 6.12: Pressure wavefield excited by shot 25 for the 1D elastic Marmousi2 model at 6 different time steps .

### 6.2.3 FWT of the simple Marmousi model

For later comparison with more simpler starting models a smoothed version of the true model is used as a perfect starting model in this chapter. To generate the starting model the same Gaussian filter Eq. (6.1) with the same correlation length  $\lambda_c$  as in case of the complex Marmousi2 model is used. The resulting starting models (Fig. 6.13) are very smooth. Even though the hydrocarbon reservoirs A and B and the water wet sand are still very dominant, especially in the density model. As in the case of the complex Marmousi model the application of a preconditioning operator is vital to suppress the large gradient values near the source and receiver positions. The preconditioning operator Eq. 6.2 for the complex Marmousi model is also used for the simple Marmousi model. No frequency filtering is applied.

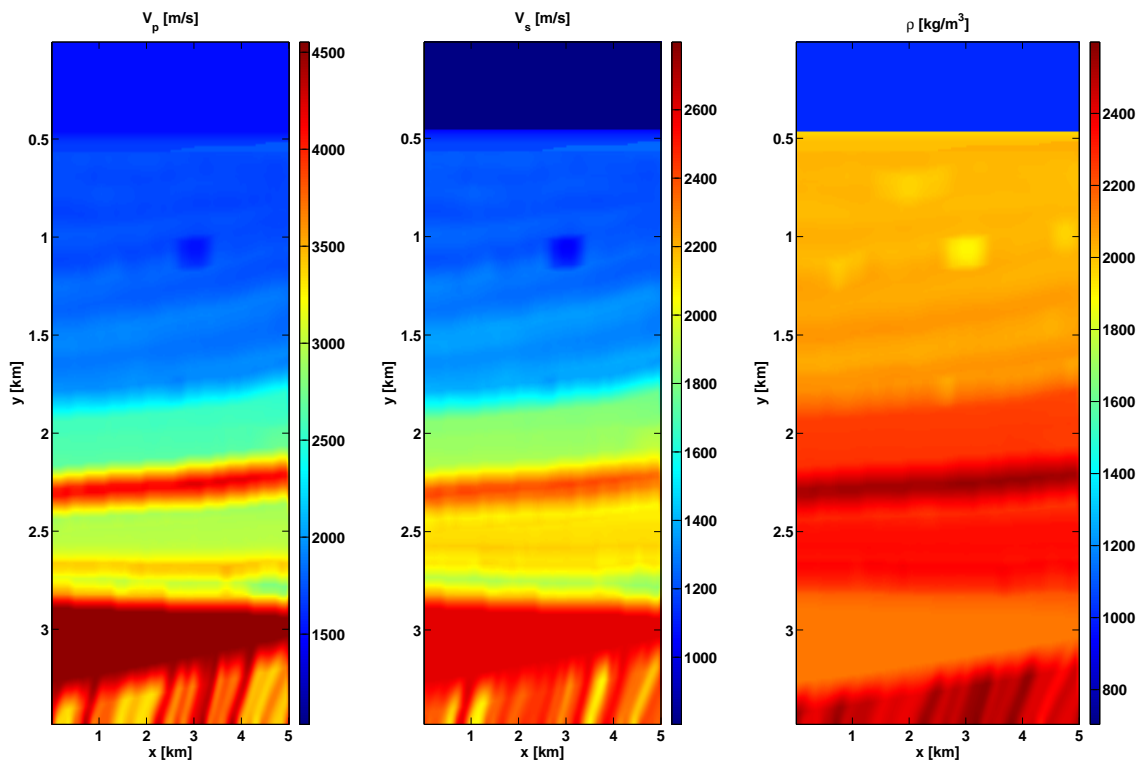


Figure 6.13: Starting models for the part of the elastic Marmousi model with 1D geology: P-wave velocity  $V_p$  (left), S-wave velocity  $V_s$  (center) and Density  $\rho$  (right).

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

---

The inversion results after 88 iterations are shown in Fig. 6.14. Fig. 6.15 shows a depth profile located at  $x_p = 3.0$  km for the starting model, the inversion result and the true model. Because the starting model already fitted the true model very well only small details are added during the inversion. The sediment layers in the upper part of the model, as well as the dominant gas reservoir A and the water wet sediments are perfectly imaged. Even the small layer containing the oil sand B is clearly visible in all three material parameters. The deeper parts of the model also contain some higher resolved layers, but is more dominated by the starting model. The quality of the inversion results is also evident in the seismic sections of shot 50 (y-component) plotted in Fig. 6.16. Notice the direct wave, the reflection from the ocean bottom, a few multiples and the dominant reflection from the interface between low and high velocity sediments, but the lack of small details beyond the first arrivals in the seismic section of the starting model (a). The initial data residuals (c) show no direct wave anymore, so it is fitted perfectly. The residuals only contain small scale features, like the strong reflections from the gas reservoir A between 1.5 and 2.2 s. Residuals in the reflection from the high impedance interface at 2.25 km depth are due to the large model misfit in the sediment layers above this interface. Therefore this reflection contains a lot of crucial transmission information. The seismic section of the FWT result (b) is nearly identical with the seismic section of the true model (f), so the final data residuals (d) are close to zero. All reflection events could be fitted perfectly. The misfit function (e) decreases very fast and smooth, no peaks or local minima are visible.

In conclusion for unrealistic perfect starting models the inversion results are also perfect. Now lets turn to the estimation of more realistic starting models and see what impact they will have on the quality of the FWT image.

## 6.2. THE SIMPLE MARMOUSI2 MODEL

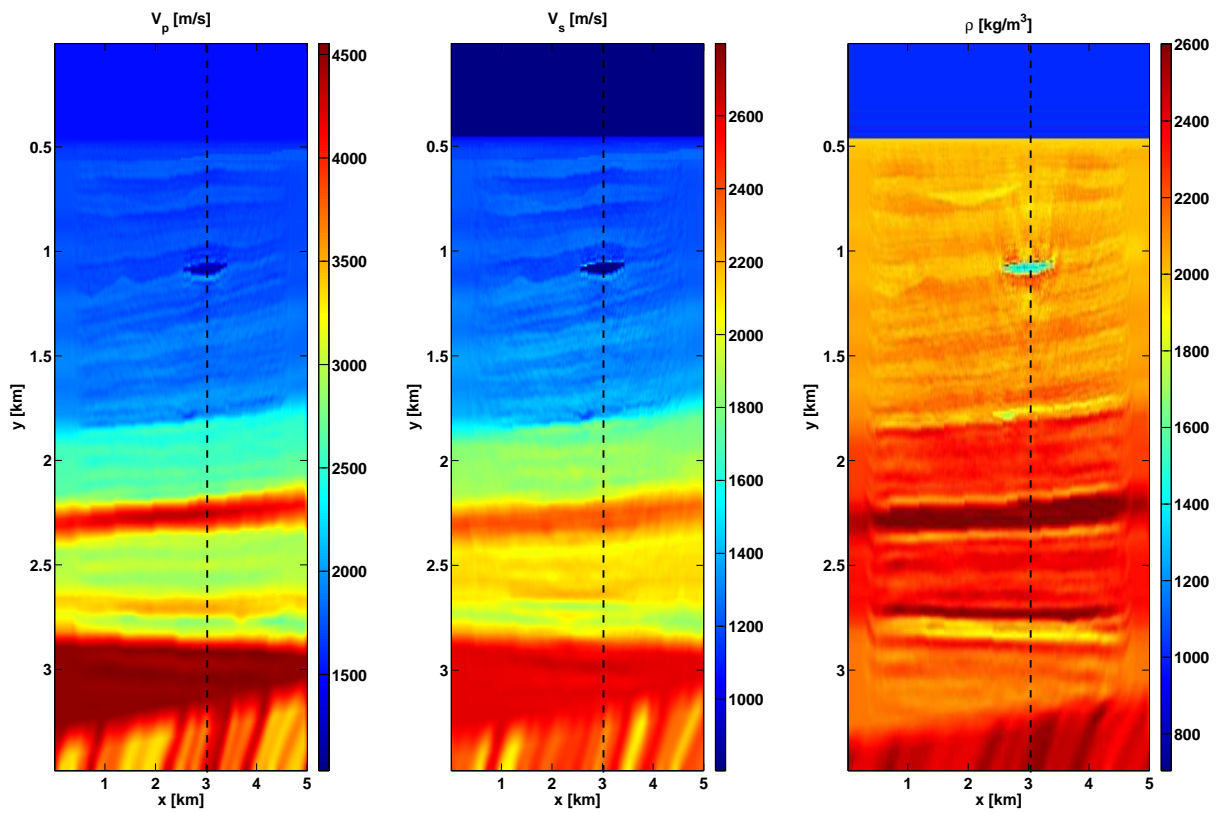


Figure 6.14: FWT results for the part of the elastic Marmousi model with 1D geology: P-wave velocity  $V_p$  (left), S-wave velocity  $V_s$  (center) and Density  $\rho$  (right). The dashed lines denote the positions of the depth profile shown in Fig. 6.15.

## CHAPTER 6. A GEOLOGICAL TEST PROBLEM - THE ELASTIC MARMOUSI2 MODEL

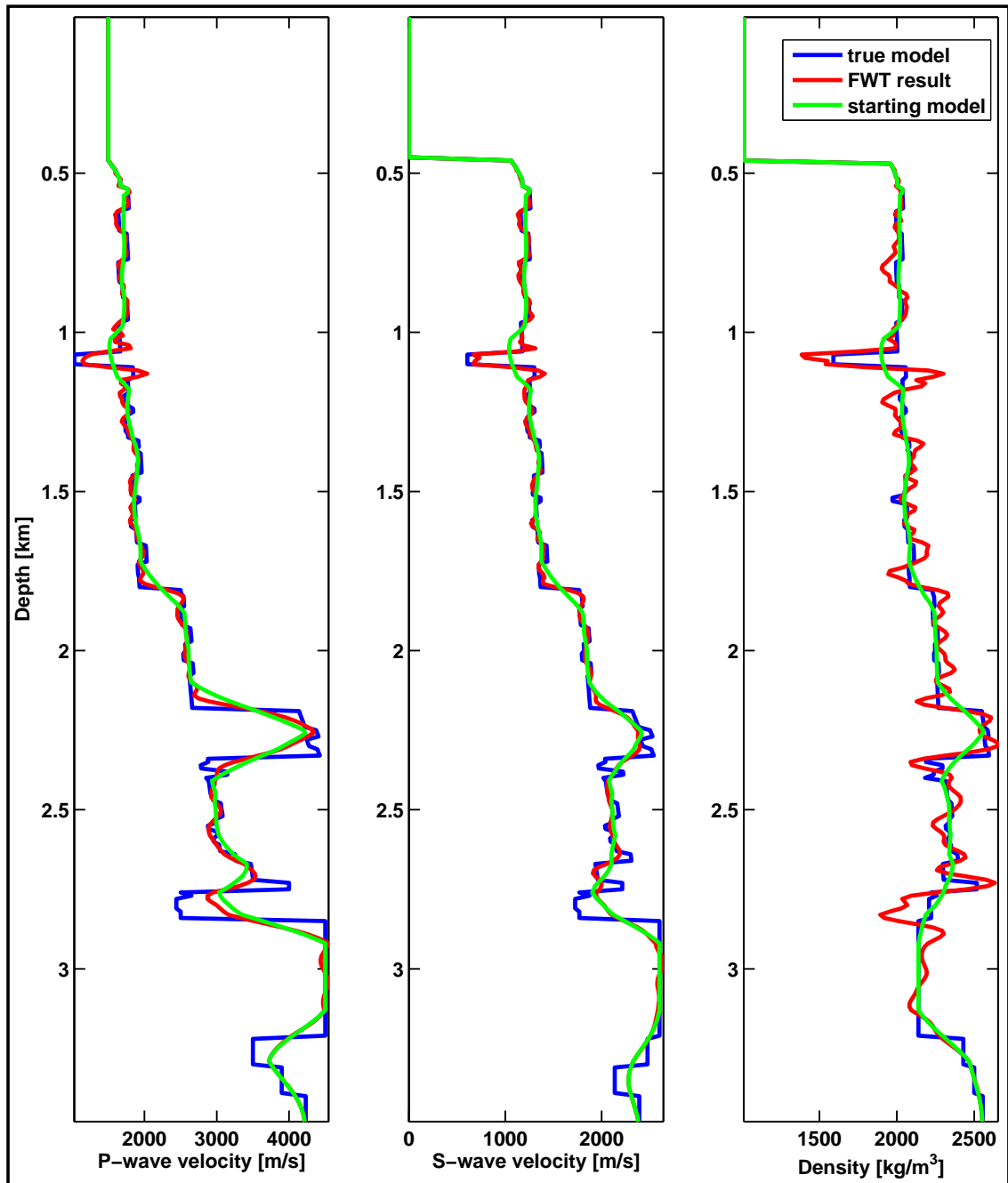


Figure 6.15: Comparison of the depth profiles at  $x_p = 3.0$  km for the starting model and FWT result with the true model for the simple Marmousi-II model: P-wave velocity (left), S-wave velocity (center) and density (right).



## 6.2. THE SIMPLE MARMOUSI2 MODEL

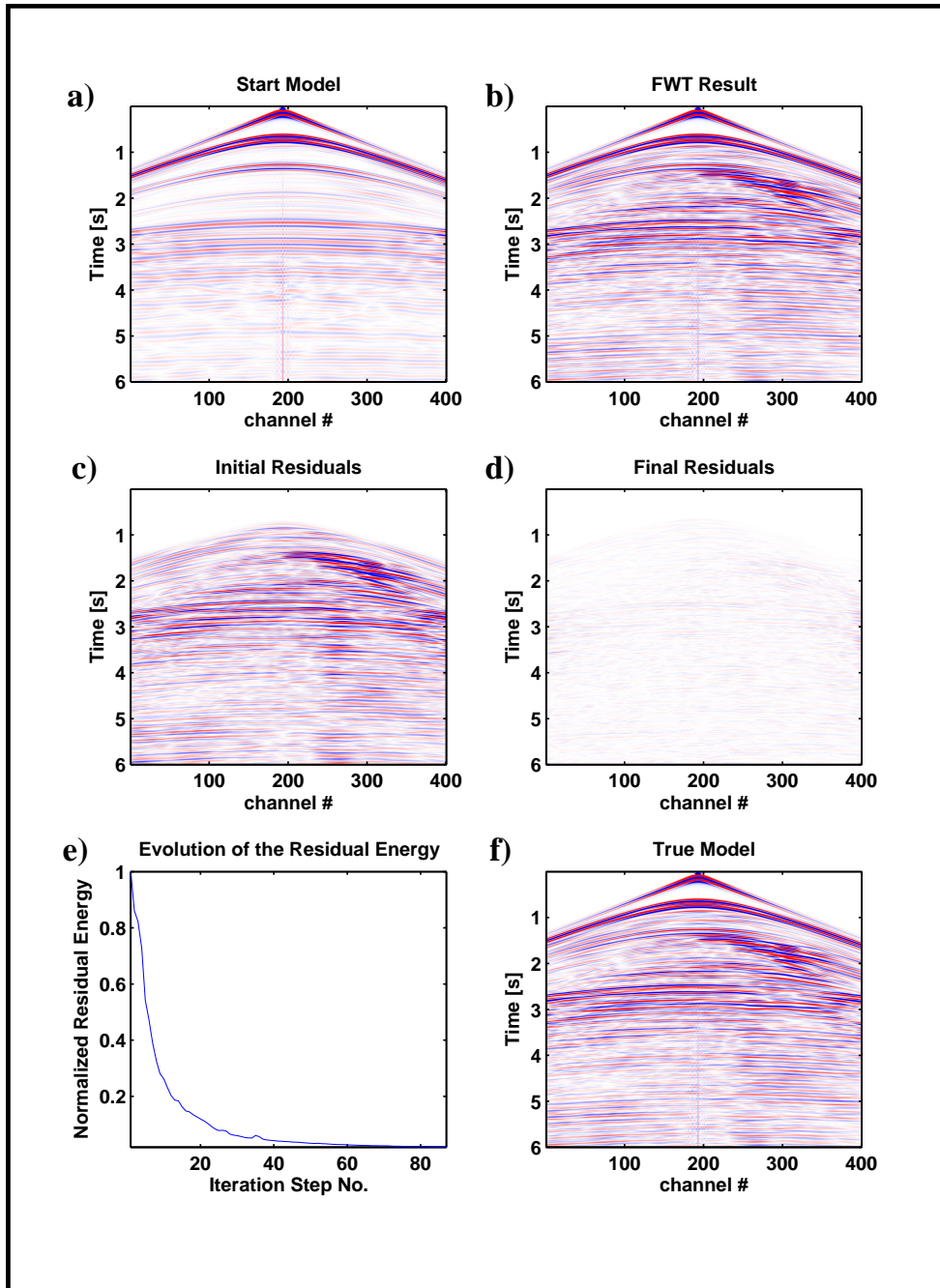


Figure 6.16: Seismic sections for the simple Marmousi-II model (shot 25, y-component). (a) starting model, (b) FWT result, (c) initial residuals, (d) final residuals, (e) evolution of the residual energy, and (f) true model.



# 7

## 1D Starting model estimations with Evolution Strategy

In the last four chapters I have demonstrated the performance of the elastic FWT for starting models which are very close to the true model and therefore the nonlinearity of the problem was not really taken into account. In this chapter the influence of the starting model will be discussed for the geological "simple" part of the elastic Marmousi model (section 6.2).

### 7.1 Basic starting models

---

Because the geology of the test problem is approximately 1D a very simple 1D starting model should be sufficient. Possible starting models (Fig. 7.1) consist of a water layer with the acoustic material parameters  $\mathbf{m}_w = [V_p, \rho]$  and a thickness  $d_w$ . Beneath the seafloor is an elastic half space, where the material parameters are increasing linear with depth  $d$

$$\begin{aligned} V_p(d) &= (V_{p0} - g_{vp}d_w) + g_{vp}d, \\ V_s(d) &= (V_{s0} - g_{vs}d_w) + g_{vs}d, \\ \rho(d) &= (\rho_0 - g_\rho d_w) + g_\rho d, \end{aligned} \tag{7.1}$$

where  $\mathbf{m}_0 = [V_{p0} \ V_{s0} \ \rho_0]$  are the values of the material parameters at the seafloor and  $\mathbf{g} = [g_{vp} \ g_{vs} \ g_\rho]$  are the gradients of the corresponding material parameters.

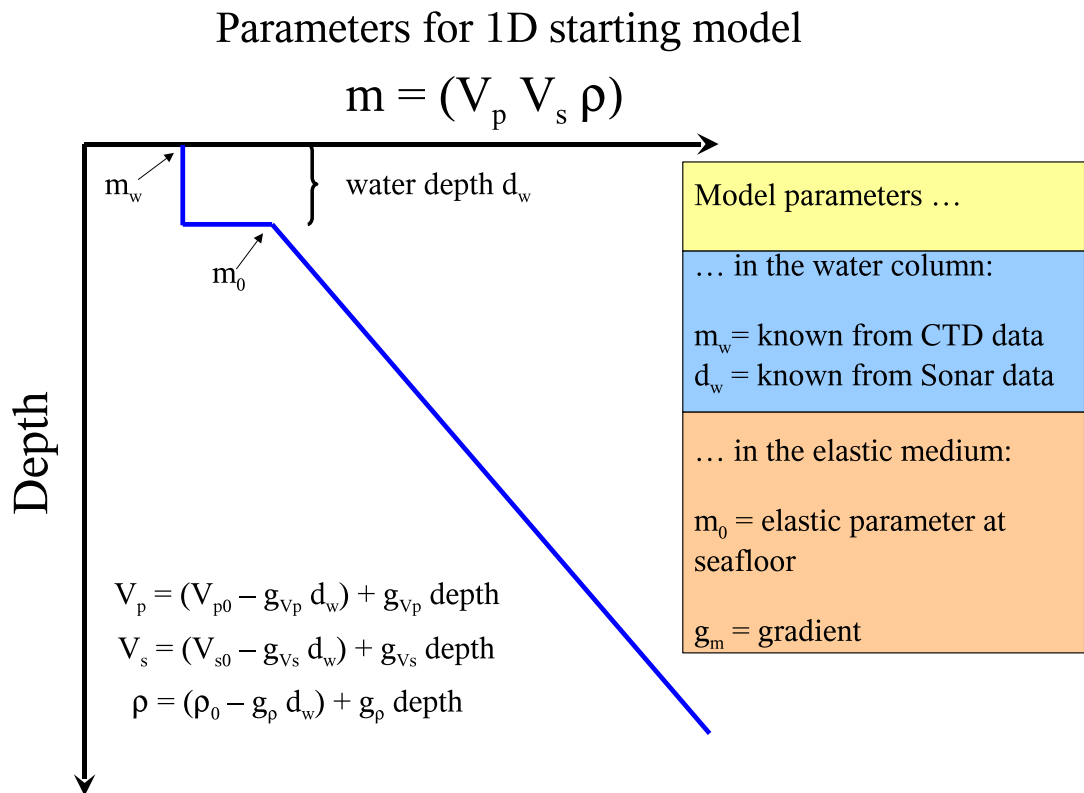


Figure 7.1: Definition of the simple 1D starting model.

## 7.2 Evolution Strategy

To find possible starting models a global search strategy will be applied. Even though this search strategy would also work if all parameters  $\mathbf{m}_w$ ,  $\mathbf{d}_w$ ,  $\mathbf{m}_0$  and  $\mathbf{g}$  are unknown I reduce the size of the parameter space by assuming, that the depth of the water column  $d_w = 470$  m is known from sonar data and the material parameters of the water layer  $\mathbf{m}_w = [V_p = 1500 \text{ m/s}, V_s = 0.0 \text{ m/s and } \rho = 1010 \text{ kg/m}^3]$  are estimated from CTD data. Therefore only the 6 parameters describing the elastic halfspace  $\mathbf{m}_0$  and  $\mathbf{g}$  are unknown. The global search strategy applied here is an **Evolution Strategy (ES)**. It is based on the ideas of adaption and evolution and was created and developed in the 1960s and 1970s by Rechenberg, Schwefel and co workers (Rechenberg (1971), Schwefel (1974)). I explain the basics of the evolution strategy using the simple example in Fig. 7.2. One iteration, or speaking in terms of evolution generation, of the Evolution Strategy consists of the following steps:

1. In evolution strategy the unknown model parameters are interpreted as genetic code, which is schematically sketched as chromosomes in Fig. 7.2. For our simple 1D starting model the genetic code consists of 6 chromosomes. Two different models 1 (colored blue) and 2 (colored red) have different genetic codes. In the following these two models are called the **parental population**.
2. A third model 3 (called **offspring population**) can be generated from the parental population by taking randomly half of the genetic information from model 1 and the other half of the genetic information from model 2. In this case the parameters  $V_{s0}$ ,  $\rho_0$  and  $g_\rho$  are taken from model 1 and the parameters  $V_{p0}$ ,  $g_{V_p}$  and  $g_{V_s}$  are taken from model 2.
3. Now comes the evolution in the evolution strategy. Each model parameter of the offspring population is perturbed by a random mutation  $\delta m_{\text{evo}}$  which is simply a scaled Gaussian random number  $d_{\text{rand}}$  defined on the interval  $[0, 1]$

$$\delta m_{\text{evo}} = (2.0 * \delta m_{\text{max}} * d_{\text{rand}}) - \delta m_{\text{max}}, \quad (7.2)$$

so that the maximum mutation variation is restricted to the interval  $[-\delta m_{\text{max}}, \delta m_{\text{max}}]$ .

4. Calculate the misfit function for the parental and offspring populations using the FD forward modeling code which is also used for the FWT.
5. Choose the best fitting models (in this case 2) from the parental and offspring populations and use them as parental population in the next generation.

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

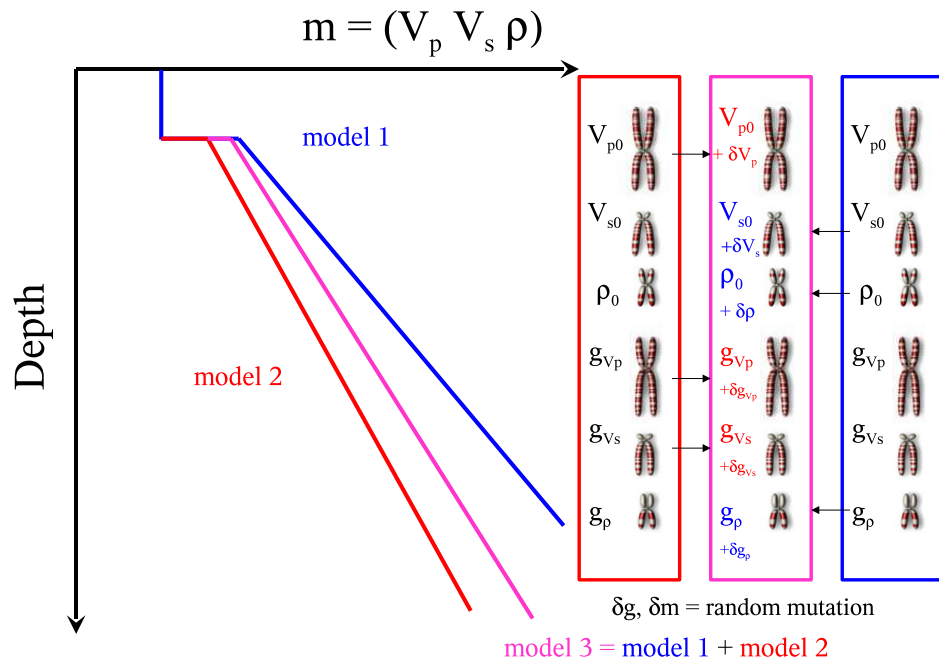


Figure 7.2: The concept of evolution strategy.

In practice there are some common implementations of the evolution strategy. If the parental population consists of  $\mu$  models, they can produce  $\lambda$  models in the offspring population. According to Schwefel (1974) we can distinguish

- A  $(\mu + \lambda)$ -ES where the best fitting models for the next parental generation are chosen from all the models of the parental and offspring population.
- A  $(\mu, \lambda)$ -ES where the best fitting models for the next parental generation are chosen from the models of the offspring population only. The models in the parental generation are neglected.
- A  $(\mu + 1)$ -ES where the offspring population only consists of the best fitting model of the parental population.
- A  $(1+1)$ -ES where each population consists of only one model. One model of the parental population creates one model in the offspring population.

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

The choice of the number of members in the parental and offspring population and the size of the mutation is essential for a successful evolution. If the population and the mutation are too small the models might concentrate near one local minimum of the misfit function. This can be compensated to some extent by introducing a variable random mutation  $\delta m_{\text{evo}}$ . For the first generations the mutation might be large, so that the populations consist of a lot of different individual models and afterwards the mutation is reduced, so that the models can cluster near local minima of the misfit function. However there is no general approach to solve this problem.

### 7.3 Starting model estimations by (10+45)-ES

To estimate 1D starting models for the FWT a (10+45)-ES is used. The parameters of the 1st parental population are listed in table 7.1.

Model No.	$V_{p0}$ [m/s]	$V_{s0}$ [m/s]	$\rho_0$ [kg/m <sup>3</sup> ]	$g_{vp}$ [1/s]	$g_{vs}$ [1/s]	$g_\rho$ [kg/m <sup>4</sup> ]
1	800	800	1000	0.2	0.2	0.2
2	1000	1000	1200	0.25	0.25	0.25
3	1200	1200	1400	0.3	0.3	0.3
4	1400	1400	1600	0.35	0.35	0.35
5	1600	1600	1800	0.4	0.4	0.4
6	1800	1800	2000	0.45	0.45	0.45
7	2000	2000	2200	0.5	0.5	0.5
8	2200	2200	2400	0.55	0.55	0.55
9	2400	2400	2600	0.6	0.6	0.6
10	2600	2600	2800	0.65	0.65	0.65

Table 7.1: Model parameters for the parental population of the first generation of the (10 + 45)-ES.

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

---

They cover a quite large part of the parameter space. The seismic velocities  $V_p$  and  $V_s$  vary between 800 and 2600 m/s, the density  $\rho$  between 1000 and 2800 kg/m<sup>3</sup>. The slopes of the linear gradients for all parameters range from 0.2 to 0.65 1/s and kg/m<sup>4</sup>. The maximum mutation variation  $\delta m_{\max}$  for the model parameters  $\mathbf{m}_0$  and  $\mathbf{g}$  as a function of the generation  $n$  are defined as

$$\delta m_{0\max} = \begin{cases} 100.0 \text{ m/s [kg/m}^3] & \text{if } n \leq 100 \\ 20.0 \text{ m/s [kg/m}^3] & \text{if } n > 100 \end{cases}$$

$$\delta \mathbf{g}_{\max} = \begin{cases} 0.1 \text{ 1/s [kg/m}^4] & \text{if } n \leq 100 \\ 0.05 \text{ 1/s [kg/m}^4] & \text{if } n > 100 \end{cases}$$

Therefore during the first 100 generations the model parameters can change very much and can scan the long wavelength parameter space, while the variation is strongly reduced for the last 80 generations to refine the models near the local minima. Fig. 7.3 shows the results of the first generation using the (10+45)-ES for the  $V_p$ ,  $V_s$  and  $\rho$  model, respectively. The black lines denote the average of the corresponding true model, which has to be fitted by the ES. The red models are members of the parental population, while the dashed blue models belong to the offspring population. The parental and offspring population of the first generation cover a large part of the parameter space. In the 2nd generation, shown in Fig. 7.4, a lot of unrealistic models already disappeared. This has 3 reasons:

1. The FD code is unstable, if the seismic velocities are too high and violate the Courant criterion. This leads to high or NaN values of the misfit function.
2. If the seismic velocities are too low grid dispersion will be present which also results in high values of the misfit function.
3. If  $V_s$  is larger than  $V_p$  the FD model also produces nonphysical results which leads to high values of the misfit function.

By taking a closer look at the individual model parameters it can be noted, that the  $V_p$  models show a separation in 2 populations. One large population that fits the upper part of the true model down to a depth of 1800 m and a smaller population which fits the velocity structures at larger depth of the model. The value of  $V_{p0}$  range from 1500 m/s to 2500 m/s for population 1 and unrealistic large values of around 3500 m/s for population 2. The slopes  $g_{V_p}$  in both populations are comparable and vary between 0.2 and 0.6 1/s. The models for  $V_s$  and  $\rho$  each show one large population



---

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

---

of models around the true average model. For  $V_{s0}$  the values range between 1500 m/s and 2500 m/s. The minimum and maximum slope values  $g_{V_s}$  are 0.1 and 0.7 1/s. The parameters for the density are in a similar range.  $\rho_0$  varies between 1500 and 2600 kg/m<sup>3</sup>, the slopes  $g_\rho$  between 0.2 and 0.6 kg/m<sup>4</sup>. The parameters of the starting models are still dominating the models at generation 2. After 100 generations (Fig. 7.5) the models for  $V_p$ ,  $V_s$  and  $\rho$  all fit the global 1D model quite well. The two  $V_p$  model populations from the 2nd generation vanished and instead one large population of quite similar models is present. The variation for  $V_{p0}$  is very small, between 1400 and 1500 m/s. The slopes  $g_{V_p}$  are large and range from 0.78 to 0.85 1/s. The model parameters for  $V_s$  and  $\rho$  also show only little variation at the seafloor,  $V_{s0}$  and  $\rho_0$  vary between 1000 and 1100 m/s respectively 1900 and 2000 kg/m<sup>3</sup>. On the other hand the corresponding slope variations are large and range from 0.1 to 0.3 1/s for  $g_{V_s}$  and 0.01 to 0.6 kg/m<sup>4</sup> for  $g_\rho$ . After 80 additional generations (Fig. 7.6) with smaller mutation parameters the  $V_p$  and  $V_s$  models are much more focused and fit the global trend of the average true model very well. The parameters for  $V_{p0}$  and  $V_{s0}$  show only small derivations from the values  $V_{p0} = 1400$  m/s and  $V_{s0} = 1070$  m/s. The slopes also consolidated at values of  $g_{V_p} = 0.9$  1/s and  $g_{V_s} = 0.365$  1/s. The possible density models range from very good fits which show approximately no density variation with depth ( $g_\rho = 0.02$  kg/m<sup>4</sup>) and moderate fitted models with stronger gradients ( $g_\rho = 0.3$  kg/m<sup>4</sup>). The estimation of the density value at the seafloor is very good ( $\rho_0 = 2070$  kg/m<sup>3</sup>).

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

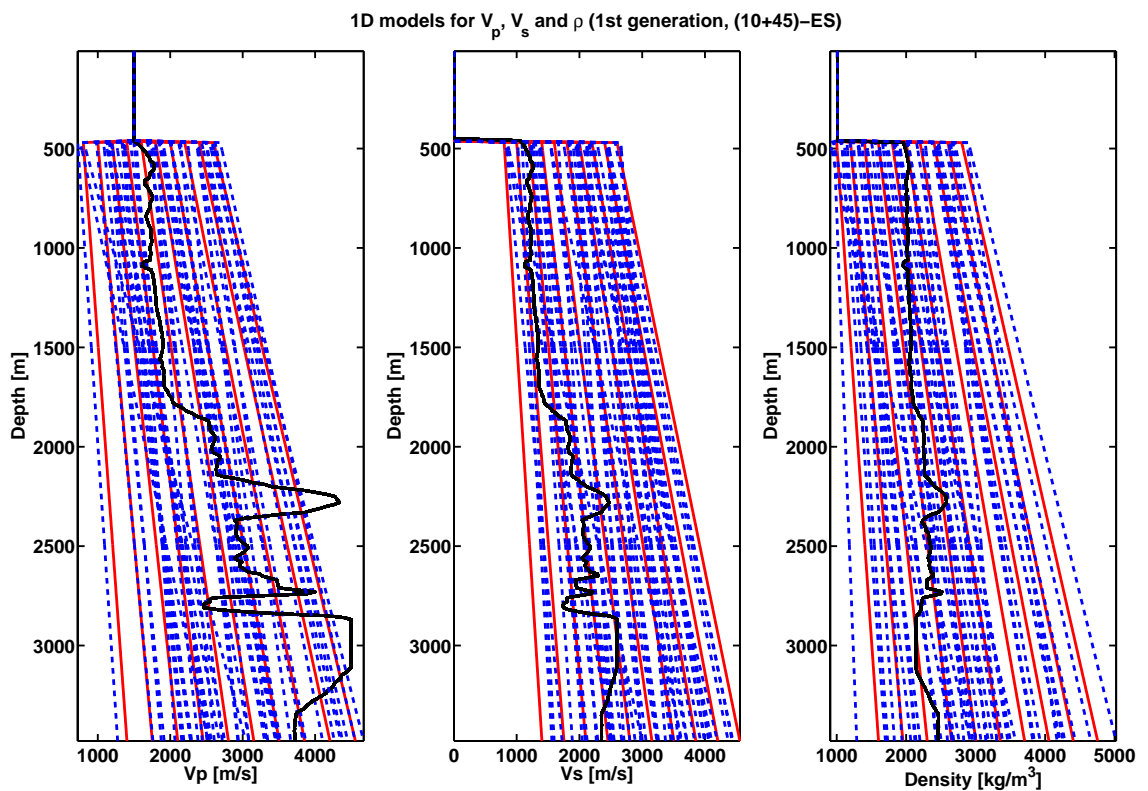


Figure 7.3: Models of the parental population (red models) and offspring population (blue dashed models) compared with the average true model (black model) for the 1st generation using a (10+45)-ES.

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

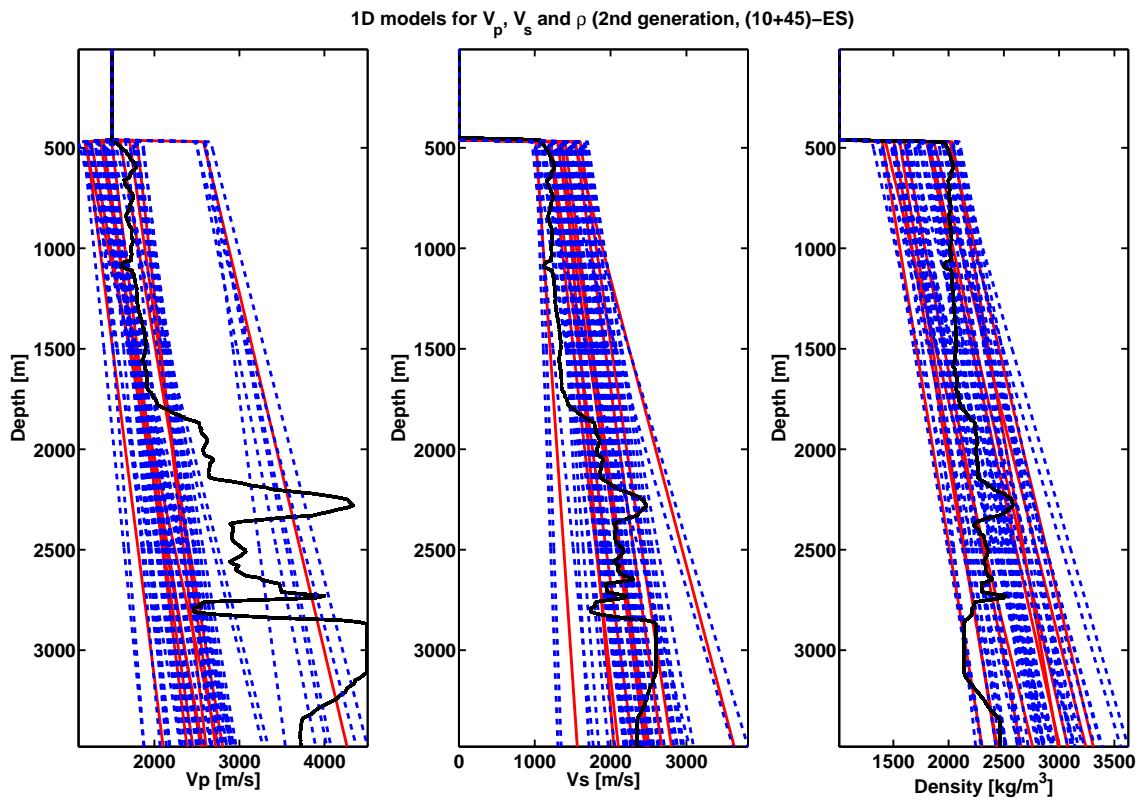


Figure 7.4: Models of the parental population (red models) and offspring population (blue dashed models) compared with the average true model (black model) for the 2nd generation using a (10+45)-ES.

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

---

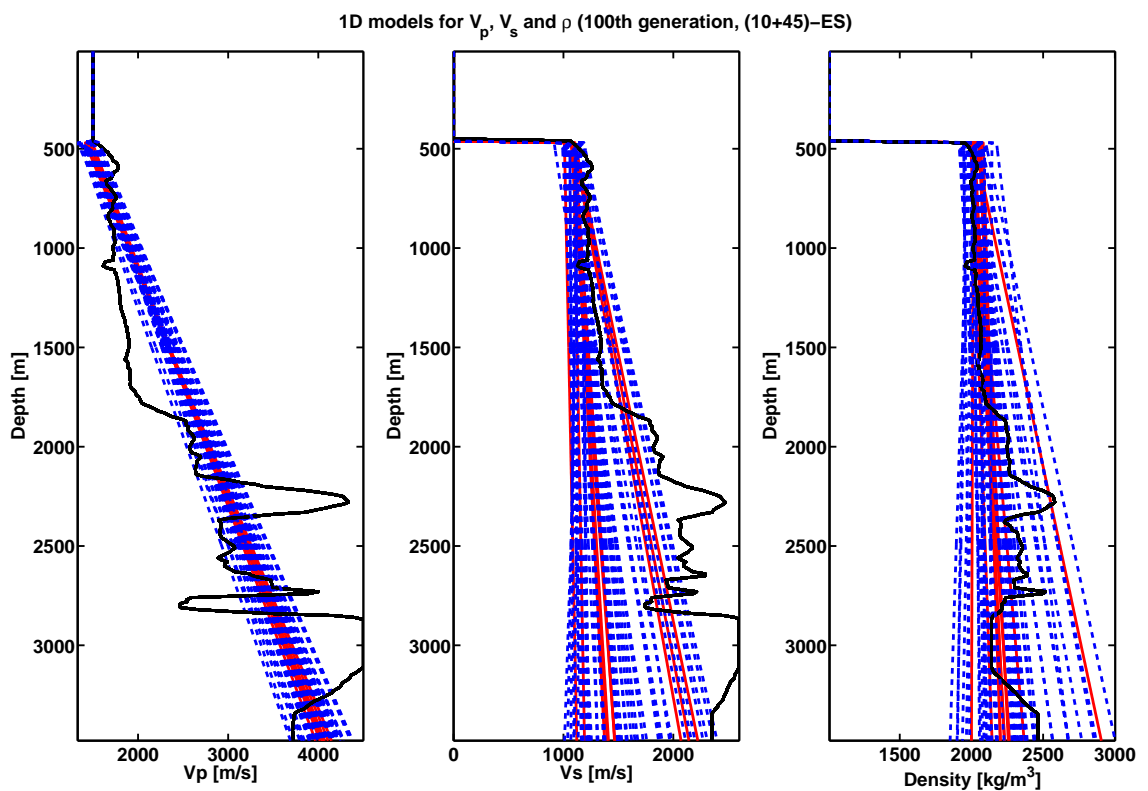


Figure 7.5: Models of the parental population (red models) and offspring population (blue dashed models) compared with the average true model (black model) for the 100th generation using a (10+45)-ES.

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

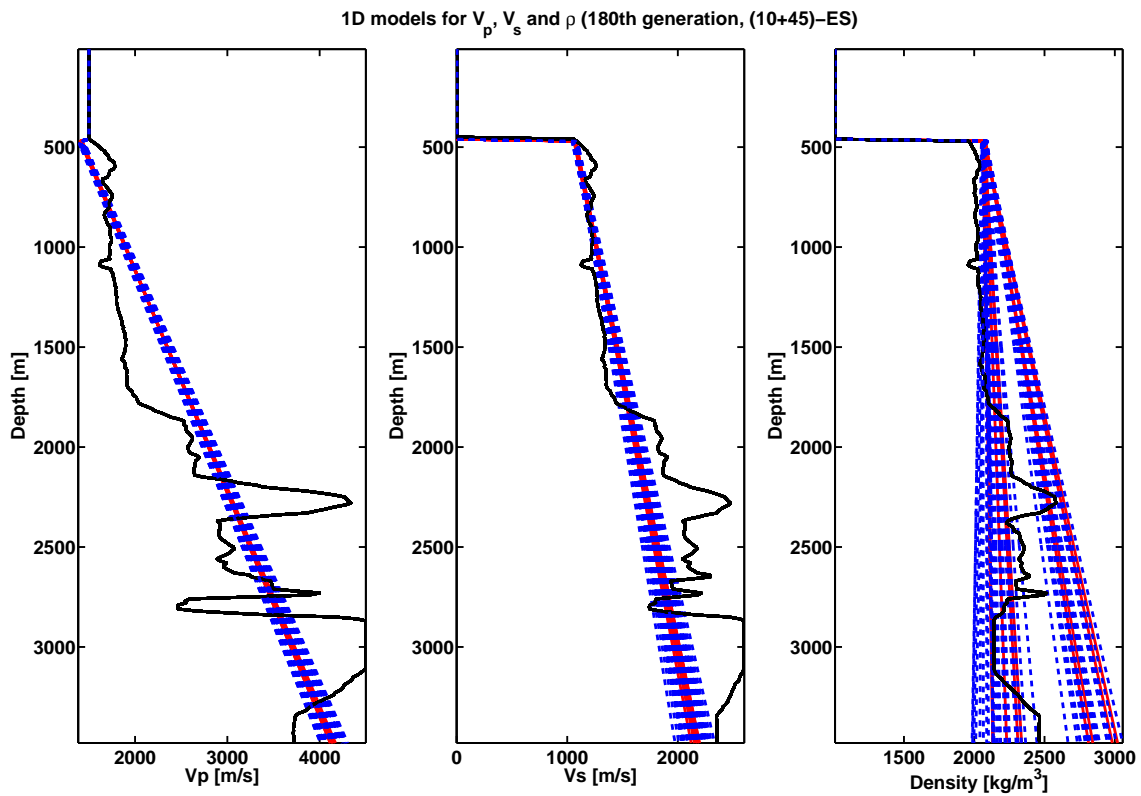


Figure 7.6: Models of the parental population (red models) and offspring population (blue dashed models) compared with the average true model (black model) for the 180th generation using a (10+45)-ES.

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

---

Fig. 7.7 - 7.9 show the parameter space for  $V_p$ ,  $V_s$  and  $\rho$  covered by the (10+45)-ES. The black dots in the upper pictures denote all tested models, including the rejected ones, of the parental and offspring populations. In total the misfit function of 9900 forward models are calculated, which requires about 1 day of computation time on 30 CPUs of an HP XC3000. The bottom picture shows a scatterplot of all models with a misfit value of less than  $1.7e - 17$ . The colors of each dot denote the corresponding values of the misfit function. Note the clustering of models near the minima of the misfit function. In case of the P-wave velocity (Fig. 7.7, bottom) 3 clusters of models exist which cover the area between  $V_{p0} = 1250$  and  $1750$  m/s with gradients ranging from  $g_{V_p} = 0.55$  to  $1$  1/s. These clusters form long elongated narrow parallel valleys of minima of the misfit function. The lowest misfit values can be found in cluster 2. A 4th cluster extends between  $V_{p0} = 1250$  and  $1500$  m/s with a few models having larger values of up to  $V_{p0} = 1800$  m/s. The gradients are quite small ranging from  $g_{V_p} = 0.4$  and  $0.6$  1/s. A few models also have values down to  $g_{V_p} = 0.2$  1/s. In contrast to the clusters 1-3, cluster 4 is shaped more spherical but also shows a dominant valley structure. The clusters 1-3 are models which fit the global true average 1D Marmousi model, while cluster 4 mainly fits the upper 1.8 km of the model with very small gradient values. For the S-wave velocity (Fig. 7.8, bottom) the minima are quite large. Nevertheless two solution clusters can be identified. Cluster 1 is elongated in shape and covers the area between  $V_{s0} = 1000$  and  $1200$  m/s with gradients ranging from  $g_{V_s} = 0.2$  to  $0.7$  1/s. The second cluster is smaller and more spherical with models ranging from  $V_{s0} = 1000$  and  $1200$  m/s and gradients ranging from  $g_{V_s} = 0.0$  to  $0.2$  1/s. In contrast to the P-wave velocity clusters the S-wave velocity clusters show no distinct structure. The misfit minima are rather diffuse. Cluster 1 covers the models fitting the global trend of the average 1D Marmousi model, while cluster 2 fits the upper part of the model. The parameter space of the density (Fig. 7.9, bottom) shows 6 clusters. The clusters 1-4 are very small and distinct, with a few models located near the points  $(\rho_{p0}, g_\rho)_1 = (1950 \text{ kg/m}^3, 0.9 \text{ kg/m}^4)$ ,  $(\rho_{p0}, g_\rho)_2 = (1800 \text{ kg/m}^3, 0.9 \text{ kg/m}^4)$ ,  $(\rho_{p0}, g_\rho)_3 = (1950 \text{ kg/m}^3, 0.6 \text{ kg/m}^4)$  and  $(\rho_{p0}, g_\rho)_4 = (1800 \text{ kg/m}^3, 0.6 \text{ kg/m}^4)$ . Cluster 5 is spherical and covers the area between  $\rho_0 = 2000 \text{ kg/m}^3$  and  $2350 \text{ kg/m}^3$  with gradients from  $g_\rho = 0.3 \text{ kg/m}^4$  to  $0.6 \text{ kg/m}^4$ . Cluster 6 is elongated and extends between  $\rho_0 = 1750 \text{ kg/m}^3$  and  $2200 \text{ kg/m}^3$  with gradients from  $g_\rho = -0.1 \text{ kg/m}^4$  to  $0.4 \text{ kg/m}^4$ . It is not really surprising, that the ambiguity for the density is very high. Even worse the models from cluster 1-5 are all not fitting the model at all. Only the solutions from cluster 6 with small gradient values near zero are plausible.

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

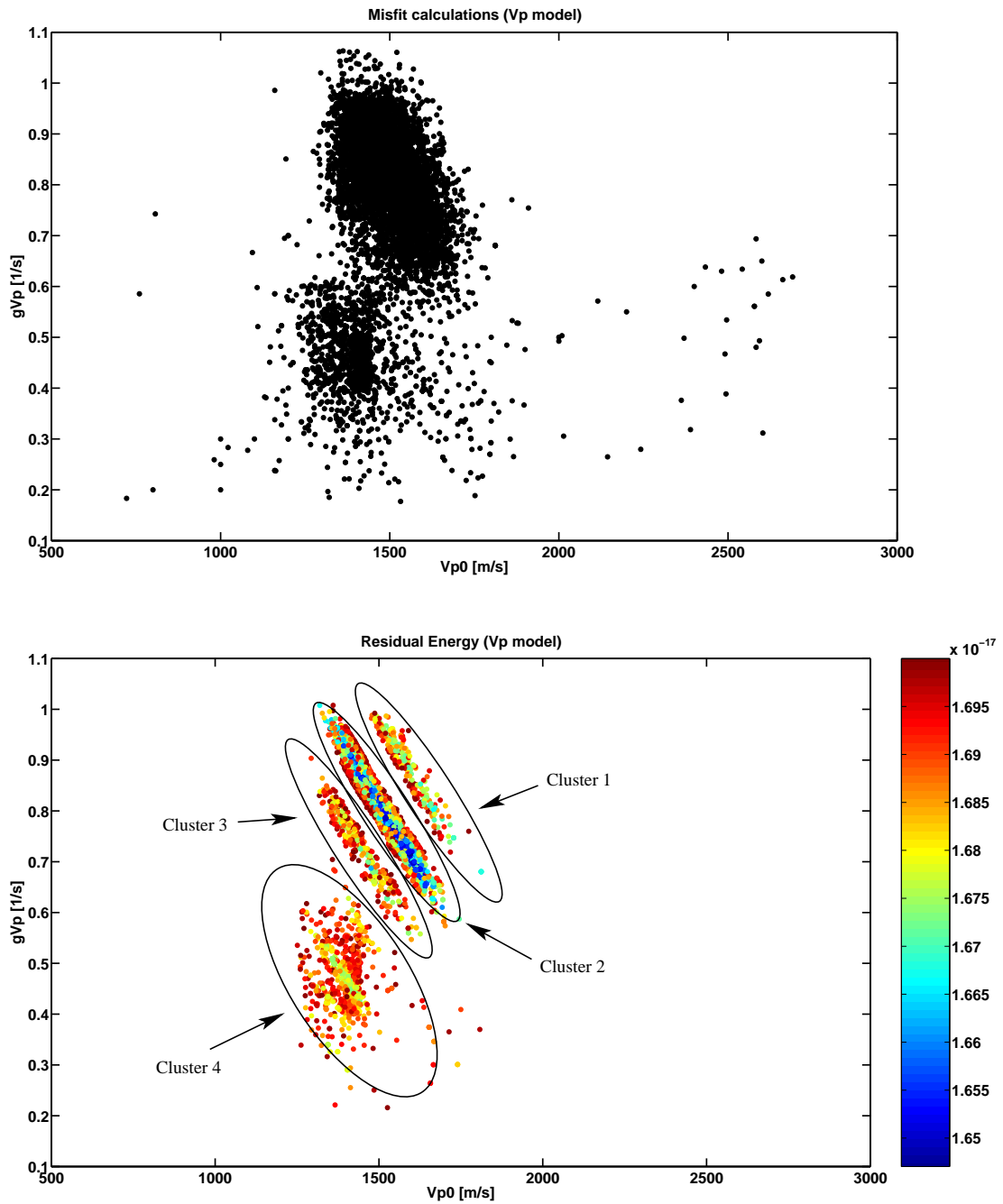


Figure 7.7: Scatterplot of all P wave velocity models evaluated by the (10+45)-ES (top) and all models with an  $L_2$ -norm smaller than  $1.7e - 17$ .

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

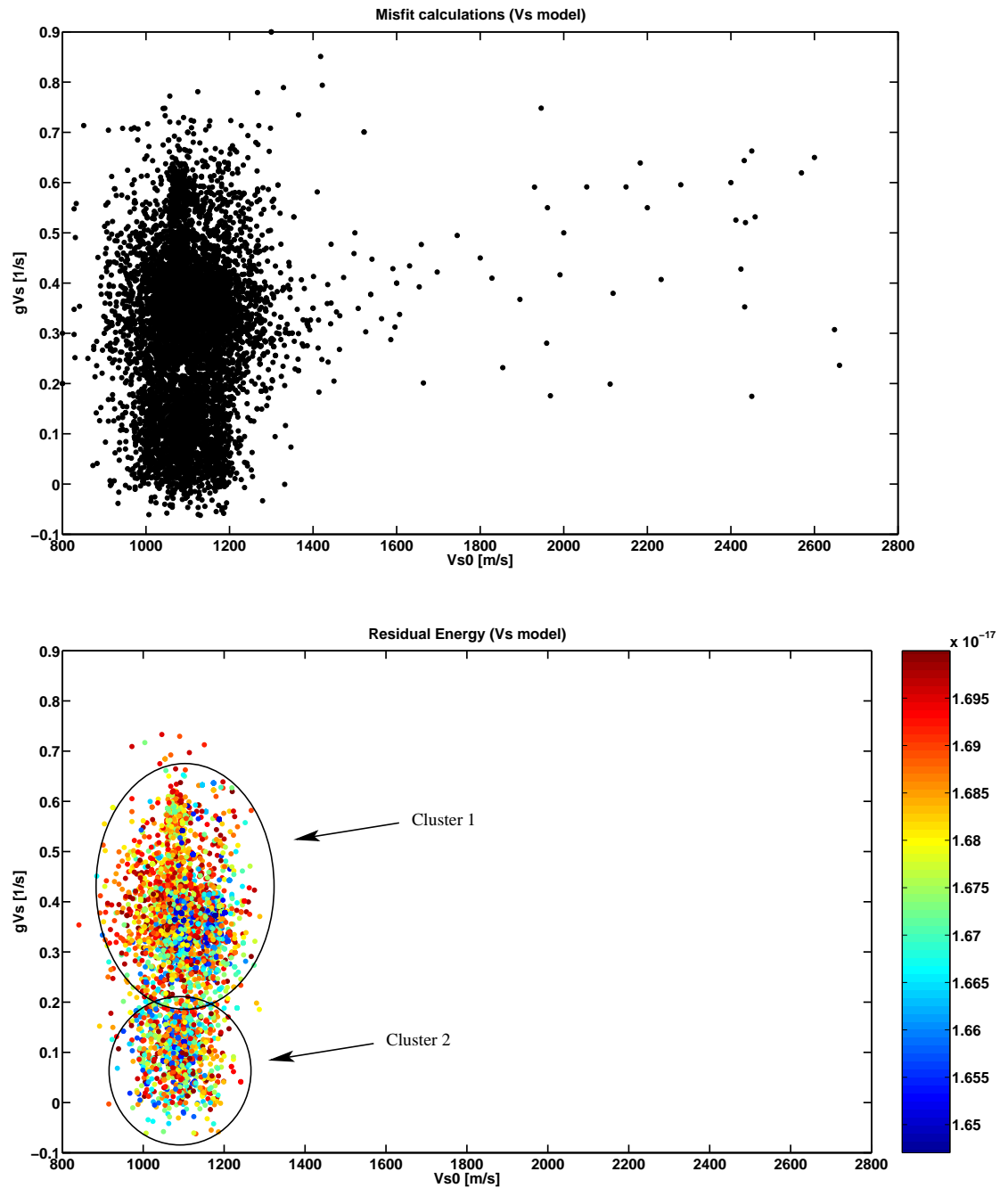


Figure 7.8: Scatterplot of all S wave velocity models evaluated by the (10+45)-ES (top) and all models with an  $L_2$ -norm smaller than  $1.7e - 17$ .



### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

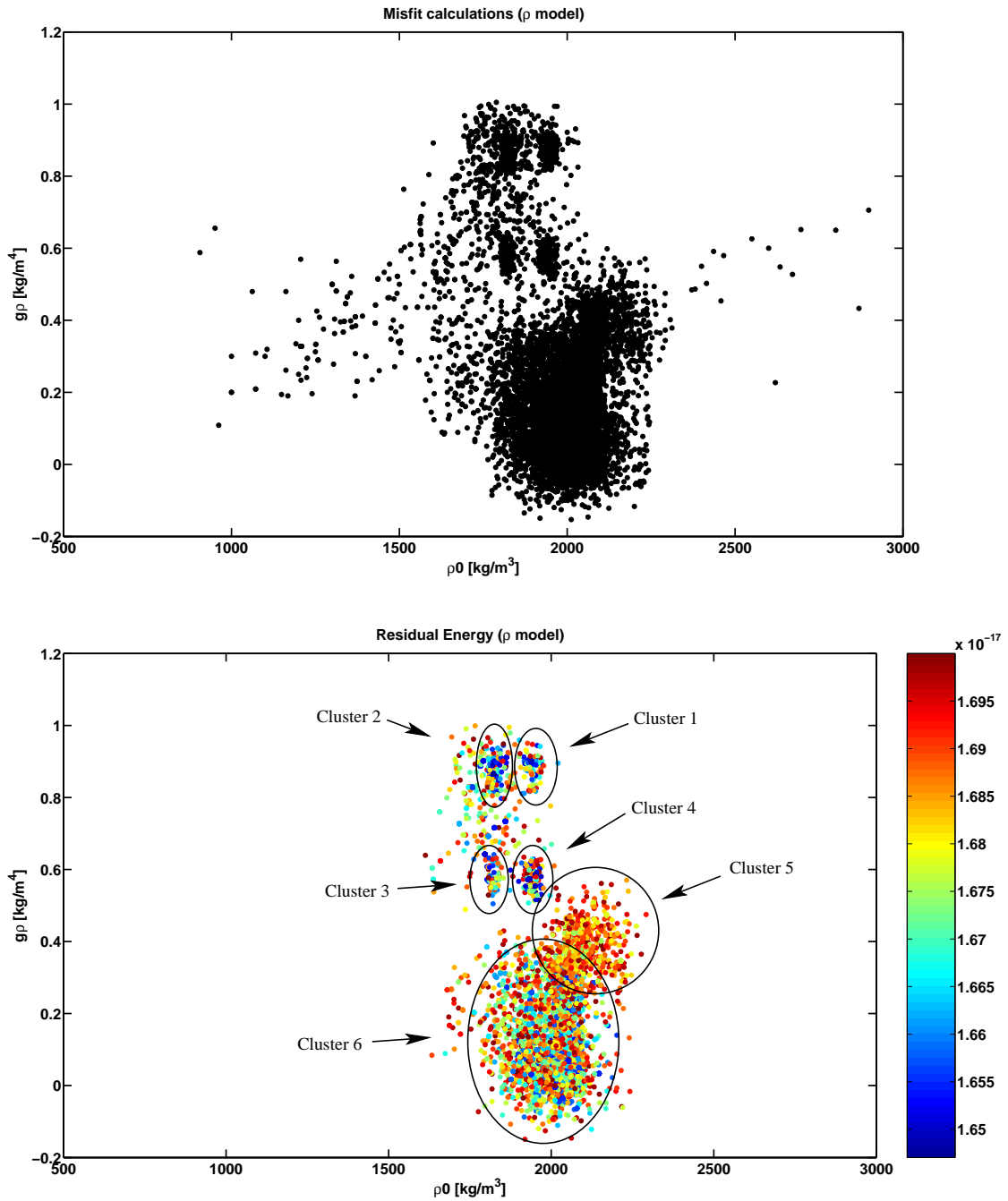


Figure 7.9: Scatterplot of all density models evaluated by the (10+45)-ES (top) and all models with an  $L_2$ -norm smaller than  $1.7e - 17$ .

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

---

From all the tested 1D models the TOP 10 with the smallest values of the misfit function are listed in table 7.2. For a special reason I also picked the model on place 16. While the top 10 models fit the global average true model more or less well, model 16 fits only the upper part extremely well. I have chosen three ES models as 1D starting models for the FWT: Model 1, 3 and 16, in the following chapters called ES-model 1, ES-model 3 and ES-model 16. In Fig. 7.10 the model parameters  $V_p$ ,  $V_s$  and  $\rho$  for these three models are plotted as a function of depth. The choice of ES-model 1 (colored in green) is obvious, it is the model with the lowest value of the misfit function out of 9900 other models. It fits the global average  $V_p$  and  $V_s$  very well, but it has the disadvantage, that the density model is completely wrong, despite the value of the density at the seafloor. Therefore ES-model 3 (colored in blue) is chosen which has  $V_p$  and  $V_s$  models nearly identical with the ES-model 1, but the global density model is fitted very well. Finally ES-model 16 (colored in red) fits not the global, but only the upper part of the average Marmousi model. The deviations of the ES-models from the average Marmousi model as a function of depth are plotted in Fig. 7.11. for the parameters  $V_p$  (left),  $V_s$  (center) and  $\rho$  (right). The green, blue and red models are ES-model 1, ES-model 3 and ES-model 16, respectively. Notice the strong misfit of the ES-models 1 and 3 in the upper part of the model down to a depth of 1.8 km, with a maximum error of 500 m/s for  $V_p$  and an error of 250 m/s for  $V_s$ . The maximum error of the density model for ES-model 1 is 500 kg/m<sup>3</sup>, while the ES-model 3 fits the density very well. ES-model 16 shows nearly no deviations for all three material parameters in the upper part of the model. The fits of the lower parts of the simple Marmousi model are poor for all 3 models. The maximum misfit for  $V_p$  are -1000 m/s for ES-model 1 and 3 and -2000 m/s for ES-model 16. For the  $V_s$  model the deviations are a little bit smaller ranging from -1000 m/s for ES-model 16 and -500 m/s for ES-model 1 and 3. The density misfits for all three models are comparable with  $\pm 500$  kg/m<sup>3</sup>. Now we have a few plausible starting models to test if the FWT algorithm can fit these models as well as the smooth true model described in chapter 6.2.3.

Model No.	$V_{p0}$ [m/s]	$V_{s0}$ [m/s]	$\rho_0$ [kg/m <sup>3</sup> ]	$g_{vp}$ [1/s]	$g_{vs}$ [1/s]	$g_\rho$ [kg/m <sup>4</sup> ]	misfit function
1	1420	1070	2071	0.9005	0.3684	0.3035	1.650396e-17
2	1404	1068	2077	0.9194	0.3682	0.2546	1.650519e-17
3	1415	1067	2074	0.9059	0.3596	0.0522	1.650820e-17
4	1415	1064	2071	0.9077	0.3755	0.3026	1.650859e-17
5	1427	1060	2066	0.8928	0.3716	0.3168	1.650923e-17
6	1425	1064	2058	0.8951	0.3767	0.2893	1.650988e-17
7	1409	1065	2070	0.9149	0.3700	0.0887	1.651005e-17
8	1427	1070	2075	0.8932	0.3722	0.0783	1.651072e-17
9	1411	1069	2066	0.9700	0.3605	0.0233	1.651085e-17
10	1428	1077	2070	0.8929	0.3449	0.2857	1.651155e-17
16	1657	1241	2010	0.2640	0.1250	0.0167	1.692980e-17

Table 7.2: The TOP 10 1D models estimated by the (10 + 45) evolution strategy. Model 16 will also become important later on. The highlighted models are used as starting models for the elastic FWT.

## CHAPTER 7. 1D STARTING MODEL ESTIMATIONS WITH EVOLUTION STRATEGY

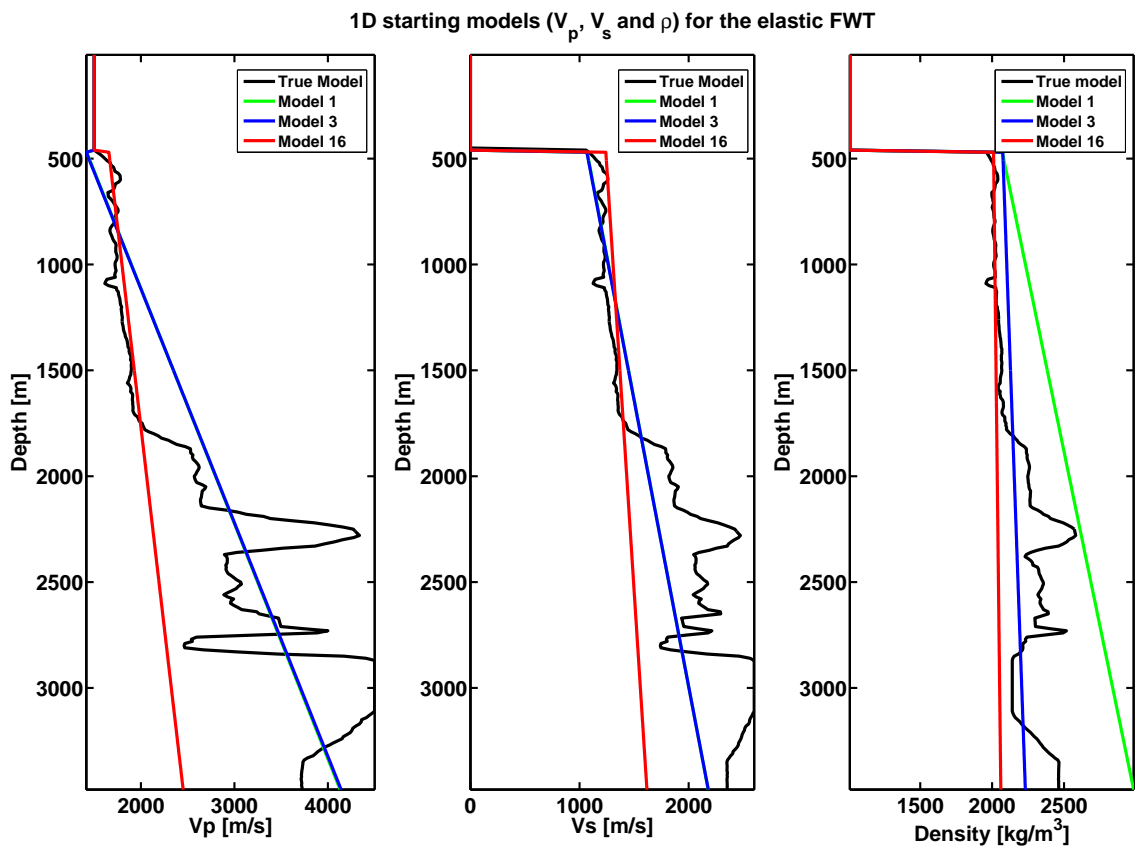


Figure 7.10: 1D starting models for the elastic FWT estimated by the (10+45) evolution strategy: ES-model 1 (green), ES-model 3 (blue), ES-model 16 (red) and the average simple Marmousi2 model (black).

### 7.3. STARTING MODEL ESTIMATIONS BY (10+45)-ES

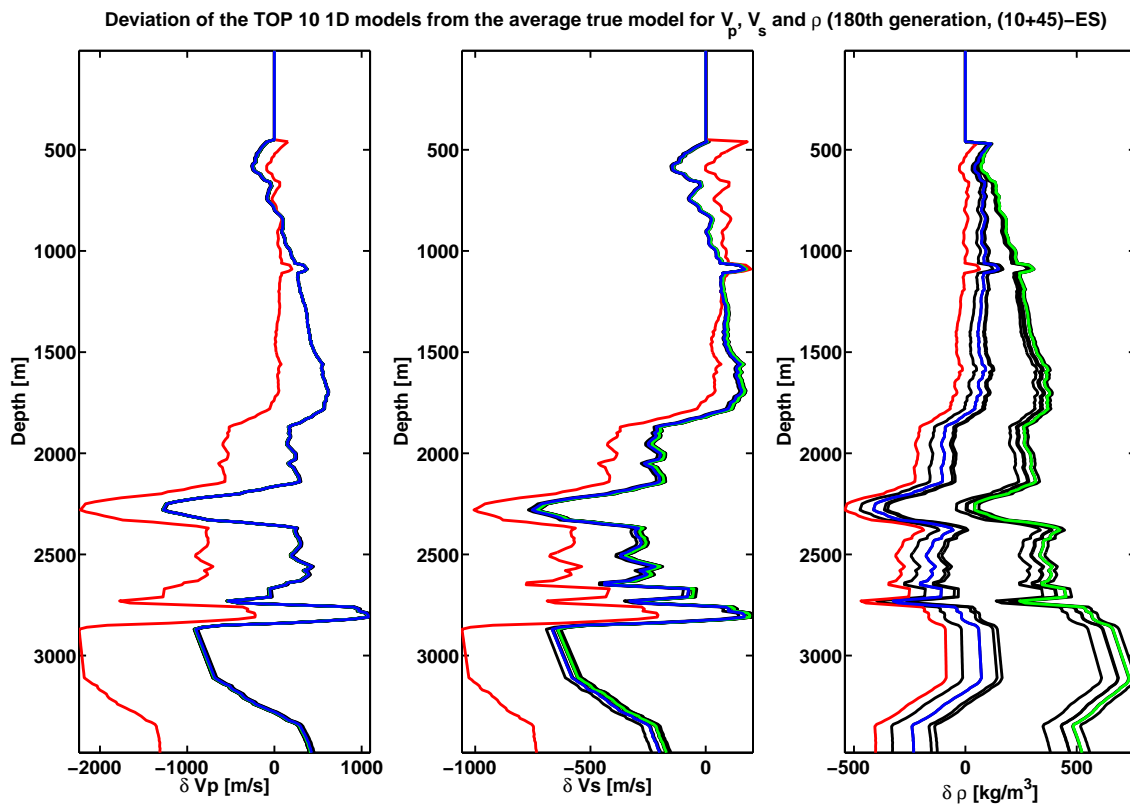


Figure 7.11: Deviations of the TOP 10 ES-models from the average simple Marmousi model as a function of depth for the parameters  $V_p$  (left),  $V_s$  (center) and  $\rho$  (right). The green, blue and red models are ES-model 1, ES-model 3 and ES-model 16, respectively.



# 8

## Reducing the nonlinearity of the elastic inverse problem I: Time Windowing

Before applying the FWT to the 1D starting models estimated in the last section I want to discuss my experience with the application of time windows in FWT. Fig. 8.1 shows the initial residuals of shot 25 (y-component) for model 1 estimated in the last chapter by the ES (left and center). For comparison the residuals of the perfect starting model presented in chapter 6.2 are shown on the right. At first glance the data residuals of the perfect starting model and the ES-model 1 seem to be very similar, despite a few minor differences. But as we have seen in the last chapter the ES model shows large model misfits at greater depth (Fig. 7.11). To test the influence of this strong misfit on the FWT result two different approaches are used:

1. Backpropagate the whole data residuals during the ES-model 1 inversion (Fig. 8.1, center).
2. Apply a time window on the data residuals of the ES-model 1 data (Fig. 8.1, left) to invert only the upper part of the model and not the strong reflections from the deeper parts of the simple Marmousi2 model. The time window is very simple, it sets all amplitudes after a defined time limit  $t_1$ , in this case 2.0 s, to zero. To avoid ringing the amplitudes are damped using a Gaussian taper function:

$$P = \begin{cases} 1 & \text{if } t_0 = 0.0 \text{ s} \leq t < t_1 = 2.0 \text{ s} \\ \exp(-\frac{1}{2}(a\frac{t-t_1}{\Delta l/2})^2) & \text{if } t_1 = 2.0 \text{ s} \leq t \leq t_2 = 2.275 \text{ s} \\ 0 & \text{if } t_2 = 2.275 \text{ s} < t \end{cases}$$

with  $a = 3.0$ ,  $\Delta l = 0.275 \text{ s}$ .

## CHAPTER 8. REDUCING THE NONLINEARITY OF THE ELASTIC INVERSE PROBLEM I: TIME WINDOWING

---

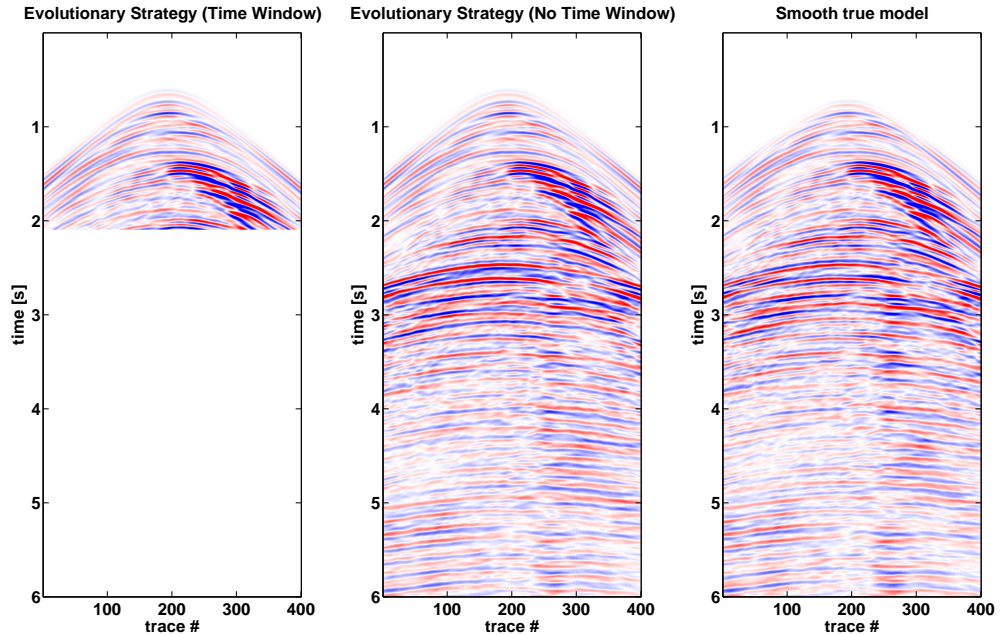


Figure 8.1: Initial data residuals for the ES-model 1 with (left) and without the application of a time window Eq. (8.1) (center) (shot 25, y-component). For comparison the initial data residuals of the smooth true model (right) is also shown.

Before testing the time window on the ES-model 1 I want to test the influence of the time window on the FWT result for the smooth starting model. Fig. 8.2 shows a comparison of the FWT result for the P-wave velocity with (left) and without (right) the application of the time window. The obvious difference between these results are the artefacts surrounding gas reservoir A when applying the time window, especially the diffractions from the edges of the reservoir seem to be imaged at the wrong positions. Additionally the resolution in the upper parts of the model are comparable, but decreases a little bit with depth when applying the time window. Both problems can be explained by the missing deeper reflections and internal multiples which contain information about these deeper layers. In conclusion the time window can produce some artefacts and decreases the resolution of the deeper model parts a little bit, but the overall impact on the inversion result is very small, at least in this case. Fig. 8.3 - 8.4 demonstrate that the time windowing is necessary. These figures show the FWT results of all three elastic parameters of the ES-model 1 with (left) and without the application of the time window (center). For comparison the results for



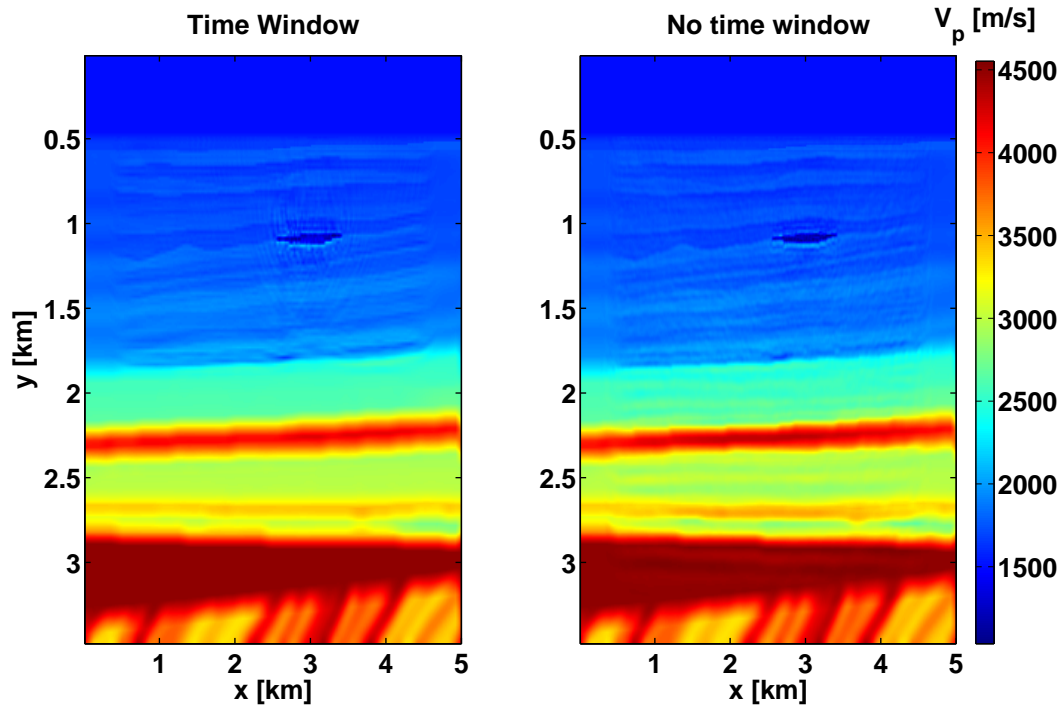


Figure 8.2: FWT results for the P-wave velocity of the simple Marmousi2 model using the smooth starting model with (left) and without the application of a time window (right).

the perfect smooth starting model without time windowing is shown on the right. It can be clearly seen, that the FWT fails to converge if no time window is applied. The gas reservoir A is imaged well, but all the layers which are present in the "perfect" inversion result are missing, instead strong artefacts are visible. In the deeper parts of the model the FWT algorithm tries to fit the strong amplitude reflections from the strong impedance contrast in 2.25 km depth. But due to the inaccurate model parameters in the upper part of the model this reflector is misplaced and therefore all traveltimes, phases and amplitudes of reflections from this interface are imaged at erratic positions in space. This nonlinearity is reduced, if the time window is applied. The corresponding FWT results do not suffer from strong artefacts anymore. The gas reservoir A is clearly visible as well as the layers and water wet sands. Some imaging problems occur above and below the gas reservoir A which are due to imaging errors of the reservoir during the first iterations. When comparing the time

## CHAPTER 8. REDUCING THE NONLINEARITY OF THE ELASTIC INVERSE PROBLEM I: TIME WINDOWING

---

windowed ES-model 1 results with the results of the "perfect" starting models the layers in the ES result doesn't fit the distribution of the real material properties, as in the case of the "perfect" starting model. Instead only the interfaces between the layers seem to be imaged. It looks more like the result of a migration than a FWT. This problem can be partly explained by the time window. With the time window we have avoided to fit the strong reflector, which can not be explained by the simple 1D starting model, but we also deleted all the transmission information which was accumulated by this strong reflections while passing the sediment layers above the reflector. The weak reflections from the low velocity sediment layers with nearly no impedance contrasts are only capable to image the interfaces. Therefore we now have the dilemma, that we need the strong reflection from the high impedance layers for the FWT, but also have to avoid them because the high impedance layers can not be explained by the starting model and lead to instabilities of the inversion scheme. But this is only one part of the problem, the other will be explained in the next chapter.

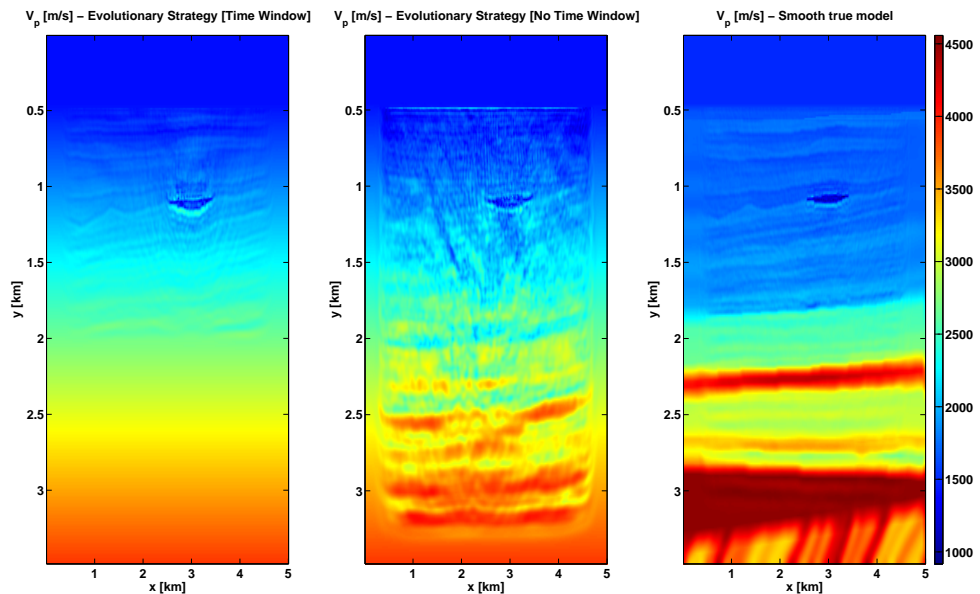


Figure 8.3: FWT results for the P-wave velocity using the ES-model 1 with time window (left) and without the time window (center) and the smoothed true starting model without a time window (right).

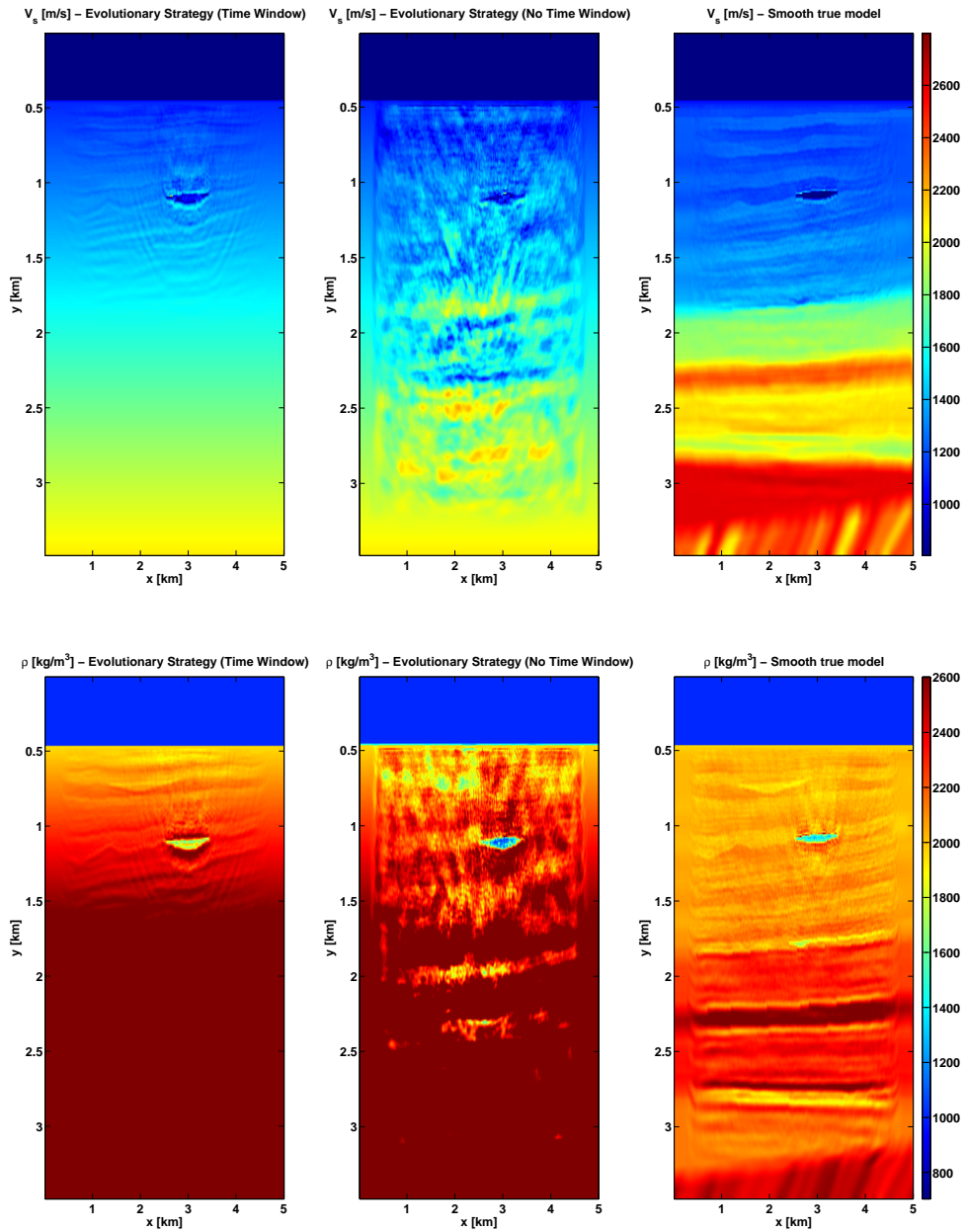


Figure 8.4: FWT results for the S-wave velocity (top) and density (bottom) using the ES-model 1 with time window (left) and without the time window (center) and the smoothed true starting model without a time window (right).



# 9

## Reducing the nonlinearity of the elastic inverse problem II: Frequency Filtering

The shape of the objective function in the elastic inversion problem shows a strong dependence with the frequency. At lower frequencies the objective function is smooth, while a lot of local minima are present at higher frequencies. Therefore it is common sense, that the inversion should start at low frequencies and add higher frequency content after a given number of iteration steps until the whole frequency content of the data is inverted. Depending on the size of the problem this "low" frequency content is covered by different frequencies. For hydrocarbon reflector seismics most of the information about the material parameters seem to be included in the very low frequency content (up to approximately 5 Hz) of the seismic data (Choi et al. (2008a)). If this frequency range is not covered in the dataset the corresponding starting model has to contain this information. That is one reason why the FWT result of the Gaussian filtered true model converges very well and shows a lot of details even for the high frequency data with a dominant frequency of 10 Hz (Fig. 9.1, (b)). The 1D starting model estimated by the ES does not contain this low frequency information and therefore the FWT result is very poor (Fig. 9.1, (a)). The quality of the FWT result for the 1D model should improve, if data is inverted separately for different frequency windows. To test this assumption the FWT is speparated in two parts.

1. In part I a low pass filter with the following parameters

$$P = \begin{cases} 1 & \text{if } f_0 = 0.0 \text{ Hz} \leq f < f_1 = 5.0 \text{ Hz} \\ \exp(-\frac{1}{2}(a\frac{f-f_1}{\Delta f/2})^2) & \text{if } f_1 = 5.0 \text{ Hz} \leq f \leq f_2 = 10.0 \text{ Hz} \\ 0 & \text{if } f_2 = 10.0 \text{ Hz} < f \end{cases}$$

with  $a = 1.6$  and  $\Delta f = 5 \text{ Hz}$  is applied to the fourier transformed residual seis-

## CHAPTER 9. REDUCING THE NONLINEARITY OF THE ELASTIC INVERSE PROBLEM II: FREQUENCY FILTERING

---

mograms. For frequencies below 5 Hz all the frequency content is passed. Between  $f = 5$  and 10 Hz a Gaussian taper is applied and above  $f = 10$  Hz all the frequency data is zeroed.

2. In part II the FWT result of part I is used as a starting model for an inversion of the whole frequency content of the data residuals.

In Fig. 9.1 the FWT results of the different approaches are compared for the P-wave velocity model. The inversion result after 90 iterations without any frequency filtering, but the application of a time window (chapter 8), for ES-model 1 shows only the main interfaces and is dominated by the starting model (Fig. 9.1, (a)). The smooth starting model already contained the low frequency information and therefore the result is perfect without frequency filtering (Fig. 9.1, (b)). The result after part I of the frequency filtering approach for ES-model 1 is shown in Fig. 9.1, (c). The long wavelength structures are fitted very well. Not only the interfaces between the different layers are visible, but also the material parameters within the layers are imaged well. There are still a few imaging problems above and below the dominant gas reservoir A. The parts of the model below a depth of 1.5 km are obviously not changed due to the application of the time window. The inversion result after part II of the frequency inversion is shown in Fig. 9.1, (d). Especially the layer in 1 km depth is now imaged better, compared with the inversion result without any frequency filtering. The layer is not only bounded by two interfaces, but also contains the correct material parameters. Some artefacts surrounding the gas reservoir A are also suppressed. The overall result of the frequency dependent FWT shows some minor improvements, but the differences are not really spectacular. This can be explained by the already low center frequency of 10 Hz for the source signal, therefore enough frequency content below 5 Hz should be already present. To illustrate what happens, if this frequency content is missing I apply a highpass frequency filter on the data set which only passes frequencies above 5 Hz. Therefore most of the information about the material parameters should be vanished. After the application of the FWT on this highpass filtered dataset, using a time window and the ES-model 1 as starting model, which also contains no low frequency information about the model, the P-wave velocity results look like Fig. 9.2 (left). When comparing these results with the image achieved by applying the two step frequency inversion, they are really devastating. Some layers down to a depth of 1 km could be correlated with the true model, but everything below 1 km is imaged completely wrong, even the dominant gas reservoir A is not visible. This result proves that the frequency content below 5 Hz is essential for a successful FWT application. Now we have the tools to apply the FWT to the 1D starting models estimated by the ES.

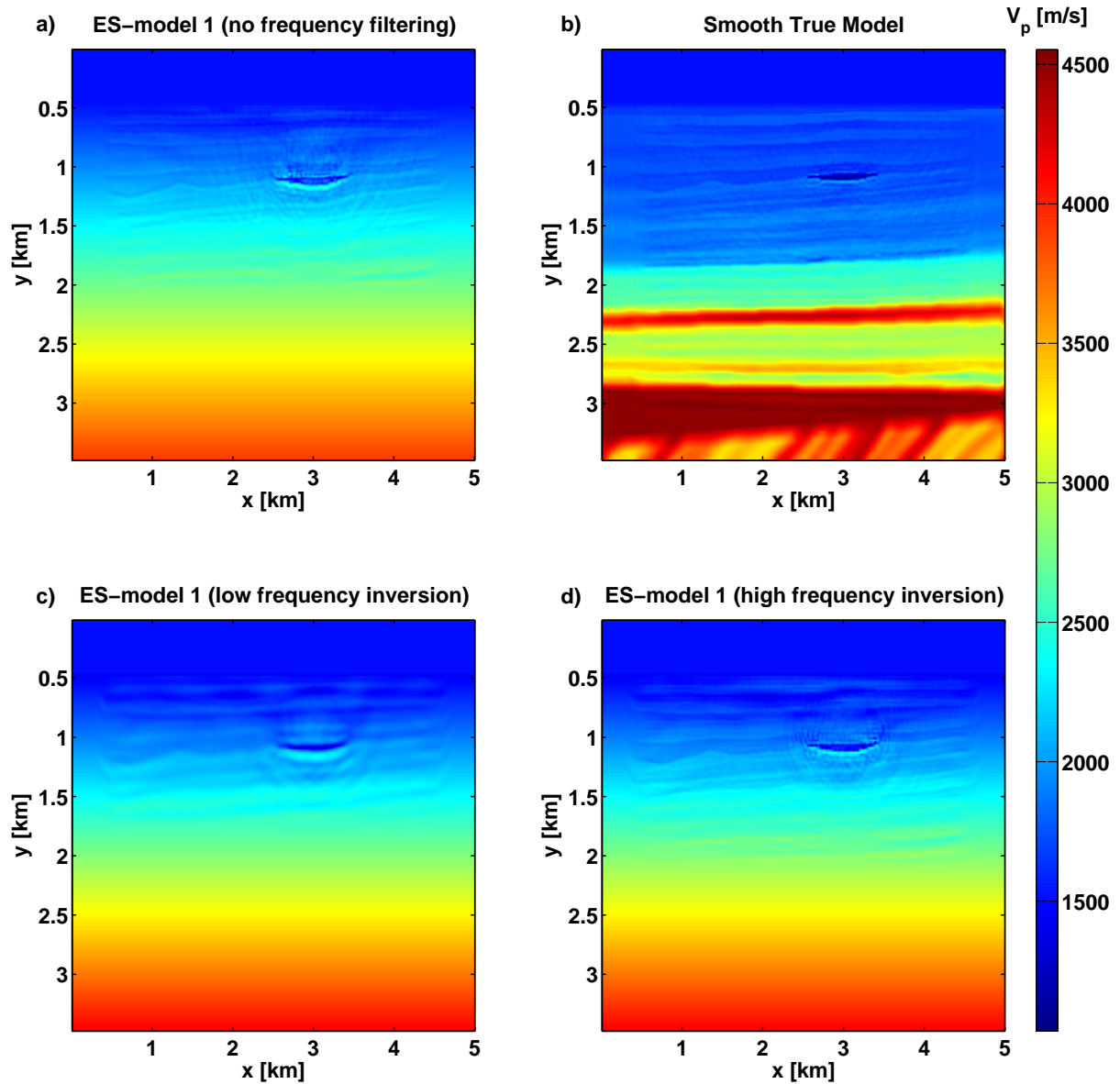


Figure 9.1: FWT results for the P-wave velocity model after 90 iterations using (a) no frequency filtering, but time windowing for ES-model 1, (b) no frequency filter, no time window for the smooth starting model, (c) low frequency filtering and time windowing for ES-model 1, (d) no frequency filtering and time windowing for ES-model 1.

## CHAPTER 9. REDUCING THE NONLINEARITY OF THE ELASTIC INVERSE PROBLEM II: FREQUENCY FILTERING

---

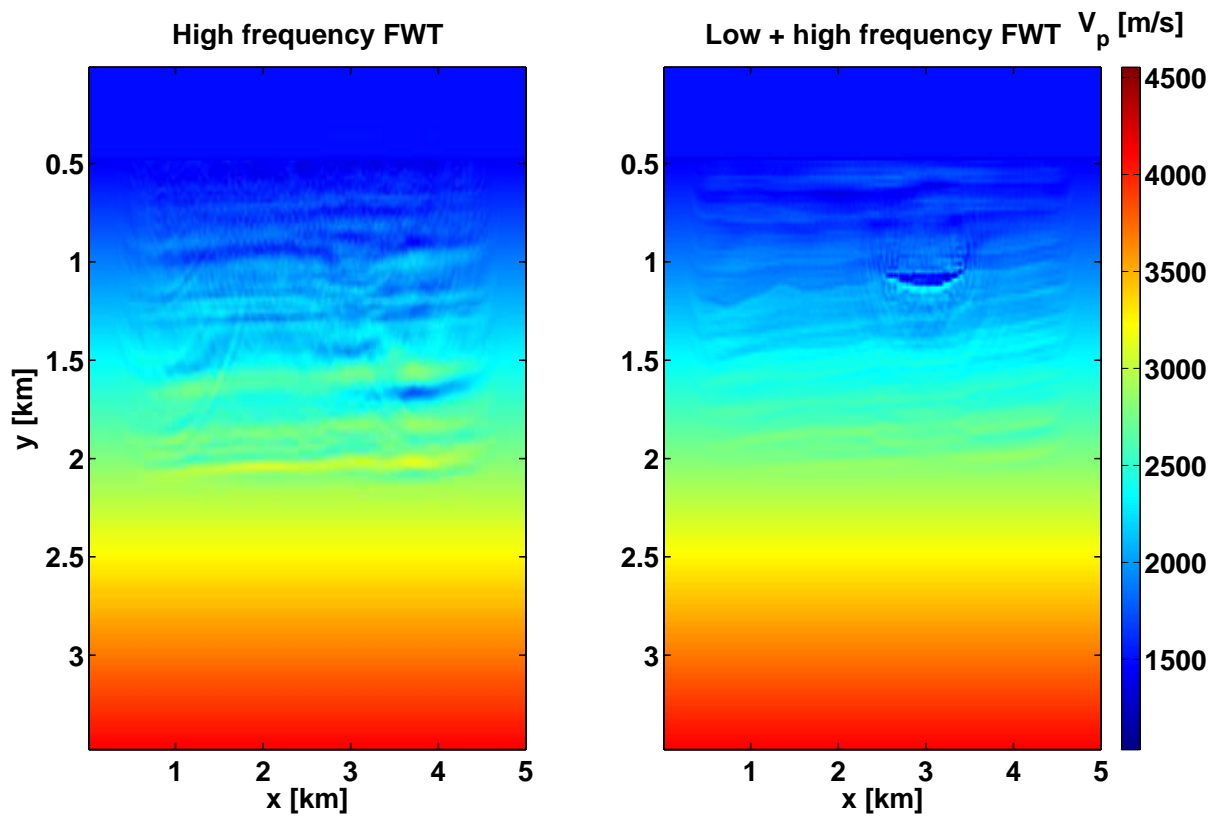


Figure 9.2: FWT results for the P-wave velocity model after 90 iterations using only high frequency information and time windowing for ES-model 1 (left) and application of the two step frequency inversion, with time window for ES-model 1 (right).



# 10

## The Influence of the starting model

With the knowledge from the last two chapters about reducing the nonlinearities of the elastic inverse problem we can apply the FWT to the 3 models estimated by the evolution strategy. The parameters for the time windowing and frequency filtering are the same as described in chapter 8 and 9. The FWT results of the ES-models 1, 3 and 16 for the seismic velocities  $V_p$ ,  $V_s$  and the density  $\rho$  are shown in Figs. 10.1 - 10.3 ((a)-(c)). For comparison the results for the smooth starting model is also displayed (d). Additionally Fig. 10.4 shows a depth profile located at  $x_p = 3.0$  km for the starting model, the inversion result and the true model. The FWT results for the seismic velocities of ES-models 1 and 3 are nearly identical, despite some very minor differences (Fig. 10.1, Fig. 10.2, (a) and (b)). Especially in the upper part of the model, down to a depth of 0.7 km the layer properties of the inversion results are not imaged very well, when compared with the result of the smoothed starting model. The thickness of the layers is too large, while the P-wave velocity is too small. The S-wave velocity model is a little bit better resolved, maybe due to the shorter wavelength of the S-wave. Even the water wet sands can be resolved with the S-waves, while they are not visible in the  $V_p$  model. The part between 0.7 and 1.5 km depth is imaged very well for both material parameters, especially the gas reservoir A and the folded layers. Below a depth of 1.5 km the time window prohibited any update of the starting model and therefore any improvements of the model parameters. While ES-model 1 did not fit the true average density model at all, ES-model 3 could fit the density model of the Marmousi model very well. Therefore the FWT density results for the ES-models 1 and 3 differ very much. Even though the starting density model for ES-model 1 was wrong, the inversion result could resolve some structures in the most upper parts of the model (Fig. 10.3, (a)). Beside the dominant gas reservoir A, the water wet sands and some layers in the upper part of the model are imaged fairly well. However the resolution and accuracy decrease very fast with depth. The density starting model for ES-model 3 was much better

## CHAPTER 10. THE INFLUENCE OF THE STARTING MODEL

---

and therefore the resolution of the layers in the inversion result is also better when compared with ES-model 1, especially in the deeper parts of the model. However the estimated density values are larger than the density in the true model. Over all the FWT results are quite surprising, because the density should have a major impact on the impedance contrasts and as a result also on the amplitudes, but the velocity models are not influenced at all by the completely different density models. This leads to the conclusion, that ...

- the seismic velocity models are mainly deduced from the phase and not the amplitude information.
- the density inversion seem to be completely independent from the velocity inversion and may be mainly depending on the amplitude and not the phase information.

As we have seen in chapter 7 the ES-models 1 and 3 fitted mainly the global trend of the true average model, but deviate very much in the upper parts of the models. Therefore these models are not very useful when applying a time window to resolve the geology in the shallow parts of the models. However ES-model 16 fits only these upper model parts and should therefore be a perfect starting model when applying a time window. Indeed the FWT results for ES-model 16 displayed in Figs. 10.1-10.3, (c) are comparable with the results of the smooth starting model. Especially in the upper parts, down to a depth of 0.7 km, the resolution of the layers is much better when compared with the results of ES-models 1 and 3. The water wet sands, the gas reservoir A and even the oil sand B are well resolved. However the computational costs and man power to achieve this resolution in reality could be very high. You have to remember, that ES-model 16 deviates in the upper parts of the seismic velocity model by 200 m/s and the density only by 60 kg/m<sup>3</sup> from the true average model, so the requirements on the starting model are very high to really achieve a high resolution image of the underground, as predicted by the synthetic test problems. For this example ES-model 16 was picked out of 9800 models because I know the correct average model. In reality it may be unknown, if model 1, 5, 100, 5000 or 9799 correctly describes the geology of the underground, because the model exactly at the mathematically global minimum might be completely nonsense in terms of a correct physical description. Therefore a more extensive global search, hand picking analysis and a lot more FWT calculations are required to seriously invert real field data.

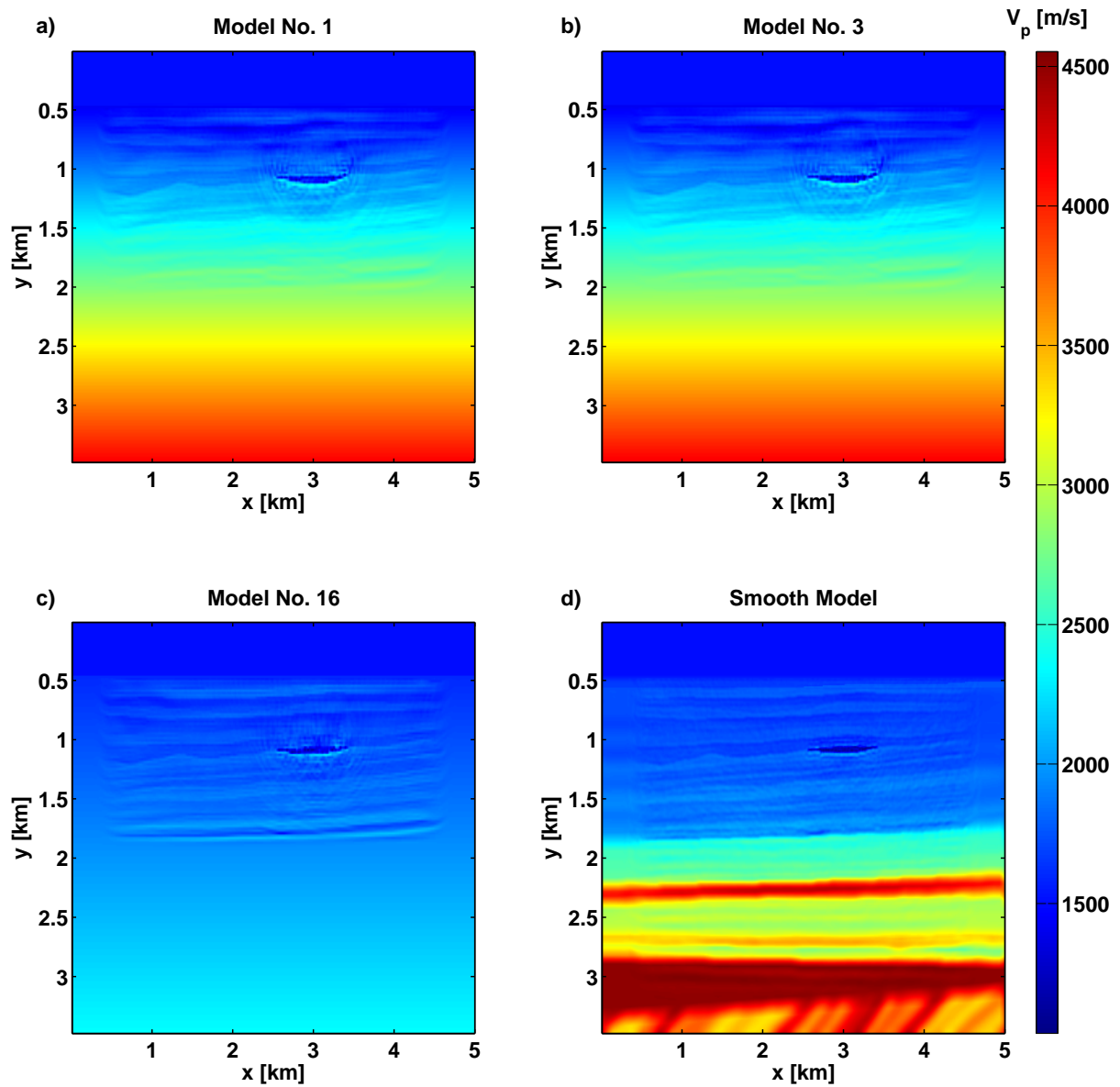


Figure 10.1: FWT results for the P-wave velocity  $V_p$  after 90 iterations using the 1D models ES-model 1 (a), ES-model 3 (b) and ES-model 16 (c) as a starting model for the inversion. For comparison the result of the smooth starting model is also shown (d).

## CHAPTER 10. THE INFLUENCE OF THE STARTING MODEL

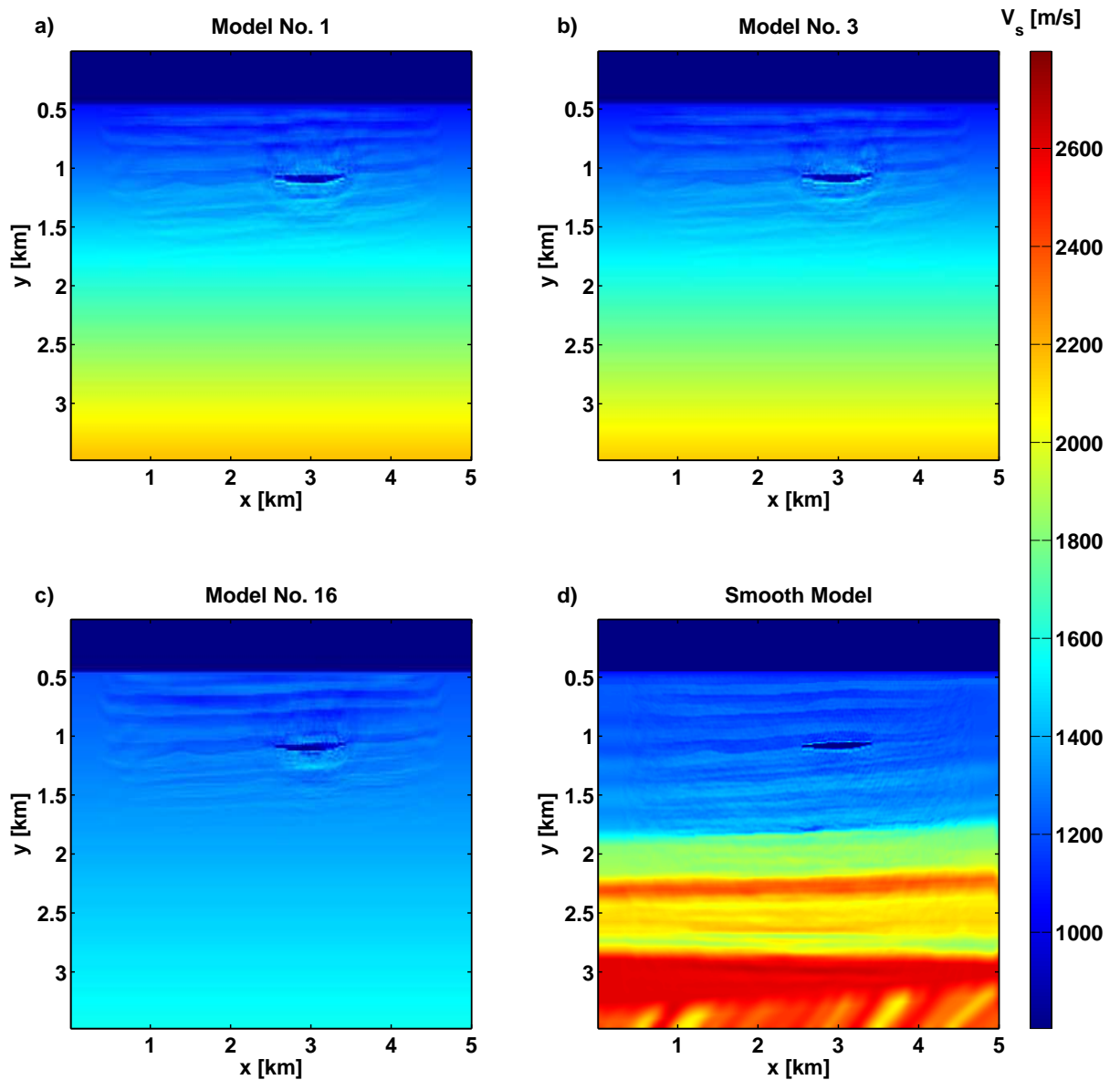


Figure 10.2: FWT results for the S-wave velocity  $V_s$  after 90 iterations using the 1D models ES-model 1 (a), ES-model 3 (b) and ES-model 16 (c) as a starting model for the inversion. For comparison the result of the smooth starting model is also shown (d).

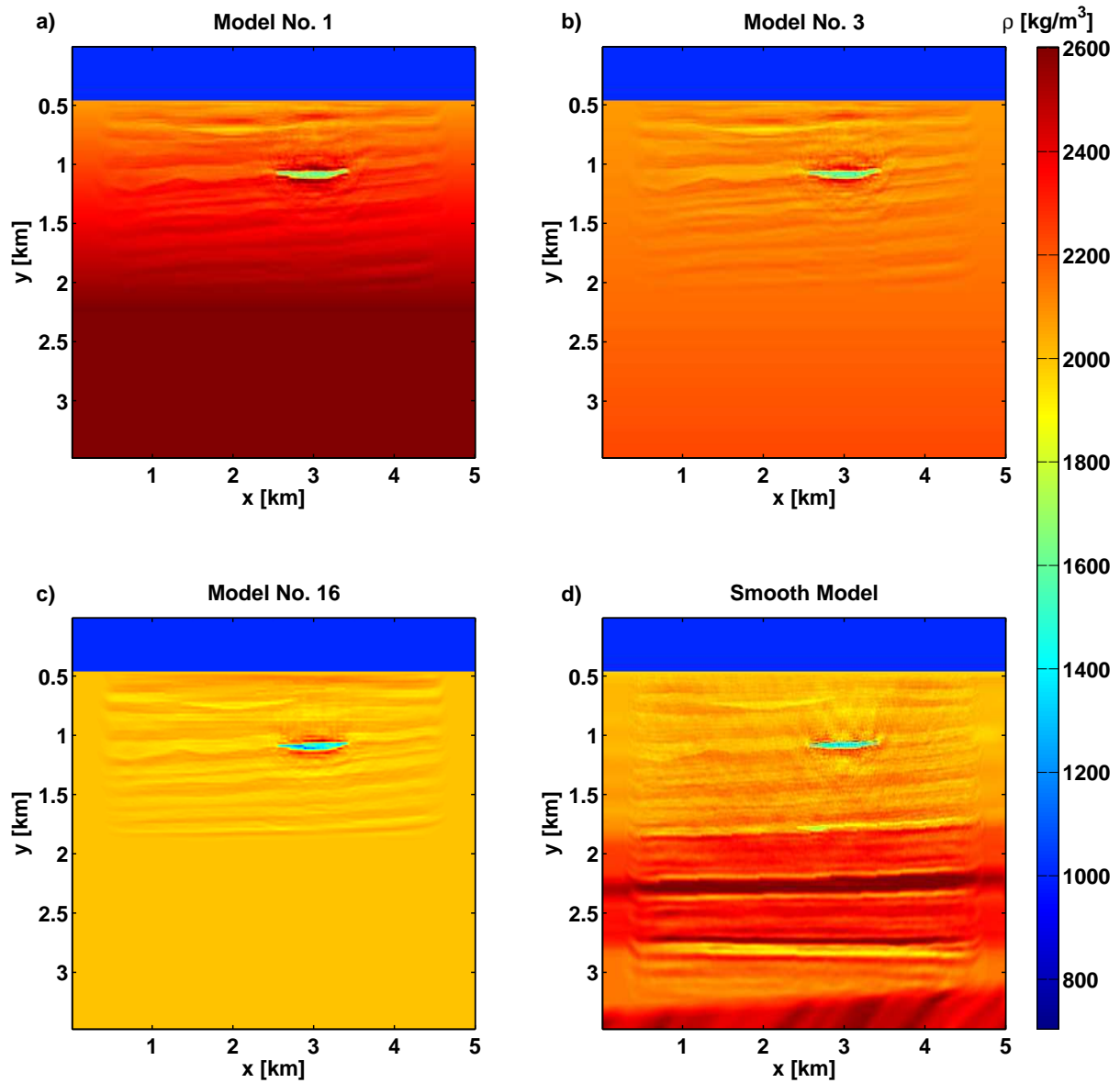


Figure 10.3: FWT results for the density  $\rho$  after 90 iterations using the 1D models ES-model 1 (a), ES-model 3 (b) and ES-model 16 (c) as a starting model for the inversion. For comparison the result of the smooth starting model is also shown (d).

## CHAPTER 10. THE INFLUENCE OF THE STARTING MODEL

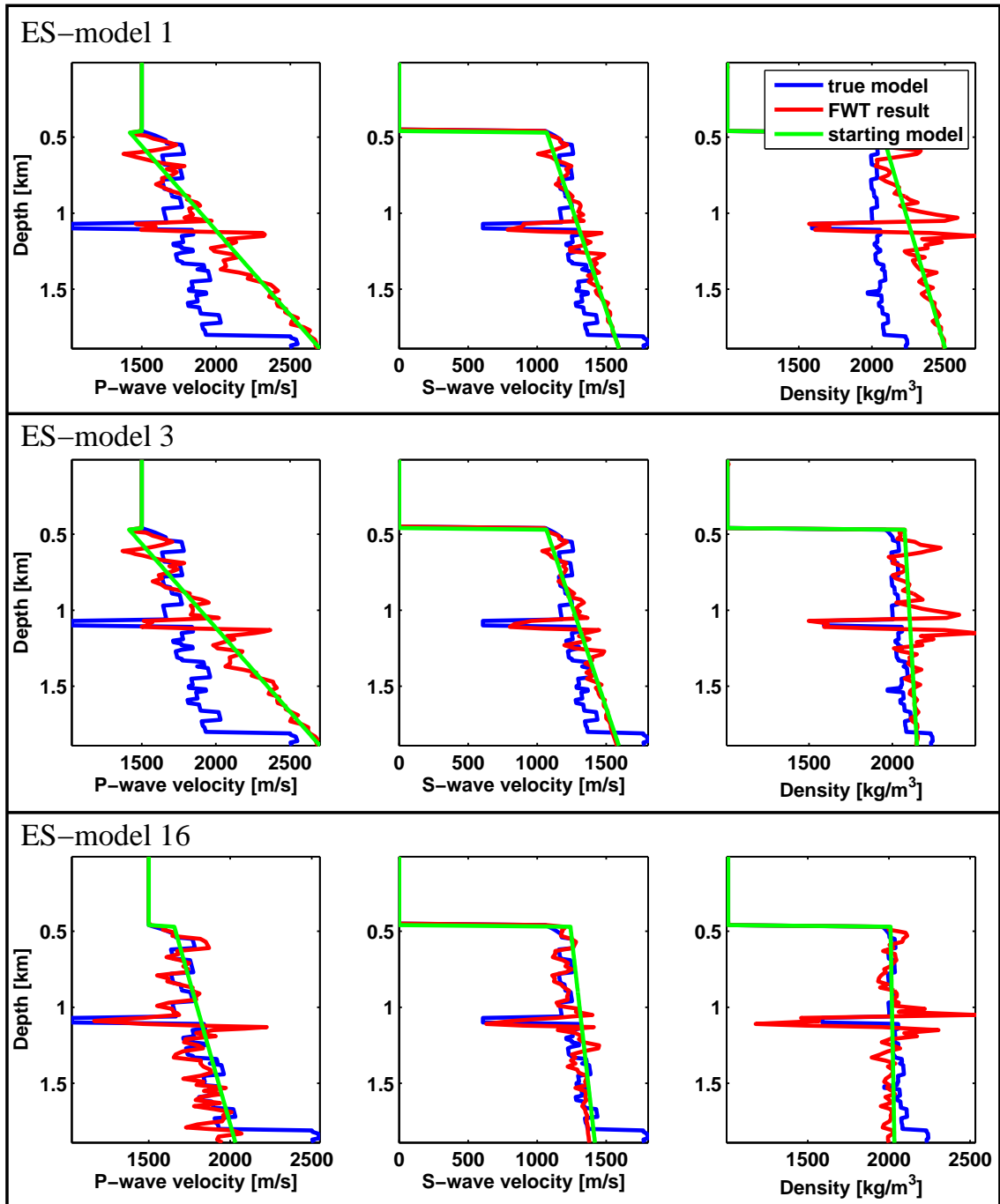


Figure 10.4: Depth profiles for the 1D ES-models 1, 3 and 16: P-wave velocity (left), S-wave velocity (center) and density (right).

# 11

## Inversion of 3D data with a 2D FWT code

In this chapter a fundamental problem of the 2D FWT code will be discussed, concerning the real data application. In the real field dataset the geometrical spreading is 3D. The energy excited by a point source in a homogeneous fullspace is approximately distributed along the surface of a sphere (Fig. 11.1, bottom). Lets imagine two wave fronts, which make 2 spherical shells whose centers are the coincided location of the source. The radius to the outer shell is  $r_2$ , which is greater than that of the radius of the inner shell  $r_1$ . Thus, the surface areas of the outer and inner shells are  $4\pi r_2^2$  and  $4\pi r_1^2$ , respectively. By energy conservation, the total energy flowing through the outer shell and the inner shell at a given time should be keep the same so that we have

$$\begin{aligned} E_1 &= E_2 \\ 4\pi r_1^2 I_1 &= 4\pi r_2^2 I_2, \end{aligned} \tag{11.1}$$

where  $I$  denotes the total energy flow through a unit surface at a unit time. The total energy flow can be related to the displacement amplitude in  $r$ -direction  $u_r$  via

$$I \approx u_r^2. \tag{11.2}$$

Inserting Eq. (11.2) in Eq. (11.1) leads to

$$\begin{aligned} 4\pi r_1^2 u_{r1}^2 &= 4\pi r_2^2 u_{r2}^2, \\ r_1 u_{r1} &= r_2 u_{r2}, \\ u_{r1} &= \frac{r_2}{r_1} u_{r2}. \end{aligned} \tag{11.3}$$

Therefore the geometrical spreading for a point source in the far-field in a homogeneous 3D fullspace can be approximately described by

$$u_r(\mathbf{r})_{3D} \approx \frac{1}{r} \tag{11.4}$$

## CHAPTER 11. INVERSION OF 3D DATA WITH A 2D FWT CODE

---

where  $r$  is the radius of the wave front. In the 2D FWT code (Fig. 11.1, top) the energy of an infinitely extending line source is distributed along the lateral surface of an infinite extending cylinder in  $z$ -direction  $\lim_{L \rightarrow \infty} 4\pi rL$ . Thus, the conservation of energy can be written as

$$\begin{aligned}
 E_1 &= E_2 \\
 \lim_{L \rightarrow \infty} 4\pi r_1 L I_1 &= \lim_{L \rightarrow \infty} 4\pi r_2 L I_2, \\
 \lim_{L \rightarrow \infty} 4\pi r_1 L u_{r1}^2 &= \lim_{L \rightarrow \infty} 4\pi r_2 L u_{r2}^2, \\
 r_1 u_{r1}^2 &= r_2 u_{r2}^2, \\
 u_{r1} &= \sqrt{\frac{r_1}{r_2}} u_{r2}.
 \end{aligned} \tag{11.5}$$

Therefore the geometrical spreading for a line source in the far-field in a 2D homogeneous fullspace can be approximately described by

$$u_r(\mathbf{r})_{2D} \approx \frac{1}{\sqrt{r}}. \tag{11.6}$$

In conclusion the amplitude decay with offset is different in a 3D and 2D medium, respectively. Additionally the 2D code is not capable to model 3D point sources, only infinite line sources extending in  $z$ -direction. Therefore not only the geometrical spreading is different, but also the radiation characteristic. The influence of the spreading correction on the inversion results for the simple 1D and the complex 2D Marmousi2 model are discussed in the following two sections.

### 11.1 3D to 2D data transformation

---

To generate a perfect 2D geology, the Marmousi2 model is transformed in a 3D-model by duplicating the 2D model slices in the  $z$ -direction, so only the geometrical spreading and the source characteristics are different (Fig. 11.1, bottom). For this 3D model a data set with the same acquisition geometry as in the 2D case is calculated. The sources and receivers are located along a 2D line, but the sources in the 3D case are point sources.

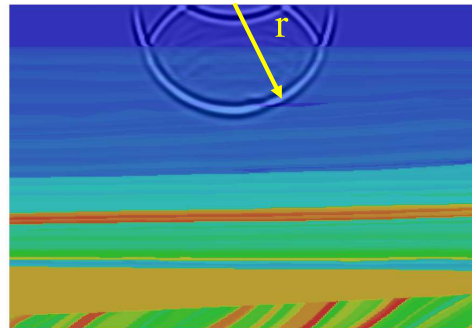
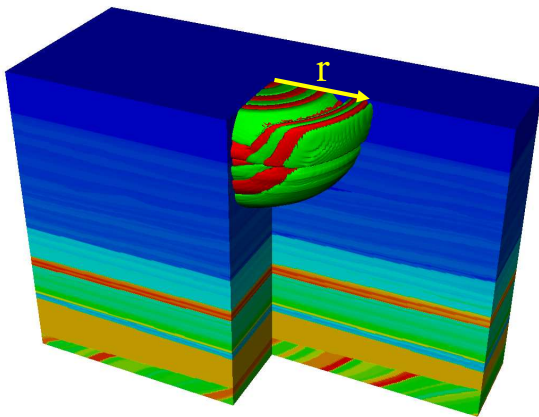


## 11.1. 3D TO 2D DATA TRANSFORMATION

---

Marmousi II Model: 2D vs 3D geometrical spreading

2D: Amplitude  $\sim 1/r$



3D: Amplitude  $\sim 1/r^2$

Figure 11.1: Comparison between 2D (top) and 3D (bottom) geometrical spreading.

## CHAPTER 11. INVERSION OF 3D DATA WITH A 2D FWT CODE

---

As an example the seismic section of shot 25 (y-component) for this 3D model is compared with the corresponding result of the 2D modeling (Fig. 11.2 (top)). The normalization of the 2D and 3D seismic sections are identical. Note the stronger amplitude decay with offset in 3D compared with the 2D case. The amplitudes and phases show a large misfit for near (Fig. 11.2 (center)) and far offset traces (Fig. 11.2 (bottom)). Applying the 2D FWT to the uncorrected 3D data would result in strong velocity artefacts and a divergence of the inversion scheme. In case of a homogeneous full-space the 3D data  $\mathbf{u}_{3D}$  can be easily transformed to 2D data  $\mathbf{u}_{2D}$  (Bleistein (1986), Pica et al. (1990), Shipp and Singh (2002)). To correct the amplitude misfit due to the geometrical spreading the 3D seismic section has to be multiplied by  $\sqrt{t}$ , where  $t$  denotes the time

$$\mathbf{u}_{2D} \approx \sqrt{t} \mathbf{u}_{3D} \quad (11.7)$$

The phase misfit is corrected by convolving the dataset with  $1/\sqrt{t}$ .

$$\mathbf{u}_{2D} \approx \frac{1}{\sqrt{t}} * \mathbf{u}_{3D} \quad (11.8)$$

This transforms the point source approximately into a line source. Afterwards the amplitudes of the individual traces are normalized to the maximum amplitude of the direct wave.

### 11.2 The simple 1D Marmousi model

---

For the approximately 1D geology this simple correction works very well for the near and far offset traces (Fig. 11.3). There are still some small amplitude and phase misfits, especially at larger offsets and for later arrivals, but most of the phase and amplitude information is transformed correctly. A few near offset traces could not be corrected at all. It is unknown if this is an effect of the different near field radiation characteristics or the different FD codes used to model the 3D and 2D data, respectively. During the FWT these traces were simply omitted during the backpropagation of the data residuals. In Fig. 11.4 - Fig. 11.6 the results of a 2D FWT with the corrected 3D dataset (left) is compared with the 2D FWT of the 2D dataset (right).

## 11.2. THE SIMPLE 1D MARMOUSI MODEL

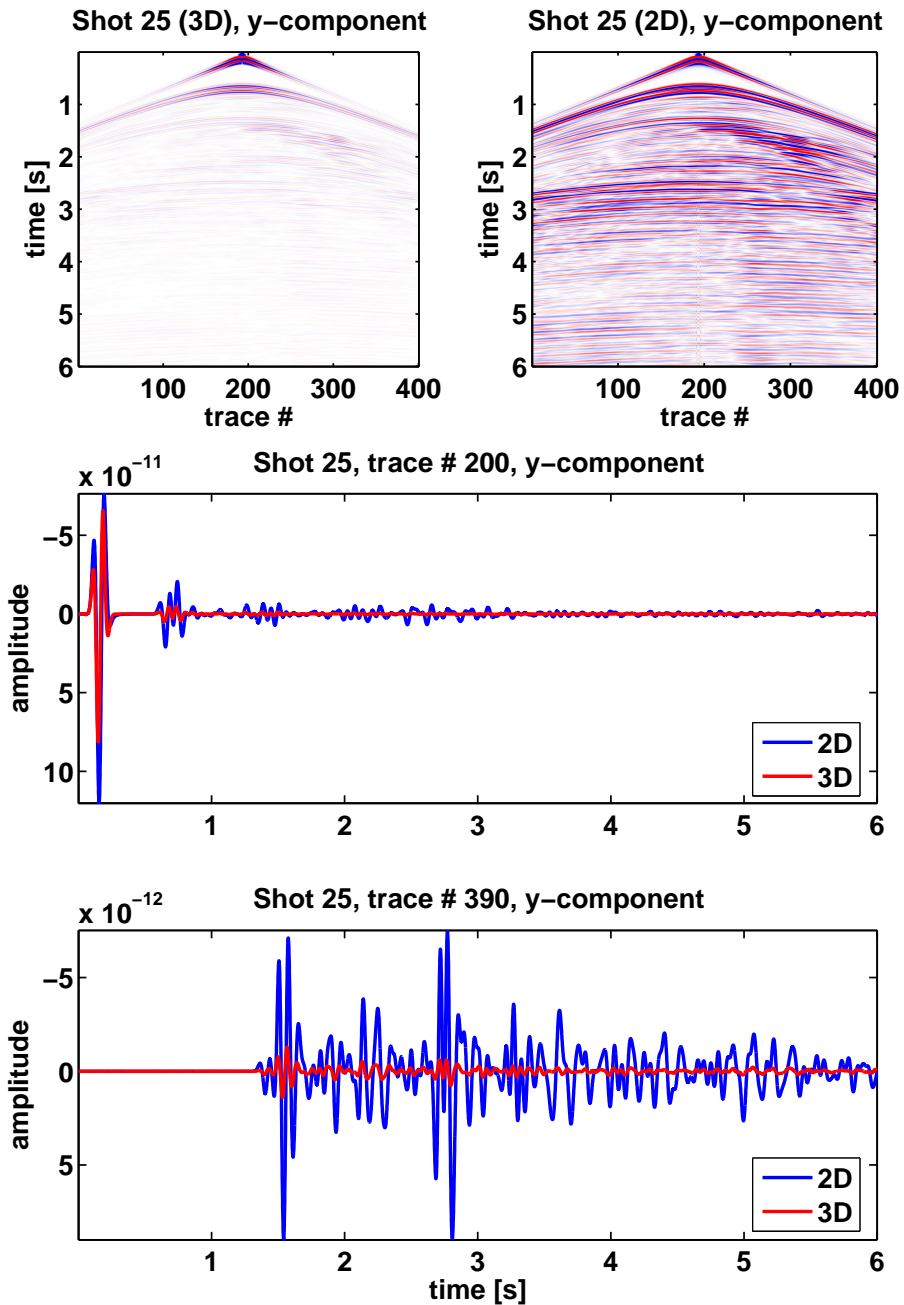


Figure 11.2: Comparison of the seismic section (y-component, shot 50) for the simple Marmousi2 model between the uncorrected 3D data and the 2D data (top). Near offset trace (center) and far offset trace comparison (bottom) for the uncorrected 3D data (red line) and the 2D data (blue line).

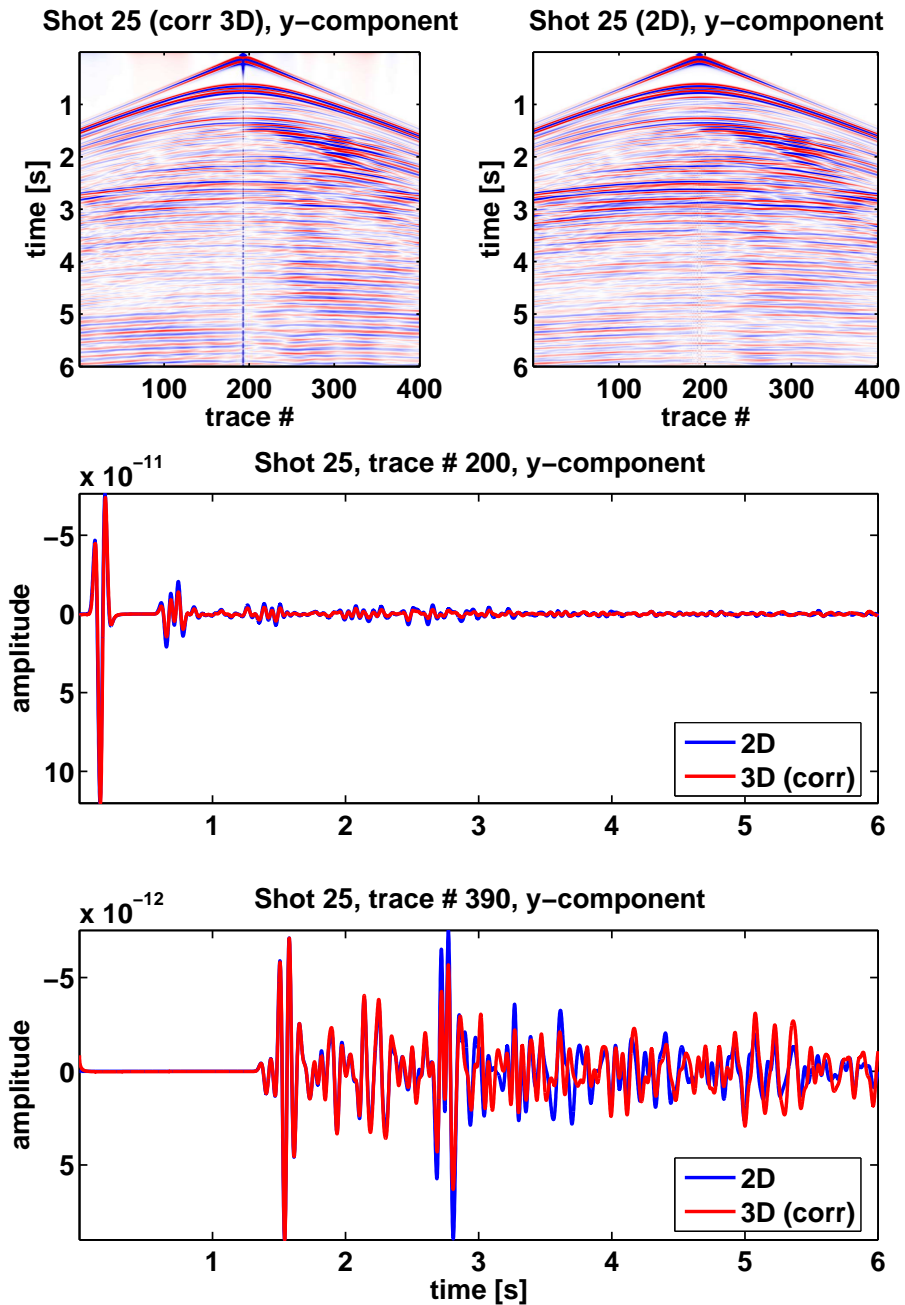


Figure 11.3: Comparison of the seismic section (y-component, shot 50) for the simple Marmousi2 model between the corrected 3D data and the 2D data (top). Near offset trace (center) and far offset trace comparison (bottom) for the corrected 3D data (red line) and the 2D data (blue line).

## 11.2. THE SIMPLE 1D MARMOUSI MODEL

The FWT result for the P-wave velocity model  $V_p$  shows a weak "V" shaped artefact down to a depth of approximately 2.2 km (Fig. 11.4, left). A possible explanation is the different efficiency of the absorbing boundary condition in the 2D and 3D code. Due to the stronger amplitude decay in the 3D code, the amplitudes in the boundary frame are smaller and therefore the damping in the 3D code is more effective compared with the 2D code. Beside these artefacts no major differences are visible. The gas reservoir A, the oil sands B and the structures of the sediment layers are perfectly imaged.

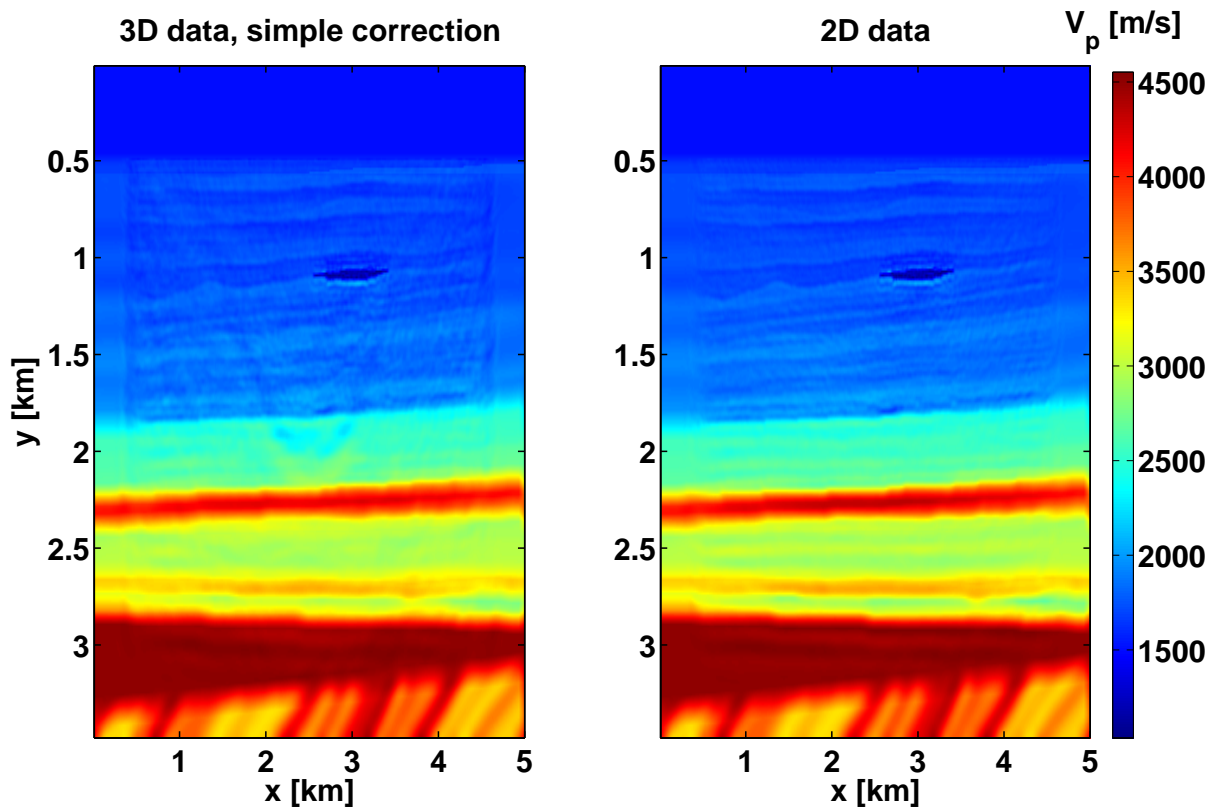


Figure 11.4: FWT results for complex Marmousi2 model. The P-wave velocity model  $V_p$  estimated by the corrected 3D data (left) is compared with the result of the 2D data inversion (right).

## CHAPTER 11. INVERSION OF 3D DATA WITH A 2D FWT CODE

In the FWT result for the S-wave velocity model  $V_s$  (Fig. 11.5, left) the "V" shaped artefact is missing, because of the shorter wavelength of the S-waves and therefore a better performance of the absorbing boundary frame. Down to a depth of 1.2 km the resolution of the inversion results for the 2D and corrected 3D data is comparable. At larger depths the resolution decreases and some imaging errors occur, especially below the gas reservoir A. This can be explained by the larger amplitude and phase misfits of the later S-waves and P-S converted modes.

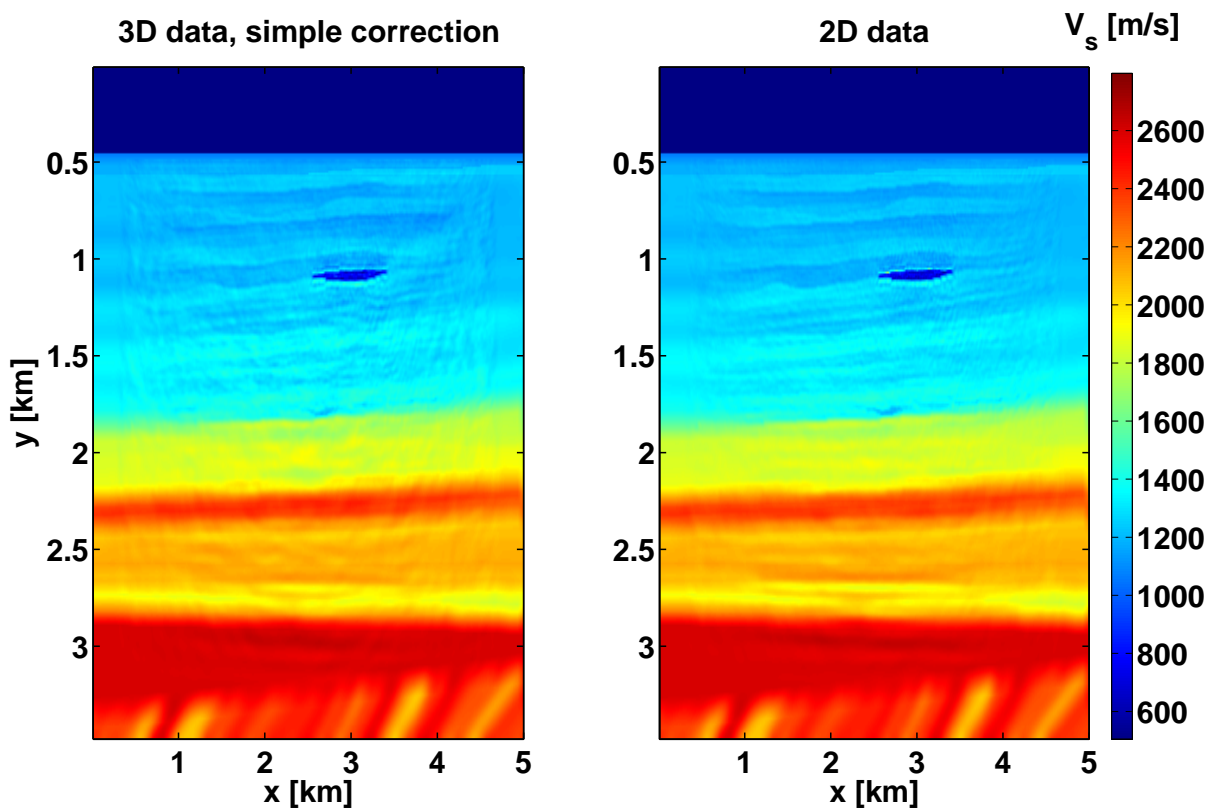


Figure 11.5: FWT results for complex Marmousi2 model. The S-wave velocity model  $V_s$  estimated by the corrected 3D data (left) is compared with the result of the 2D data inversion (right).

## 11.2. THE SIMPLE 1D MARMOUSI MODEL

The remaining amplitude and phase differences in the corrected dataset have a huge impact on the quality of the density inversion. The large scale structures like the hydrocarbon reservoirs A and B or the water wet sands are still imaged correctly, but the overall image is rather "noisy" compared with the FWT result of the 2D data (Fig. 11.6). The clear images for the P- and S-wave velocity lead to the conclusion, that these images are mainly based on the phase information, which was fitted perfectly by the simple spreading correction. The density image however seems to be calculated from the amplitude information and the strong "noise" in the image could be explained by the strong amplitude misfits.

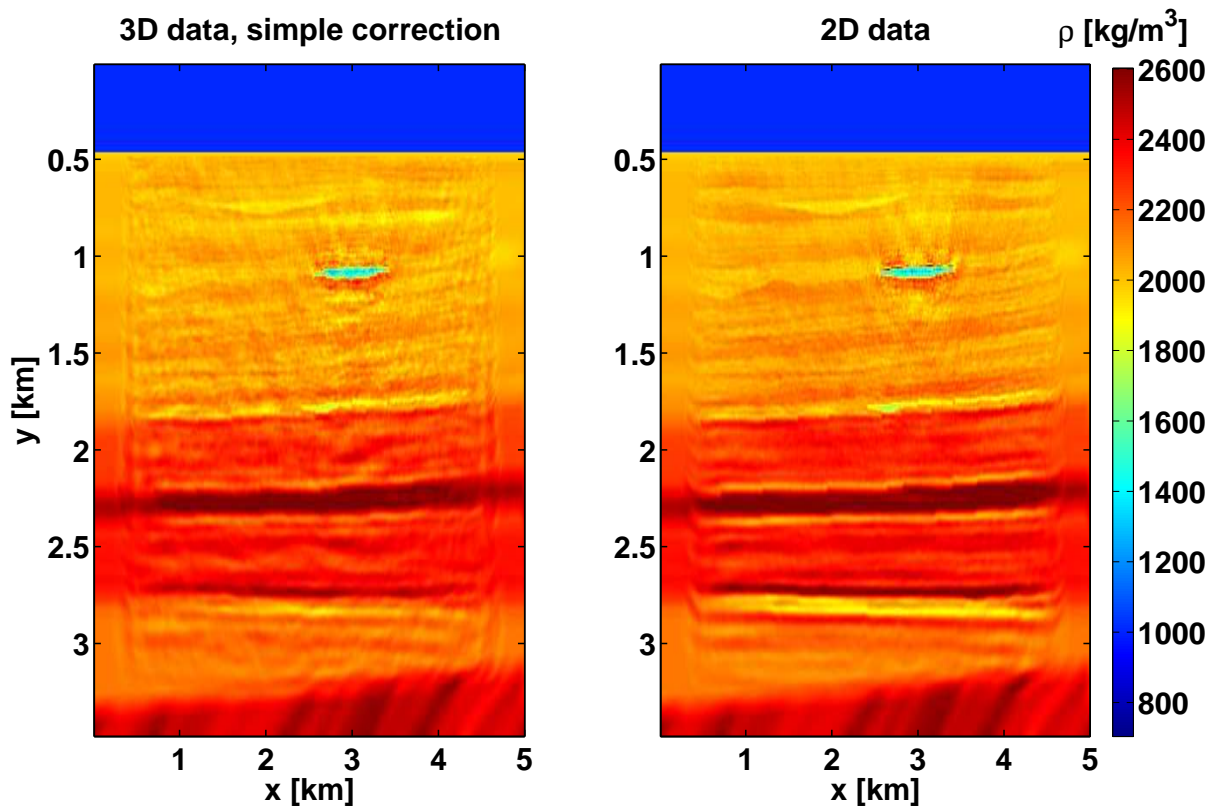


Figure 11.6: FWT results for complex Marmousi2 model. The density model  $\rho$  estimated by the corrected 3D data (left) is compared with the result of the 2D data inversion (right).

### **11.3 The complex 2D Marmousi model**

---

The problem in the last section consisted of a rather simple geology. How does the simple spreading correction work for a more complex geological setting ? To answer this question I build a 3D model of the complex 2D Marmousi model by duplicating the 2D slices in z-direction. In Fig. 11.7 the resulting seismic sections and traces of different offsets are compared with the results of the 2D model. The amplitude misfits for the near and far offsets between the 3D data (red line) and 2D data (blue line) are very obvious. After applying the simple spreading correction on the 3D data, the seismic sections of the corrected 3D and 2D data seem to fit very well (Fig. 11.8, top), except for the few near offset traces which could also not be corrected in case of the simple 1D Marmousi model. But a closer look at the near offset traces reveals that the data could not be fitted as well as in the case of the simple Marmousi2 model. The phase and amplitude misfits are much larger (Fig. 11.8, center). For the far offset trace (Fig. 11.8, bottom) the misfits are very evident when compared with the result for the simple Marmousi2 model. In Figs. 11.9-11.11 the FWT for the corrected 3D data set (top) are compared with the results of the 2D data inversion (bottom). In both cases the smoothed true model is used as a starting model for the FWT. The P-wave velocity model Fig. 11.9 estimated from the corrected 3D data (top) still looks good in the areas with 1D geology where the data correction seem to work well, especially between  $x = 0$  and 4.5 km as well as  $x = 7$  and 9 km. However the thrust fault region is dominated by strong artefacts. Some sediment layers and even the trapped gas sands C1, C2 and C3 can be seen, if you know where to look, but when compared with the results of the 2D data the resolution is strongly decreased. The same behaviour can be observed in the S-wave image of the corrected 3D (Fig. 11.10, (top)). The approximately horizontal layers are imaged well, while the structures in the thrust faults can not be distinguished from inversion artefacts. The bad images for the P- and S-wave velocity models in the vicinity of the thrust fault system can be explained by the strong phase misfits which lead to incorrect velocity models. The same problem is visible in the density image of the corrected 3D data (Fig. 11.11, (top)). While the horizontal layers are imaged correctly, the structures in the thrust fault system look even worse than in case of the seismic velocities. A plausible reason are the strong amplitude misfits in the corrected 3D data.



### 11.3. THE COMPLEX 2D MARMOUSI MODEL

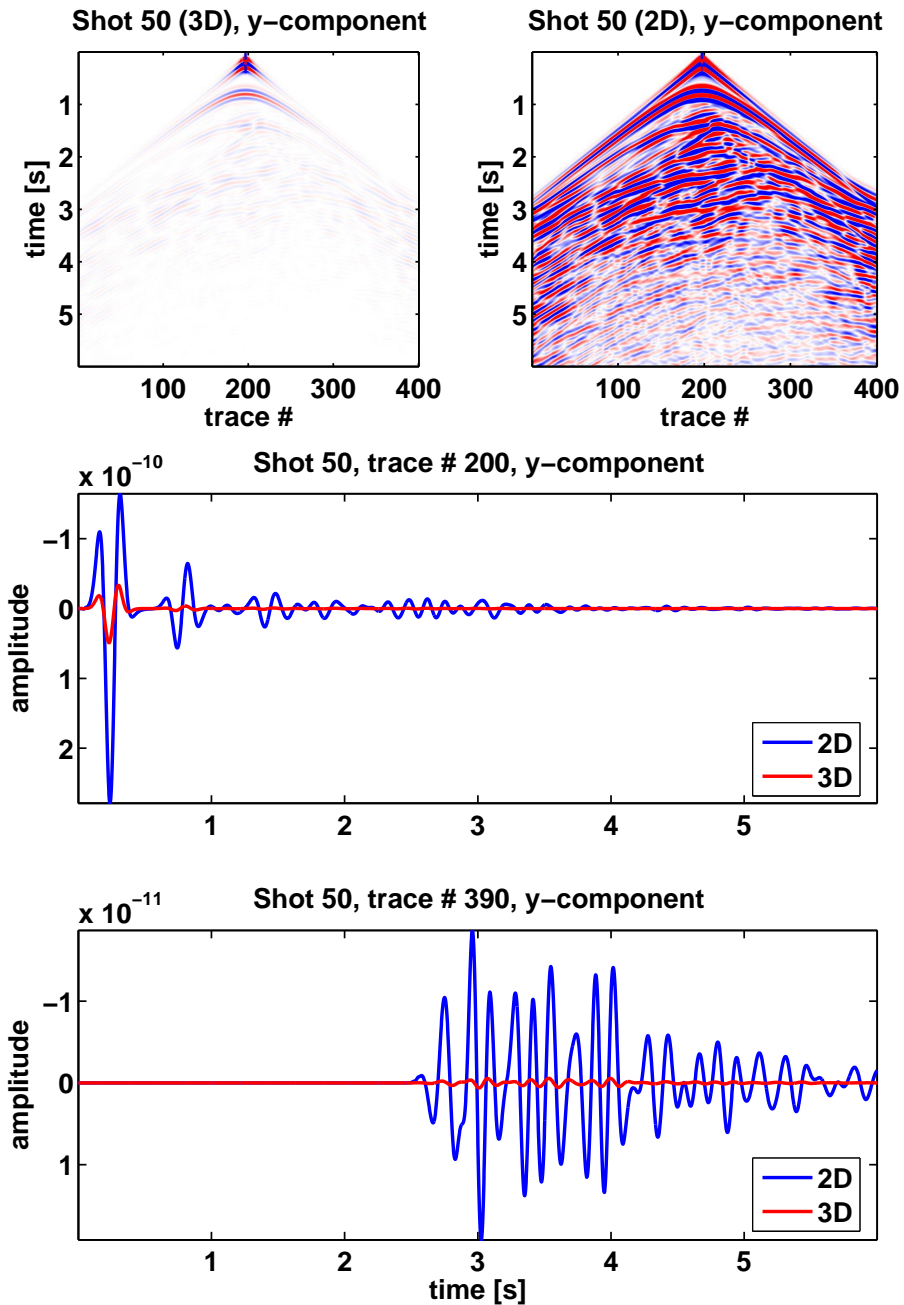


Figure 11.7: Comparison of the seismic section (y-component, shot 50) for the complex Marmousi2 model between the uncorrected 3D data and the 2D data (top). Near offset trace (center) and far offset trace comparison (bottom) for the uncorrected 3D data (red line) and the 2D data (blue line).

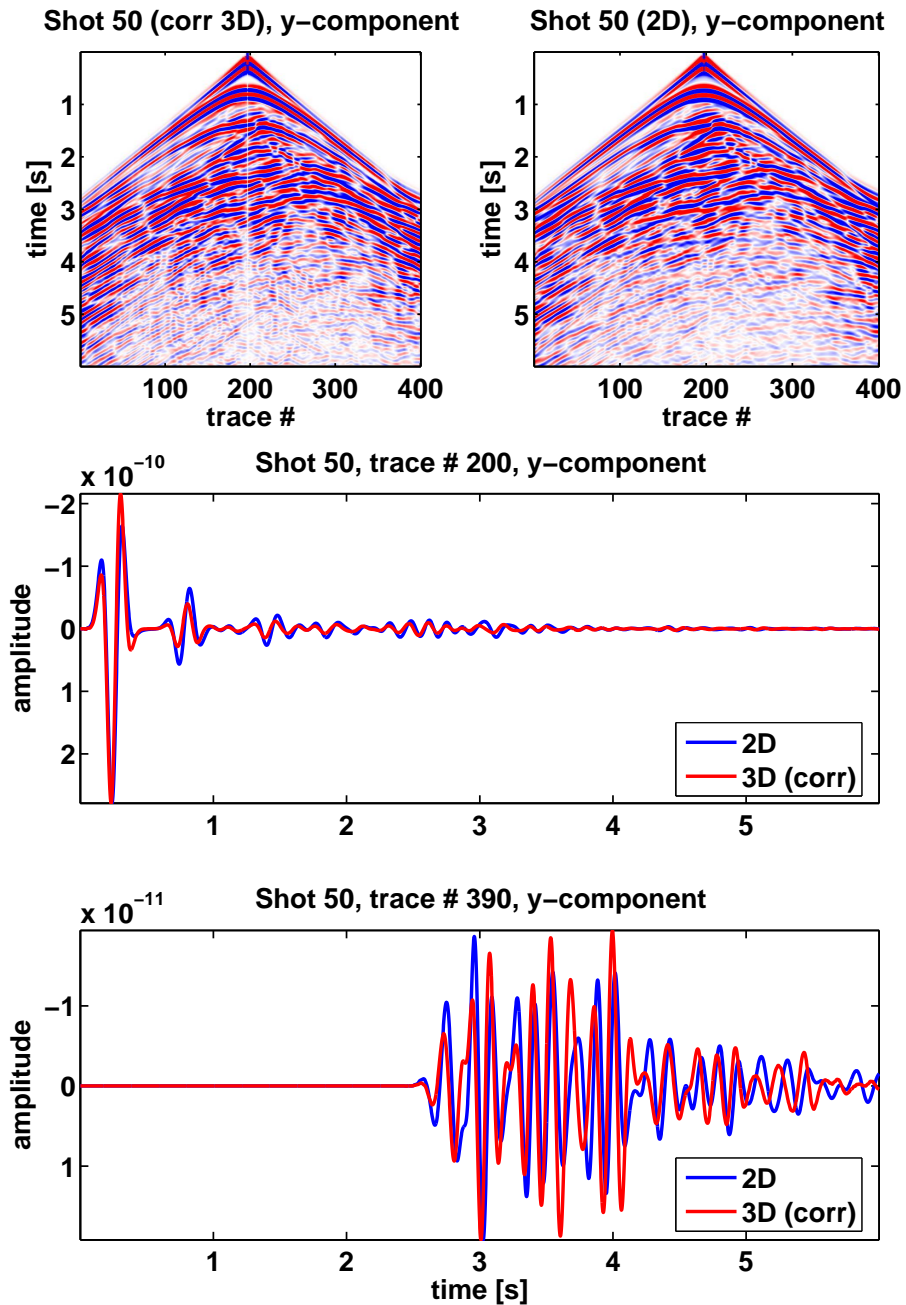


Figure 11.8: Comparison of the seismic section (y-component, shot 50) for the complex Marmousi2 model between the corrected 3D data and the 2D data (top). Near offset trace (center) and far offset trace comparison (bottom) for the corrected 3D data (red line) and the 2D data (blue line).

### 11.3. THE COMPLEX 2D MARMOUSI MODEL

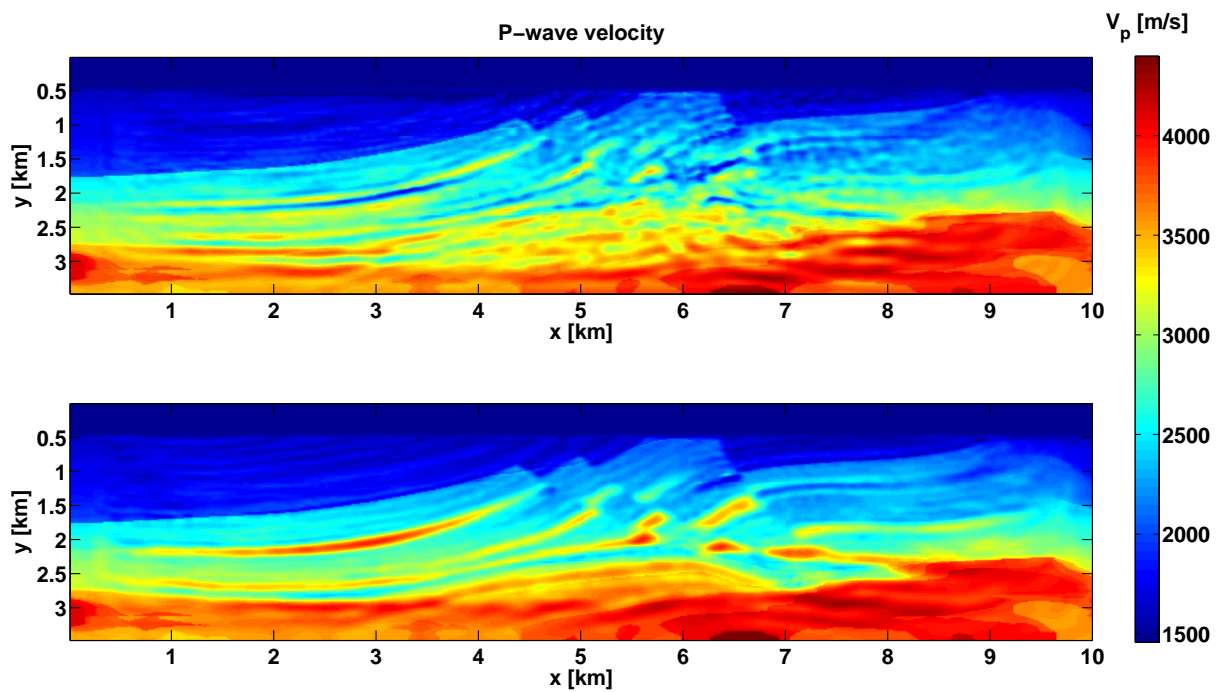


Figure 11.9: FWT results for complex Marmousi2 model. The P-wave velocity model  $V_p$  estimated by the corrected 3D data (top) is compared with the result of the 2D data inversion (bottom).

## CHAPTER 11. INVERSION OF 3D DATA WITH A 2D FWT CODE

---

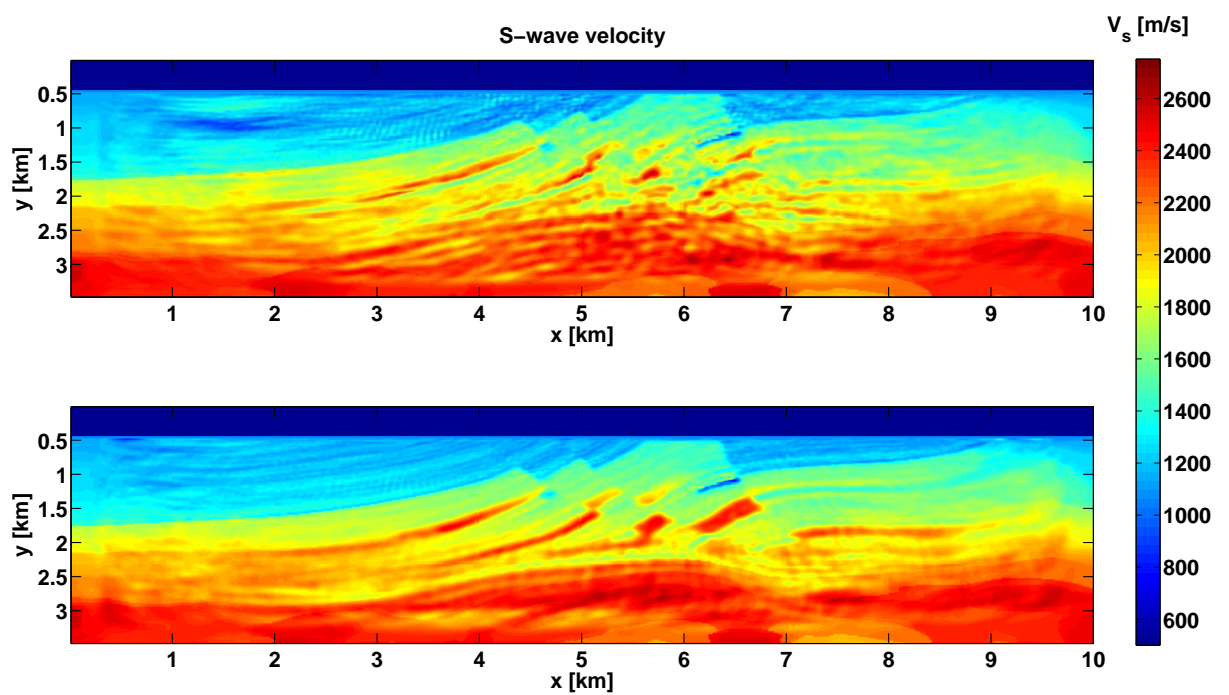


Figure 11.10: FWT results for complex Marmousi2 model. The S-wave velocity model  $V_s$  estimated by the corrected 3D data (top) is compared with the result of the 2D data inversion (bottom).

### 11.3. THE COMPLEX 2D MARMOUSI MODEL

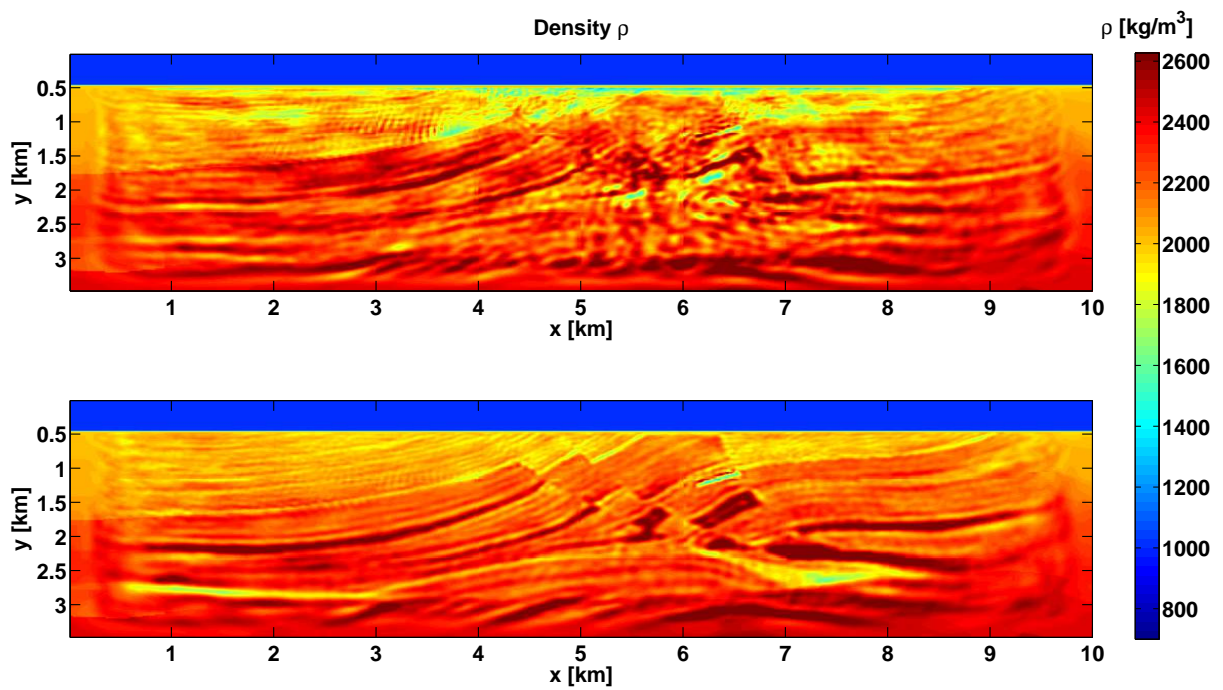


Figure 11.11: FWT results for complex Marmousi2 model. The density model  $\rho$  estimated by the corrected 3D data (top) is compared with the result of the 2D data inversion (bottom).



# 12

## Summary, Conclusions and Outlook

In this work I have presented an implementation of a 2D time domain FWT algorithm and tested the feasibility of a later real data application. Numerical tests range from very simple test problems like a spherical low velocity anomaly to the complex geological elastic Marmousi2 model.

A systematic parameter study with the CTS test problem demonstrates the strong influence of the model parametrization on the resolution and ambiguity of the inversion result, especially the density image. Even though other authors claim that other parametrizations like seismic impedances (Tarantola (1986)) or the Poisson-ratio (Igel et al. (1993)) theoretically lead to better resolutions I think that the seismic velocities are the best choice for the real data application. Imagine you have a very good starting model for the P-wave velocity, estimated by other imaging techniques like travelttime or NIP-wave tomography, a poor estimate of the S-wave velocity model and a very bad estimate for the density model. When converting these models to other parametrizations, e.g. P-wave impedance models  $I_p$  and S-wave impedance models  $I_s$  using the relationship

$$\begin{aligned} I_p &= \rho V_p \\ I_s &= \rho V_s \end{aligned} \tag{12.1}$$

you get a poor starting model for the P-wave impedance and a bad starting model for the S-wave impedance due to the mixing of models with different accuracy. Any other parametrization, which has to be converted from the seismic velocity models leads to a mixing of models and therefore worsen starting models for all three model parameters. This is also crucial when implementing different parametrizations in the inversion code. The parametrization has to be 100% consistent. If seismic velocities are used, seismic velocity models have to be read, used as model parameters in the forward FD code and as inversion parameters in the adjoint problem. Otherwise the inversion results, especially the parameters with high ambiguity like the density,

## CHAPTER 12. SUMMARY, CONCLUSIONS AND OUTLOOK

---

suffer from strong numerical artefacts.

Another problem is the nonlinearity of the elastic adjoint problem. This has three major issues:

1. The nonlinearity of the objective function is frequency dependent. It is smooth at low frequencies, but shows a lot of local minima at higher frequencies.
2. Multiple reflections or strong reflections which are not explained by the starting model increase the ambiguity of the problem.
3. Most of the information about the seismic velocities is contained in the phases, while the amplitudes contain also density information. However the time domain can not separate phase and amplitude information and therefore the nonlinearity of the problem increases.

Problem 1 can be reduced by inverting for low frequencies first and increase the frequency content after a certain amount of iteration steps (see chapter 9). The results for the simple Marmousi2 model in combination with a starting model estimated by the evolution strategy shows only a minor advantage of this approach. Similar results are obtained by Choi et al. (2008a). This could be explained by the low center frequency of the source signal. If the low frequency content (up to 5 Hz) is eliminated by a highpass filter, the quality of the inversion result strongly decreases (see chapter 9). In this time domain implementation the frequency selective inversion is realized by applying a low pass frequency filter with a varying bandwidth on the data residuals before the backpropagation. Two other approaches were also tested. Instead of filtering the data residuals in the frequency domain, a spatial wavelength filter can be applied to the gradients in the time domain. The corresponding inversion results are comparable. Another very smart way to implement a selective frequency inversion is to do the wavefield modeling in the time domain, convert the time domain wavefields "on the fly" during the FD time stepping to the frequency domain and do the inversion in the frequency domain. A nice byproduct of this approach is a strong reduction of the required memory because the forward wavefield needs to be saved only for a few selected frequencies. Sirgue et al. (2008) presented an acoustic 3D implementation of this algorithm. I tested a 2D elastic implementation, however the results are not really satisfying. The resolution is comparable with the time domain results, but only if the discrete frequencies for the inversion are selected correctly. Not to mention that most of the frequency information modelled by the FD code is neglected during the frequency domain inversion.

Problem 2 can be compensated by applying a time window on the data residuals



---

(chapter 8). Thus the upper part of the model is updated first and multiple reflections are neglected before the causative structures have been reconstructed. Even though in most cases the time window reduces the influence of multiples, as described in chapter 8, I experienced for some problems the occurrence of large artefacts and a divergence of the FWT algorithm, e.g. if the time window cuts through strong reflection events and reduces the velocity information contained in the far offset amplitudes and phase moveout of this reflection. Therefore the application of the time window has to be done with caution. Other forms of time windowing might be even better, like parabolic time windows which are consistent with reflection events or time windows based on the first arrival times of the direct or refracted waves.

Problem 3 is debatable. On the one hand, one can claim that the time domain code already separates the phase and amplitude inversion automatically, because the strongest decrease of the objective function is achieved by changing the phase information first and afterwards the amplitudes. On the other hand the misfit function does not explicitly separate phase and amplitude information, it just minimizes the residuals of the displacement amplitude. A true separation of phases and amplitudes can only be achieved in frequency domain. As discussed by Shin and Min (2006) for the acoustic case, the nonlinearity of the problem can be reduced by introducing a new logarithmic misfit function in the frequency domain approach, which separates the phase and amplitude inversion. Their results show that most of P-wave velocity information is contained in the phases. My conclusions of the starting model study (chapter 10) show, that the inversion of the seismic velocity and density model seem to be separated. A bad starting model for the density has no influence on the quality of the velocity models. Therefore if the velocity models are estimated from the phase information and density models from amplitude information, respectively, than this could prove a separate inversion of phase and amplitude information, even though I don't understand why, based on the definition of the objective function. To answer this question additional research has to be done for the elastic case, perhaps with the combined TD/FD approach mentioned above. Alternative approaches to separate phases and amplitudes in the time domain, e.g. by destroying the amplitude information using a wavefield normalization (De Nil (2010), personal communication) are not tested yet. A separate phase inversion would also solve the problem of the 3D to 2D spreading correction for complex 2D media (chapter 11).

Another problem discussed in this thesis is the application of a preconditioning operator. In most older papers (Tarantola (1986), Mora (1987)) it is claimed, that the preconditioning operator is only necessary to increase the convergence speed of the FWT code. However numerical experiments for different acquisition geometries like the cross-well geometry for the spherical inclusion (chapter 4) and the reflection

## CHAPTER 12. SUMMARY, CONCLUSIONS AND OUTLOOK

---

geometry in the case of the Marmousi2 model (chapter 6) demonstrate, that the preconditioning is not only increasing the convergence speed of the code, but is absolutely vital for the convergence at all.

For the reflection acquisition geometry three problems have to be compensated by the preconditioning operator

1. The amplitude decrease in larger depth due to the geometrical spreading and reflections in the upper parts of the model.
2. Artefacts due to large amplitudes at the source and receiver positions.
3. Artefacts near the free surface due to problems with the numerical free surface boundary condition of the backpropagation of the data residuals.

In this FWT implementation problem 1 is compensated by a very crude spreading correction which simply scales the gradient with depth. More sophisticated approaches require the computation of the inverse Hessian matrix  $H_{ij}^{-1}$ . Shin et al. (2001) demonstrated that the inverse Hessian corrects spreading effects very well. Unfortunately the computation of the inverse Hessian in time domain leads to large computational costs (Sheen et al. (2006)). Approximations like the Pseudo-Hessian approach proposed by Choi et al. (2008a) might also work in the time domain.

The artifacts (problem 2 and 3) can be effectively corrected by setting the gradient to zero in the vicinity of these artifacts. However this approach requires a very good knowledge of the model in these regions (see also Ben-Hadj Ali et al. (2008)), because no gradient information means no model update. For a marine environment this approach might work, but it is not very likely that it works also with land data, where heterogeneous weathering layers near the surface are a very common problem. With regard to a desired application to field data two problems were studied in detail.

1. The influence of the starting model.
2. The influence of 3D geometrical spreading.

For an approximately horizontally layered subseafloor model (chapter 6.2) the 1D starting models are estimated using a (10 + 45) evolution strategy. The results are able to describe the long wavelength part of the true model for all three elastic model parameters very well. However the FWT algorithm has problems to fit events with very large amplitudes, e.g. the reflections from the salt layers which are not described by the starting model. This problem leads to a divergence of the FWT algorithm. This problem can be explained in different ways. One explanation is

---

given by the Born approximation, which was used to derive the gradient directions of the individual material parameters. The Born approximation is only valid, if the model perturbations are small. The fitting of the salt layers would require higher order terms or a better starting model. The second explanation is the nonlinearity of the problem. If the starting model is far away from the minimum describing the correct geological subsurface model, it is very likely that the inversion will stuck in the nearest local minimum.

I already pointed out extensively in the conclusions of chapter 10 that the requirements on the starting model are very high to really achieve a high resolution image of the underground, as predicted by the synthetic test problems. In this study a model was picked out of 9800 models because I know the correct average model. In reality it is not easy to estimate what model correctly describes the geology of the underground. The model exactly at the mathematically global minimum of the objective function might be completely nonsense in terms of a correct physical description. Therefore a more extensive global search, hand picking analysis and a lot more FWT calculations are required to seriously invert real field data. The elastic data application in Choi et al. (2008b) does not fulfill this requirement.

The second problem, which occurs in applications to real field data, is the different geometrical spreading in 3D (real world) and 2D (2D FD code). For a 1D medium this effect could be easily compensated by a very simple spreading correction for phases and amplitudes. More sophisticated correction methods lead to similar or even better results (Amundsen and Reitan (1994), Wapenaar et al. (1992), Roberts (2005)), but are also restricted to 1D media. For 2D media, like the complex Marmousi 2 model, the simple spreading correction fails and therefore the inversion results are dominated by artefacts, due to the high data misfits. As mentioned above, the whole problem of amplitude misfits could be avoided by a separate phase inversion, which would at least resolve the seismic velocity models.

Based on the results of this thesis I listed a few recommendations for the geology/acquistion geometry for a first serious data application of the 2D elastic FWT:

1. The geology should be approximately 1D with fine layers including small hydrocarbon reservoirs to demonstrate, if the resolution observed in the synthetic test problems can also be achieved in nature. The 1D geology is absolutly necessary to correct the 3D spreading of the real data and to find plausible starting models for all three model parameters in a finite amount of time. Other a priori information, e.g. estimated from bore hole data, could also be helpful to find plausible starting models.
2. The radiation characteristics of the source should be simple and the source

## CHAPTER 12. SUMMARY, CONCLUSIONS AND OUTLOOK

---

wavelet at least approximately known, e.g. estimated from the direct wave or even better recorded by a deep towed hydrophone.

3. The signal to noise (SN) ratio should be very high. This is necessary because the  $L_2$ -norm used in this FWT implementation is very sensitive to noise. Other norms can be easily realized (Cruse et al. (1990)) with the only constraint that the resulting residual wavefields have to propagate on the time domain FD grid. Norms like the  $L_1$ -norm would result in a rather blocky residual wavefield and therefore exhibit strong grid dispersion.
4. The tests with the marine acquisition geometry for the Marmousi2 model demonstrated surprisingly well resolved results for all elastic material parameters for the chosen frequency band, with a maximum frequency of 20 Hz. Although it appears attractive to complement a relatively high frequency data set acquired on a dense grid by some low frequency data on a less dense grid, e.g. in this case by broadband OBS data, it should be kept in mind that pre-processing of a heterogeneous data set can be very complicated (De Nil, 2009, personal communication). Furthermore, hydrophone and streamer data has the advantage, that imperfect coupling of the receiver to the seafloor does not play a role. Receivers with pronounced directivity might be a problem, too. Frequency transfer functions have to be known and to cover the desired frequency range so that the true wavefield can be recovered.

In conclusion the application of 2D FWT is restricted to approximately 1D media. However the FWT will show it's full potential to image complex geological structures only with a 3D elastic FWT code - SAVA ?

# A

## 1<sup>st</sup> order perturbation of the elastic equations of motion

$$\begin{aligned}\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \sigma_{ij} &= f_i, \\ \sigma_{ij} - c_{ijkl} \epsilon_{kl} &= T_{ij}, \\ &+ \text{boundary conditions},\end{aligned}\tag{A.1}$$

where  $c_{ijkl}$  denote the stiffness tensor,  $\rho$  the density,  $\sigma_{ij}$  the stress tensor,  $\epsilon_{kl}$  the strain tensor,  $f_i$  the volume forces and  $T_{ij}$  the surface forces. In the next step every parameter and variable in the elastic wave equation is perturbed by a first order perturbation

$$\begin{aligned}u_i &\rightarrow u_i + \delta u_i, \\ \rho &\rightarrow \rho + \delta \rho, \\ \sigma_{ij} &\rightarrow \sigma_{ij} + \delta \sigma_{ij}, \\ c_{ijkl} &\rightarrow c_{ijkl} + \delta c_{ijkl}, \\ \epsilon_{ij} &\rightarrow \epsilon_{ij} + \delta \epsilon_{ij}.\end{aligned}\tag{A.2}$$

## APPENDIX A. 1<sup>ST</sup> ORDER PERTURBATION OF THE ELASTIC EQUATIONS OF MOTION

---

These substitutions yield

$$\begin{aligned}
 (\rho + \delta\rho) \frac{\partial^2(u_i + \delta u_i)}{\partial t^2} - \frac{\partial}{\partial x_j}(\sigma_{ij} + \delta\sigma_{ij}) &= f_i \\
 (\sigma_{ij} + \delta\sigma_{ij}) - (c_{ijkl} + \delta c_{ijkl})(\epsilon_{kl} + \delta\epsilon_{kl}) &= T_{ij}, \\
 &+ \text{perturbed boundary conditions,}
 \end{aligned} \tag{A.3}$$

or

$$\begin{aligned}
 \rho \frac{\partial^2 u_i}{\partial t^2} + \rho \frac{\partial^2 \delta u_i}{\partial t^2} + \delta\rho \frac{\partial^2 u_i}{\partial t^2} + \delta\rho \frac{\partial^2 \delta u_i}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} &= f_i \\
 \sigma_{ij} + \delta\sigma_{ij} - c_{ijkl}\epsilon_{kl} - c_{ijkl}\delta\epsilon_{kl} - \delta c_{ijkl}\epsilon_{kl} - \delta c_{ijkl}\delta\epsilon_{kl} &= T_{ij}.
 \end{aligned} \tag{A.4}$$

Because the perturbations are small the quadratic perturbation terms can be neglected

$$\begin{aligned}
 \rho \frac{\partial^2 u_i}{\partial t^2} + \rho \frac{\partial^2 \delta u_i}{\partial t^2} + \delta\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} &= f_i \\
 \sigma_{ij} + \delta\sigma_{ij} - c_{ijkl}\epsilon_{kl} - c_{ijkl}\delta\epsilon_{kl} - \delta c_{ijkl}\epsilon_{kl} &= T_{ij}.
 \end{aligned} \tag{A.5}$$

Collecting the terms of the unperturbed medium on the LHS leads to

$$\begin{aligned}
 \rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} + \rho \frac{\partial^2 \delta u_i}{\partial t^2} + \delta\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} &= f_i \\
 \sigma_{ij} - c_{ijkl}\epsilon_{kl} + \delta\sigma_{ij} - c_{ijkl}\delta\epsilon_{kl} - \delta c_{ijkl}\epsilon_{kl} &= T_{ij}.
 \end{aligned} \tag{A.6}$$

---

Due to the definition of the source terms in Eq. (A.1) the terms of the unperturbed medium are equal to the source terms  $f_i$ ,  $T_{ij}$  and are therefore canceled

$$\begin{aligned} \rho \frac{\partial^2 \delta u_i}{\partial t^2} + \delta \rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} &= 0 \\ \delta \sigma_{ij} - c_{ijkl} \delta \epsilon_{kl} - \delta c_{ijkl} \epsilon_{kl} &= 0. \end{aligned} \tag{A.7}$$

Distributing the terms with perturbations of the material parameters on the RHS and the perturbations of wavefield variables on the LHS yields

$$\begin{aligned} \rho \frac{\partial^2 \delta u_i}{\partial t^2} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} &= \Delta f_i \\ \delta \sigma_{ij} - c_{ijkl} \delta \epsilon_{kl} &= \Delta T_{ij}. \end{aligned} \tag{A.8}$$

where the new source terms

$$\begin{aligned} \Delta f_i &= -\delta \rho \frac{\partial^2 u_i}{\partial t^2}, \\ \Delta T_{ij} &= \delta c_{ijkl} \epsilon_{kl} \end{aligned} \tag{A.9}$$

have been introduced.





# B

## Green's Function Solution

If a unit impulse is applied as a source term at  $\mathbf{x} = \mathbf{x}'$  at time  $t = t'$  in the  $n$ -direction, then we denote the  $i$ th component of the displacement field at any point  $(\mathbf{x}, t)$  as **Green's function**  $G_{in}(\mathbf{x}, t; \mathbf{x}', t')$  (Aki and Richards (1980)). It depends on the receiver and source coordinates and satisfies the equation

$$\rho \frac{\partial^2 G_{in}}{\partial t^2} - \frac{\partial \sigma_{ij}}{\partial x_j} = \delta_{in} \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \quad (\text{B.1})$$
$$\sigma_{ij} = c_{ijkl} \epsilon_{kl}.$$

at all points in the volume  $V$ .  $\delta_{in}$  denotes the Kronecker Delta, while  $\delta(y - y')$  is the Dirac function:

$$\delta(y - y') \begin{cases} = 1 & \text{if } y = y' \\ = 0 & \text{if } y \neq y'. \end{cases}$$

for any variables  $y$  and  $y'$ . The initial conditions are

$$\left\{ \begin{array}{l} \mathbf{G}(\mathbf{x}, t; \mathbf{x}', t') = 0 \\ \frac{\partial \mathbf{G}(\mathbf{x}, t; \mathbf{x}', t')}{\partial t} = 0 \end{array} \right\}: \text{ if } t < t' \text{ and } \mathbf{x} \neq \mathbf{x}'.$$

Two important properties of the Green's function related to the boundary condition are important to notice here. If the boundary conditions are independent of time then the time origin can be shifted:

$$\mathbf{G}(\mathbf{x}, t; \mathbf{x}', t') = \mathbf{G}(\mathbf{x}, t - t'; \mathbf{x}', 0) = \mathbf{G}(\mathbf{x}, -t'; \mathbf{x}', -t). \quad (\text{B.2})$$

This is a **reciprocal relation for source and receiver time**.

## APPENDIX B. GREEN'S FUNCTION SOLUTION

---

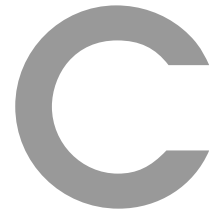
Another important property of the Green's function (see Aki and Richards (1980)) is the **space – timereciprocity**:

$$G_{ij}(\mathbf{x}, t; \mathbf{x}', t') = G_{ji}(\mathbf{x}', t; \mathbf{x}, t'). \quad (\text{B.3})$$

This means that the response at a point  $\mathbf{x}$  in the  $i$ th direction due to a source at point  $\mathbf{x}'$  in the  $j$ th direction is identical to the response at the point  $\mathbf{x}'$  in the  $j$ th direction due to a source at point  $\mathbf{x}$ .

For a given Green's function  $\mathbf{G}$ , the displacement field  $u_i$  due to arbitrary sources  $f_j$  and  $T_{jk}$  can be represented by (see Aki and Richards (1980))

$$\begin{aligned} \delta u_i(\mathbf{x}, t) = & \int_V dV \int_0^T dt' G_{ij}(\mathbf{x}, t; \mathbf{x}', t') f_j(\mathbf{x}', t') \\ & - \int_V dV \int_0^T dt' \frac{\partial G_{ij}}{\partial x'_k}(\mathbf{x}, t; \mathbf{x}', t') T_{jk}(\mathbf{x}', t'). \end{aligned} \quad (\text{B.4})$$



# Bibliography

## Bibliography

---

- Aki, K. and Richards, P. (1980). *Quantitative seismology*. W.H. Freeman and Company.
- Amundsen, L. and Reitan, A. (1994). Transformation from 2-d to 3-d wave propagation for horizontally layered media. *Geophysics*, 59(12):1920–1926.
- Ben-Hadj Ali, H., Operto, S., and Virieux, J. (2008). Velocity model-building by 3D frequency-domain full-waveform inversion of wide-aperture seismic data. 73(5):101–117.
- Bleistein, N. (1986). Two-and-one-half dimensional in-plane wave propagation. *Geophys. Prosp.*, 34:686–703.
- Bohlen, T. (2002). Parallel 3-D viscoelastic finite-difference seismic modelling. *Computers & Geosciences*, 28(8):887–899.
- Bohlen, T. and Saenger, E. (2006). Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves. *Geophysics*, 71(4):T109–T115.
- Brenders, A. and Pratt, R. (2007). Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. *Geophys. J. Int.*, 168:133–151.
- Brossier, R. (2009). *Imagerie sismique à deux dimensions des milieux visco-élastiques par inversion des formes d’ondes : développements méthodologiques et applications*. PhD thesis, Université de Nice - Sophia Antipolis.

## APPENDIX C. BIBLIOGRAPHY

---

- Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 50:705–708.
- Choi, Y., Min, D., and Shin, C. (2008a). Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix: Experience of elastic marmousi-2 synthetic data. *Bull., Seis Soc. Am.*, 98:2402–2415.
- Choi, Y., Min, D., and Shin, C. (2008b). Two-dimensional waveform inversion of multi-component data in acoustic-elastic coupled media. *Geophysical Prospecting*, 56:863–881.
- Courant, R., Friedrichs, K., and Lewy, H. (1928). Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100:32–74.
- Courant, R., Friedrichs, K., and Lewy, H. (March 1967). On the partial difference equations of mathematical physics. *IBM Journal*, pages 215–234.
- Cruse, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A. (1990). Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, 55:527–538.
- Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms - numerical results. *Geophysics*, 51(7):1387–1403.
- Holberg, O. (1987). Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophysical Prospecting*, 35:629–655.
- Igel, H., Debski, W., Djikpéssé, H., and Tarantola, A. (1993). Gradient inversion of marine seismic reflection data: Parametrization and geometrical spreading. In *SEG Expanded Abstracts*, volume 12, pages 657–660.
- Jastram, C. (1992). *Seismische Modellierung mit Finiten Differenzen höherer Ordnung auf einem Gitter mit vertikal variierendem Gitterabstand*. PhD thesis, Universität Hamburg.
- Komatitsch, D. and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. 72(5):155 – 167.
- Komatitsch, D., Michéa, D., and Erlebacher, G. (2009). Porting a high-order finite-element earthquake modeling application to NVIDIA graphics cards using CUDA. *J.Parallel Distrib. Comput.*, 69:451–460.

## BIBLIOGRAPHY

---

- Kurzmann, A., Köhn, D., Przebindowska, A., Ngyuen, N., and Bohlen, T. (2008). Performance of acoustic full waveform tomography for different acquisition geometries. In *12th annual report of the Wave Inversion Technology consortium*, pages 117–125.
- Levander, A. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- Martin, G., Wiley, R., and Marfurt, K. (2006). Marmousi2 - An elastic upgrade for Marmousi. *The Leading Edge*, 25:156–166.
- Moczo, P., Kristek, J., and Halada, L. (2004). *The Finite-Difference Method for Seismologists. An Introduction*. Comenius University, Bratislava.
- Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52:1211 – 1228.
- Müller, G. (2007). *Theory of Elastic Waves*. Samizdat Press.
- Pica, A., Diet, J., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. 55(3):284–282.
- Pratt, R. (1990). Inverse theory applied to multi-source cross-hole tomography. Part II: Elastic wave-equation method. *Geophysical Prospecting*, 38:311–329.
- Pratt, R. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, 64:888–901.
- Pratt, R. (2004). Velocity models from frequency-domain waveform tomography: Past, present and future. In *66th EAGE conference and exhibition, Expanded Abstracts*, pages 181–182, Paris, France.
- Pratt, R., Gao, F., Zelt, C., and Levander, A. (2002). The limits and complementary nature of travelt ime and waveform tomography. In *International Conference of Sub-basalt imaging, Expanded Abstracts*, pages 181–182, Cambridge, England.
- Pratt, R., Huang, L., Duric, N., and Littrup, P. (2007). Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data.
- Pratt, R. and Worthington, M. (1990). Inverse theory applied to multi-source cross-hole tomography. Part I: Acoustic wave equation method. *Geophysical Prospecting*, 38:287–310.

## APPENDIX C. BIBLIOGRAPHY

---

- Rechenberg, I. (1971). *Evolutionsstrategie - Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. PhD thesis, TU Berlin.
- Roberts, M. (2005). A line source to point source transform for seismograms. In *Lithos Science Report*, volume 7, pages 49–52.
- Robertsson, J., Levander, A., Symes, W., and Holliger, K. (1995). A comparative study of free-surface boundary conditions for finite-difference simulation of elastic/viscoelastic wave propagation. pages 1277–1280, Houston, Texas.
- Rosenbrock, H. (1960). An automatic method for finding the greatest or least value of a function. *The Computer Journal*, 3:175–184.
- Schwefel, H. (1974). *Numerische Optimierung von Computermodellen*. PhD thesis, TU Berlin.
- Sheen, D., Tuncay, K., Baag, C., and Ortoleva, P. (2006). Time domain Gauss-Newton seismic waveform inversion in elastic media. *Geophys. J. Int.*, 167:1373–1384.
- Shin, C. and Min, D. (2006). Waveform inversion using a logarithmic wavefield. *Geophysics*, 3:R31–R42.
- Shin, C., Yoon, K., Marfurt, K., Park, K., Yang, D., Lim, H., Chung, S., and Shin, S. (2001). Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion. *Geophysics*, 6:1856–1863.
- Shipp, R. and Singh, S. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophys. J. Int.*, 151:325–344.
- Sirgue, L., Etgen, J., and Albertin, U. (2008). 3d frequency domain waveform inversion using time domain finite difference methods. In *Eur. Assn. Explor. Geosci. Eng., Expanded Abstracts*, page F022, ROME.
- Soubrier, F., Operto, S., Virieux, J., Amestoy, P., and L'Excellent, J. (2009a). FWT2D: a massively parallel program for frequency domain full-waveform tomography of wide-aperture seismic data - part 1: algorithm. *Computer & Geosciences*, 35:487–496.
- Soubrier, F., Operto, S., Virieux, J., Amestoy, P., and L'Excellent, J. (2009b). FWT2D: a massively parallel program for frequency domain full-waveform tomography of wide-aperture seismic data - part 2: numerical examples and scalability analysis. *Computer & Geosciences*, 35:496–514.

## BIBLIOGRAPHY

---

- Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49:1259–1266.
- Tarantola, A. (1984b). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, 32:998–1015.
- Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, 51:1893–1903.
- Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation. *PAGEOPH*, 128:365–399.
- Tarantola, A. (2005). *Inverse Problem Theory*. SIAM.
- Versteeg, R. (1994). The marmousi experience: Velocity model determination on a complex data set. *The Leading Edge*, 13:927–936.
- Virieux, J. (1986). P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.
- Wapenaar, C. P. A., Verschuur, D. J., and Herrmann, P. (1992). Amplitude preprocessing of single- and multicomponent seismic data. *Geophysics*, 57(9):1178–1188.





# D

## Nomenclature



# List of symbols

$\langle \mu \rangle$ .....	Harmonically averaged Lamé parameter $\mu$
$\alpha$ .....	Parameter for the PML boundaries
$\beta_n^{\text{HS}}, \beta_n^{\text{PR}}, \beta_n^{\text{FR}}$ ...	Hestenes-Stiefel, Polak-Ribière and Fletcher-Reeves weighting factors $\beta$ in the conjugate gradient direction at iteration $n$
$\beta_k$ .....	FD weighting coefficient
$\delta \tilde{\mathbf{u}}, \delta \tilde{\mathbf{u}}'$ .....	First order perturbation of the data space
$\delta \mathbf{u}$ .....	data residuals
$\Delta \mathbf{f}_i$ .....	Born source for the body forces.
$\Delta \mathbf{T}_{ij}$ .....	Born source for the surface forces.
$\delta_{ij}$ .....	Kronecker's Delta
$\epsilon_{ij}$ .....	strain tensor
$\frac{\partial \mathbf{u}}{\partial \mathbf{m}}$ .....	Frechet kernel for material parameter $\mathbf{m}$ .
$\frac{\partial E}{\partial \mathbf{m}} = \delta \mathbf{m}_n$ .....	Gradient of the objective function with respect to the material parameter $\mathbf{m}$ at iteration step $n$
$\hat{\mathbf{L}}^*$ .....	Adjoint operator.
$\hat{\mathbf{L}}$ .....	Forward operator.
$\lambda_c$ .....	Correlation length
$\lambda_{\min}$ .....	Minimum wavelength
$\lambda$ .....	Lamé parameter
$\lambda$ .....	wavelength
$\delta \lambda, \delta \mu, \delta \rho$ .....	Gradients for the parametrization Lamé parameter $\lambda, \mu$ and density
$\delta \mathbf{c}_n$ .....	Conjugate gradient direction at iteration step $n$
$\delta \mathbf{m}_{\text{evo}}$ .....	Random mutation
$\delta \mathbf{m}_{\text{max}}$ .....	Maximum random mutation
$\delta \mathbf{m}_{\text{new}}$ .....	Gradients in terms of new material parameters $\mathbf{m}_{\text{new}}$
$\delta \mathbf{m}, \delta \mathbf{m}'$ .....	First order perturbation of the material parameter $\mathbf{m}$
$\delta \mathbf{m}_n^{\text{p}}$ .....	Preconditioned gradient direction at iteration step $n$
$\delta \mathbf{V}_{\text{p}}, \delta \mathbf{V}_{\text{s}}, \delta \rho_{\text{vel}}$ .	Gradients for the parametrization P-, S-wave velocity and density
$\mu_n$ .....	Step length at iteration step $n$
$\omega$ .....	Residual wavefield of the acoustic inverse problem
$\Psi$ .....	Residual wavefield.
$\mathbf{d}_w$ .....	Thickness of the water column
$\mathbf{g} = [g_{\text{vp}}, g_{\text{vs}}, g_{\rho}]$ .	Gradients of the elastic material parameters for the 1D models
$\mathbf{m}_0 = [V_{\text{p0}}, V_{\text{s0}}, \rho_0]$	Elastic material parameters at the seafloor for the 1D models

## APPENDIX D. NOMENCLATURE

---

$\mathbf{m}_w = [V_p, \rho]$ .....	Acoustic material parameters in the water column for the 1D models
$\mathbf{m}_n$ .....	Model parameters at iteration step n
$\mathbf{P}$ .....	momentum vector
$\mathbf{u}_{\text{mod}}$ .....	modelled data
$\mathbf{u}_{\text{obs}}$ .....	field data
$\mathbf{U}$ .....	Forward wavefield of the acoustic inverse problem
$\mathbf{x}$ .....	position vector
$\mu_i$ .....	Test step length for the ith test model
$\mu_{\text{ext}}$ .....	Extremum step length
$\mu_{\text{opt}}$ .....	Optimum step length
$\mu$ .....	Lamé parameter
$\nabla$ .....	Nabla Operator
$\rho_i$ .....	Arithmetically averaged density $\rho$ in the ith direction
$\rho$ .....	mass density
$\Sigma_i$ .....	traction vector
$\sigma_{ij}$ .....	stress tensor
$\theta$ .....	cubic dilatation
$\zeta_{ij}$ .....	rotation tensor
amp .....	Parameter for the absorbing boundary frame
A .....	surface
$c_{ijkl}$ .....	stiffness tensor
$D_i^+$ .....	Forward FD operator acting in the ith direction
$D_i^-$ .....	Backward FD operator acting in the ith direction
$d_{\text{rand}}$ .....	Gaussian Random number
damp .....	Parameter for the absorbing boundary frame
dA .....	infinitesimal surface element
dh .....	Spatial grid point distance on the FD grid
dt .....	Temporal grid point distance on the FD grid
dV .....	infinitesimal volume
$E_{\mathbf{r}}$ .....	Elastic energy as a function of offset $\mathbf{r}$
E .....	Residual energy/misfit function
E .....	Young's modulus
$f_0, f_1, f_2$ .....	Corner frequencies for the frequency filter
$f_r(x, y)$ .....	Rosenbrock test function.
$f_i$ .....	body force source term
$f_{\text{max}}$ .....	Maximum frequency
FW .....	Parameter for the absorbing boundary frame

---

$f$ .....	frequency
$G_{ij}(\mathbf{x}, t; \mathbf{x}', t')$ .....	Elastic Green's function.
$i, j, k, l, m$ .....	General indices or spatial and temporal grid points on the FD grid
$I_p, I_s$ .....	Impedance for P- and S-waves
$L2_i$ .....	$L_2$ -norm for the $i$ th test model
$L$ .....	Parameter for the PML boundaries
$m_i$ .....	mass
$n_j$ .....	normal vector in $j$ -direction
$NX, NY, NT$ .....	Total number of grid points in the spatial and temporal dimensions of the FD grid
$P_n$ .....	Preconditioning operator at iteration step $n$ .
$t_0, t_1, t_2$ .....	Corner times for the time window
$t$ .....	time
$u_i$ .....	$i$ th component of the displacement
$utt_i^n$ .....	Second derivative of the displacement $\mathbf{u}$ with respect to time at in $i$ direction at time step $n$
$v_i$ .....	$i$ th component of the particle velocity
$V_p, V_s$ .....	P-wave and S-wave velocity
$V_{\max}$ .....	Maximum velocity
$V_{\min}$ .....	Minimum velocity
$V_{\text{pml}}$ .....	Parameter for the PML boundaries
$V$ .....	volume
$x_0, x_{\text{grad}1}, x_{\text{grad}2}, x_{\text{grad}3}, x_{\text{grad}4}$	Corner coordinates for the Preconditioning operator in a cross-well acquisition geometry
$y_0, y_{\text{grad}1}, y_{\text{grad}2}$ ..	Corner coordinates for the Preconditioning operator in a reflection acquisition geometry





## Danksagung (Acknowledgements)

Diese Arbeit entstand am Institut für Geowissenschaften (Abteilung Geophysik) der Christian-Albrechts-Universität zu Kiel und dem Institut für Geophysik der TU Bergakademie Freiberg. Viele Personen haben zum Gelingen dieser Arbeit beigetragen. Besonders hervorheben möchte ich:

- An allererster Stelle **Denise De Nil**, die mich sowohl wissenschaftlich, wie auch moralisch immer wieder unterstützt hat. Ohne Dich und Deine intuitive Art physikalische Probleme zu lösen wäre diese Arbeit nicht das geworden was sie jetzt ist. Insbesondere danke ich Dir für die intensiven Diskussionen, die sich teilweise bis 3 Uhr morgens hingezogen haben und bei denen ca.  $10^{28}$  konstruktive Ideen entstanden sind. Schade das nicht alle Ideen in dieser Arbeit aufgearbeitet werden konnten. Desweiteren danke ich Dir dafür, daß Du mir mindestens dreimal aus einem ziemlichen Motivationstief herausgeholfen hast. Aus diesen und anderen Gründen trägt der elastische 2D Zeitbereichs FWT Code nun den Namen DENISE, der Pre-Alpha 3D Zeitbereichs FWT Code den Namen SAVA?.
- **André Kurzmann** für die vielen Diskussionen zur Implementierung von neuen Ideen in den elastischen/akustischen FWT Code, dem aufspüren von Bugs und der Optimierung der Pre-Pre-Alpha Version von DENISE, die effektiv aus mehr Bugs als lauffähigem Code bestand.
- **Anna Przebindowska** für die vielen Diskussionen zur Anwendung der FWT auf reale Daten und dem aufstöbern von interessanter Literatur aus dem Internet.
- **Lisa Rehor** für die theoretischen Diskussionen und dem Finden einiger schwerwiegender Bugs in dem Vorwärtscode.

## APPENDIX E. DANKSAGUNG (ACKNOWLEDGEMENTS)

---

- **Olaf Hellwig** danke ich für die vielen interessanten Diskussionen zur 2.5D FD Modellierung in Zylinderkoordinaten und der Implementierung von PMLs.
- **Prof. Dr. Thomas Bohlen** für die Betreuung der Arbeit und die vielen Diskussionen. Auch wenn wir nicht unbedingt immer der gleichen Meinung waren, wie die sinnvolle Vorgehensweise zur Lösung eines Problems aussieht.
- **Prof. Dr. Wolfgang Rabbel** für die Übernahme des Korreferats und die Möglichkeit diese Arbeit in Kiel abzuschließen.
- **Dr. Dennis Wilken** für die gute Atmosphäre in meinem Kieler Büro und die Diskussionen zur Inversionstheorie.
- **Nhi Nyguen** und **Stefan Jetschny** für die gute Atmosphäre in meinem Freiburger Büro.
- Den fleißigen Korrekturlesern **Denise De Nil**, **Lisa Rehor**, **André Kurzmann** und **Thomas Bohlen** für ihre sorgfältige Suche nach Fehlern in der Arbeit und die sehr konstruktiven Korrekturvorschläge.
- Den **Systemadministratoren und ihren Assistenten** der Rechenzentren Kiel, Freiberg, Chemnitz, Dresden und Karlsruhe, sowie dem Hochleistungsrechenzentrum Nord (HLRN) für die Bereitstellung von Rechenzeit zum debuggen und testen des FWT Codes.
- Der **Deutschen Forschungsgemeinschaft (DFG)**, der **TU Bergakademie Freiberg**, **WesternGeco**, dem **Wave Inversion Technology Consortium (WIT)** sowie der **Verbundnetzgas AG (VNG)** für die finanzielle Unterstützung.
- Meinen Eltern **Hermann** und **Sylvia Köhn** für die Unterstützung während der Studienzeit in Kiel. Insbesondere meinem Vater danke ich für die Begeisterung, die er bei mir für die Naturwissenschaften geweckt hat. Leider konnte er die Fertigstellung dieser Arbeit nicht mehr miterleben.
- Schließlich meinem Neffen **David** und unserem Hund **Pinto** dafür auch mal abschalten zu können.



---

# Curriculum Vitae

## Persönliche Daten

Name: Daniel Köhn

Nationalität: Deutsch

Geburtsdatum: 18. Oktober 1977

Geburtsort: Preetz, Deutschland

Familienstand: Ledig

## Schulbildung

06/1984–04/1997 Grundschule (Raisdorf) und Gymnasium (Preetz)

## Hochschulbildung

10/1997–10/2002 Studium der Physik an der Christian-Albrechts-Universität Kiel

10/2002–10/2004 Studium der Physik/Geophysik an der Christian-Albrechts-Universität Kiel.

10/2004–10/2005 Diplomarbeit am Institut für Geowissenschaften (Abteilung Geophysik), Kiel mit dem Titel "Modellierung elastischer Wellen auf einem adaptiven Finite-Differenzen-Gitter".

12/2005–11/2006 Wissenschaftlicher Mitarbeiter am Institut für Geowissenschaften (Abteilung Geophysik), Kiel in dem "Scherseis 3D Projekt".

seit 11/2006 Wissenschaftlicher Mitarbeiter am Institut für Geophysik, TU Bergakademie Freiberg, Freiberg (Sachsen).

seit 04/2010 Wissenschaftlicher Mitarbeiter am Institut für Geowissenschaften (Abteilung Geophysik).

July 12, 2011



## Eidesstattliche Erklärung

---

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig unter Anleitung meiner akademischen Lehrer und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungskommission vorgelegt und auch nicht veröffentlicht. Diese Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden.

.....  
*Kiel, den 15.03.2011*

*Daniel Köhn*