KIT
Karlsruhe Institute of Technology

Geophysical
Institute KIT-Dept. of Physics

Joint acoustic full-waveform and gravity inversion - development and synthetic application to a salt dome

Simultane Inversion akustischer Wellenformen und gravimetrischer Daten - Entwicklung und Anwendung auf einen Salzdom

Master's Thesis of

Martin Pontius

at the Department of Physics
Geophysical Institute (GPI)

Reviewer:	Prof. Dr. Thomas Bohlen
Second reviewer:	PD Dr. Joachim Ritter

22. October 2015 – 21. October 2016

Karlsruhe Institute of Technology
Department of Physics
Post office box 6980
76128 Karlsruhe

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

Karlsruhe, October 21, 2016

. . . .

(Martin Pontius)

Abstract

The seismic full-waveform inversion (FWI) has become increasingly popular in recent years because of its high resolution and its capability to handle highly heterogeneous subsurfaces. Whereas seismic velocity models can be built reliably in most cases the treatment of density is still challenging because of the low sensitivity of seismic waves to density variations and trade-off effects between the involved parameters. Moreover, seismic amplitudes depend not only on the density, but also on effects like attenuation, anisotropy and noise. In contrast, classical gravimetric inversion suffers from latent ambiguities and resolves density distributions only for long wavelengths.

To overcome these inherent limitations we perform a joint acoustic full-waveform and gravity inversion where we do not use empirical relationships or criteria based on structural similarity. For this purpose, a synthetic reconstruction test is applied to a salt dome structure which is embedded in sedimentary layers.

The reconstructed density model of the joint inversion contains short wavelength information about sedimentary layers which are not included in the gravity inversion results. Compared to the pure acoustic FWI the joint inversion suffers less from trade-off effects and the final density model explains the pseudo-observed gravity field almost perfectly. In contrast, the final gravity field of the acoustic FWI shows significant residuals.

In conclusion, the integration of gravity data can in fact reduce trade-off effects between P-wave velocity and density while a higher resolution can be achieved compared to the pure gravity inversion. Our approach allows a more confidential interpretation of density distributions than each individual method.

Zusammenfassung

Die seismische *full-waveform inversion* (FWI) genießt in jüngster Zeit wegen ihres hohen Auflösungsvermögens und der Fähigkeit auch hochgradig heterogenen Untergrund aufzulösen immer größere Beliebtheit. Seismische Geschwindigkeitsmodelle können in der Regel zuverlässig bestimmt werden, wohingegen immer noch große Schwierigkeiten bei der Inversion nach der Dichte bestehen. Dies liegt zum einen an der geringen Sensitivität seismischer Wellen in Bezug auf Dichtevariationen und zum anderen an *trade-off* Effekten zwischen den involvierten Parametern. Außerdem ist die Interpretation der seismischen Amplituden grundsätzlich mehrdeutig, da neben der Dichte auch Effekte wie Dämpfung, Anisotropie und Rauschen eine Rolle spielen. Die klassische Schwereinversion leidet dagegen an inhärenter Mehrdeutigkeit, da die Schwere ein Potentialfeld darstellt, und weist nur eine Auflösbarkeit im langwelligen Bereich auf.

Um diese Einschränkungen zu überwinden führen wir eine gemeinsame Inversion von Druckseismogrammen und Schweresignalen durch. Dabei benutzen wir weder empirische Beziehungen noch Kriterien, die auf strukturellen Ähnlichkeiten basieren. Als Testmodell für einen Rekonstruktionstest benutzen wir einen Salzdom, der in einige Sedimentschichten eingebettet ist.

Das rekonstruierte Dichtemodell der gemeinsamen Inversion enthält kurzwellige Informationen über die Sedimentschichten, welche nicht in den Resultaten der reinen Schwereinversion enthalten sind. Im Vergleich zu der reinen FWI leidet die gemeinsame Inversion weniger an gegenläufigen Abhängigkeiten zwischen den unterschiedlichen Parametern. Außerdem erklärt das finale Dichtemodell die pseudo-observierten Schweresignale fast perfekt. Dies ist bei der reinen FWI nicht der Fall.

Abschließend ist zu sagen, dass die Integration von Schweredaten in der Tat die *trade-off* Effekte zwischen der P-Wellengeschwindigkeit und der Dichte verringern kann und dass außerdem eine höhere Auflösung des Dichtemodelles im Vergleich zu der reinen Schwereinversion erreichbar ist. Unsere Inversionsstrategie erlaubt es, eine vertrauenswürdigere Interpretation von Dichteverteilungen durchzuführen als es die individuellen Methoden gewährleisten.

Contents

Abstract i

Zusammenfassung iii

1. Introduction 1

2. Theoretical background 3
- 2.1. Acoustic waves . . . 3
- 2.2. Gravitational field . . . 4
  - 2.2.1. Newton's law of universal gravitation . . . 4
  - 2.2.2. Gravitational potential and attraction . . . 5
  - 2.2.3. Gravitational field of continuously distributed bodies . . . 6
- 2.3. Inverse problems . . . 7
  - 2.3.1. *Ill-posed* problems . . . 7
  - 2.3.2. Solution of the inverse problems . . . 8
  - 2.3.3. Inversion strategies . . . 8

3. Implementation 11
- 3.1. Acoustic code . . . 11
  - 3.1.1. Forward problem . . . 11
  - 3.1.2. Inverse problem . . . 15
- 3.2. Gravity code . . . 20
  - 3.2.1. Choice of parameterization . . . 20
  - 3.2.2. Forward problem . . . 21
  - 3.2.3. Inverse problem . . . 24
- 3.3. Joint inversion . . . 25
  - 3.3.1. Joint misfit functional . . . 26
  - 3.3.2. Joint gradients . . . 26

4. Model, forward modeling and inversion parameters 29
- 4.1. Salt dome model . . . 29
- 4.2. Pseudo-observed data and forward modeling parameters . . . 31
  - 4.2.1. Acoustics . . . 31
  - 4.2.2. Gravity . . . 33
- 4.3. Inversion parameters . . . 38

5. Inversion results 41
- 5.1. Acoustic full-waveform inversion . . . 41

5.2. Gravity inversion 46
5.3. Joint inversion 48
5.4. Comparison 54

6. Conclusion and outlook **57**
6.1. Conclusion 57
6.2. Outlook 58

Bibliography **59**

A. Appendix **63**
A.1. Computational resources 63
A.2. Exemplary gradients and model update 63
A.3. Acknowledgements/Danksagungen 65

List of Figures

2.1. Gravitational attraction of a point mass and of a continuously distributed body 5

3.1. Staggered finite-difference grid 12
3.2. Parameterization with rectangular prisms which are integrated into the finite-difference grid 23
3.3. Flow chart of one iteration of the inversion algorithm 27
3.4. Weighting between acoustic full-waveform and gravity inversion 28

4.1. True density model of the salt dome structure 30
4.2. True P-wave velocity model of the salt dome structure 31
4.3. Trace-normalized pseudo-observed data for shot 6 33
4.4. Amplitude-frequency spectrum of the pseudo-observed data for shot 6 . 34
4.5. Boundary effects of the gravimetric forward modeling 37

5.1. Results of the acoustic FWI 43
5.2. Evolution of the normalized L_2-misfit during the acoustic FWI 43
5.3. Final seismograms of shot 6 of the acoustic FWI 44
5.4. Final seismograms of shot 19 of the acoustic FWI 44
5.5. Cross-sections of the P-wave velocity and density models of the acoustic FWI 45
5.6. Results of the gravity inversions 47
5.7. Evolution of the normalized L_2-misfit during the gravity inversions . . . 48
5.8. Final P-wave velocity model and seismograms for shot 6 of two joint inversions ($\gamma = 0.1$ and $\gamma = 0.3$) 50
5.9. Final density model and gravity field of two joint inversions ($\gamma = 0.1$ and $\gamma = 0.3$) 51
5.10. Cross-sections of the density models of two joint inversions ($\gamma = 0.1$ and $\gamma = 0.3$) 52
5.11. Evolution of the normalized joint misfit functional of two joint inversions ($\gamma = 0.1$ and $\gamma = 0.3$) and gravity misfit functional after reduction of the weighting parameter λ_1 53
5.12. Comparison of the final inverted density model and gravity field between the acoustic FWI, the gravity inversion and the joint inversion ($\gamma = 0.3$) . 55
5.13. Comparison of the cross-sections of the density model between the acoustic FWI and the joint inversion ($\gamma = 0.3$) 56

A.1. Exemplary gradients and model update of the joint inversion with $\gamma = 0.3$ at iteration 18 64

List of Tables

3.1. Number of grid points per minimum wavelength n and factor h for different lengths of the FD Taylor operator 14

4.1. Explanation of the geological bodies in the true model 30
4.2. Acoustic forward modeling parameters 32
4.3. Gravimetric forward modeling parameters 34
4.4. Inversion parameters 39
4.5. Multi-stage workflow of the acoustic FWI 39

A.1. Computational aspects of the inversions shown in this thesis 63

1. Introduction

A major goal of geophysical applications is to obtain knowledge about the earth's inner structure from global to local scale. Continuous information about large areas is derived by the acquisition and interpretation of different types of geophysical data such as seismic, electromagnetic and gravimetric data. Today, the seismic full-waveform inversion (FWI) is increasingly popular because of its high resolution and its capability to handle highly heterogeneous subsurfaces. When it was introduced into the seismic community by Lailly (1983), Tarantola (1984), and Mora (1987) limited computational resources restricted its application. However, due to fast improvements of the computational power FWI is now feasible even with a large number of model parameters. During the FWI full seismograms are fitted which leads to a resolution of up to half of the propagated wavelength. Generally, the seismic FWI can be used to build velocity models and density models. While seismic velocity models can be built reliably it is still challenging for density distributions in the subsurface (e.g. Przebindowska, 2013). The main reason is the low sensitivity of seismic waves to the density which is mostly restricted to density variations at layer boundaries. Moreover, trade-off effects between seismic velocities and density may also complicate the inversion. The success of the FWI and especially of the density model building highly depends on the initial model, the low frequency and wavenumber content of the seismic wavefield and the acquisition geometry (Virieux and Operto, 2009).

A strategy to mitigate the nonlinearity and trade-off effects is to apply a hierarchical inversion where seismic velocities or Lamé parameters are updated in the first stage and density in the second stage (e.g. Jeong et al., 2012). In this case, the overall velocity trend is already correct before fitting the amplitudes by updating the density model (Przebindowska et al., 2012). In contrast, a simultaneous seismic multi-parameter inversion can also be successful if different step lengths are used for each parameter class (Xu and McMechan, 2014). Instead of inverting for the density directly empirical laws like Gardner's relation are often used to update the density model. However, empirical laws are usually not valid for various rock types and may produce unsatisfactory results. In a recent study, Wehner (2015) showed that it is possible to improve the inversion for the density by using a combined elastic full-waveform and gravity inversion. He applied a gravity inversion in the first stage of the inversion to introduce long wavelength information to the density model. Later, a pure FWI was applied and led to an improved density reconstruction compared to the pure FWI. A disadvantage of this strategy is that the density update by the gravity inversion may be reversed in the subsequent pure FWI and trade-off effects cannot be reduced efficiently.

To reduce trade-off effects and to use the complementary information of different geophysical methods joint inversions are applied. However, many different joint inversion approaches have been studied in the past and there is no obvious strategy which gives satisfying results in all possible cases. Until now, most of the joint inversion strategies focussed on empirical or analytical laws between different physical parameters and on structural similarities. Physical parameters can be linked by using cross-gradient criteria (Gallardo and Meju, 2004) or by connecting the gradients by empirical relationships (Bjoern Heincke et al., 2006). A more general approach is to use empirical or analytical parameter relationships and structural similarities as a constraint during the inversion (Colombo and De Stefano, 2007; Björn Heincke et al., 2010). However, the assumption of structural similarity and the validity of empirical laws is not always justified. Consequently, we do not make use of them.

To overcome the drawbacks of joint inversions based on empirical relations and structural similarity and of the combined inversion where long wavelength information is only used in the first inversion stage we present a joint full-waveform and gravity inversion. Thereby, both methods can be weighted relative to each other to control the influence of each method. We choose acoustic instead of elastic inversion to avoid additional trade-off effects between shear wave velocity and density and because it is computationally more efficient. With this joint approach we would like to derive a high resolution density model with reliable density values. Therefore, on the one hand we would like to overcome the latent ambiguity of gravimetry and its restricted resolution by providing high resolution seismic data. On the other hand we would like to mitigate trade-off effects between P-wave velocity and density and to compensate the limited sensitivity of seismic waves to density variations by simultaneously minimizing the residuals of the gravity field.

Finally, we give a short overview of the structure of this thesis. In the **second chapter** we explain the theoretical background of acoustic waves, the gravitational field and inverse problems. Therefore, we derive the second-order linearized acoustic wave equation and the gravitational field of continuously distributed bodies. Afterwards we discuss the general challenge of inverse problems, how we can define a solution and which inversion strategies exist. In the **third chapter** we describe the implementation. It consists of the acoustic and gravimetric forward and inverse problem and the joint inversion strategy. Subsequently, the salt dome model which is subject to the synthetic reconstruction test is described in the **fourth chapter** as well as the pseudo-observed data. Additionally, the forward modeling and inversion parameters are explained. In the **fifth chapter** we present the results of the reconstruction test which are divided into four sections: acoustic full-waveform inversion results, gravity inversion results of three different inital models, two joint inversion results with different relative weightings between seismics and gravity and finally a comparison of the reconstructed density models. In the last, **sixth chapter** we summarize the results and give an outlook to future studies and possible further improvements.

2. Theoretical background

In this chapter, we describe the basics of acoustic wave propagation, the gravitational field and inverse problems. Therefore, the linearized acoustic wave equation is derived shortly as well as the gravitational field of continuously distributed bodies based on Newton's law of universal gravitation. Subsequently, we introduce the concept of inverse problems and their challenges. Finally, we adress the question how we can define a solution of inverse problems and how to choose a suitable inversion strategy.

2.1. Acoustic waves

Acoustic waves are the simplest type of seismic waves. They occur in a medium with reference density ρ_0 and reference pressure p_0 if the state of equilibrium is disturbed, e.g. by a local change of pressure. As a result the medium will be deformed and alternately compressed and dilatated. Therefore, acoustic waves are also known as compressional or pressure waves. In the following we will shortly derive the linear acoustic wave equation. A more detailed description can be found in Ehrenfried (2004) and Chapman (2004).
In general three different equations are needed, the equation of motion (conservation of momentum), the continuity equation (conservation of mass) and a constitutive equation (in thermodynamics also called equation of state):

1. equation of motion:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \mathrm{grad})\vec{v} \right] = \mathrm{grad}(p), \tag{2.1a}$$

2. continuity equation:

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho \vec{v}) = 0, \tag{2.1b}$$

3. constitutive equation:

$$K = V \left(\frac{\partial p}{\partial V} \right)_S . \tag{2.1c}$$

Here, we neglected volumetric forces and friction and assumed that acoustic waves can be regarded as an adiabatic process which excludes convection. Moreover, it should be mentioned that the pressure tensor σ_p is not defined in the conventional way as the negative of the stress tensor, but as $\sigma_p = p\mathbf{I}$. Hence the bulk modulus K also has to be defined with opposite sign. Nevertheless, the wave equation will be exactly the same.
In the next step, equations 2.1a and 2.1b will be linearized (cf. Ehrenfried, 2004). Therefore, we divide the pressure p, the density ρ and the particle velocity $\vec{v}$ into a constant (denoted

by 0) and perturbated (denoted by a stroke) part. We assume that the medium is in a static condition at the beginning so that $\vec{v}_0 = 0$ is fulfilled:

$$\begin{aligned} p &= p_0 + p', \\ \rho &= \rho_0 + \rho', \\ \vec{v} &= \vec{v}_0 + \vec{v}' = \vec{v}'. \end{aligned} \tag{2.2}$$

As a result we obtain the following linearized equations:

1. equation of motion

$$\frac{\partial \vec{v}'}{\partial t} = \frac{1}{\rho_0} \operatorname{grad}(p'), \tag{2.3a}$$

2. continuity equation

$$\frac{\partial \rho'}{\partial t} + \rho_0 \operatorname{div}(\vec{v}') = 0, \tag{2.3b}$$

3. constitutive equation

$$K = -\rho_0 \left(\frac{\partial p'}{\partial \rho'} \right)_S . \tag{2.3c}$$

Now, equations 2.3b and 2.3c can be combined yielding a new constitutive equation:

$$\frac{\partial p'}{\partial t} = K \ \operatorname{div}(\vec{v}'). \tag{2.4}$$

If we calculate the time derivative of equation 2.4 and substitute equation 2.3a, we obtain the second-order linearized acoustic wave equation for variable density:

$$\frac{1}{K} \frac{\partial^2 p'}{\partial t^2} - \operatorname{div} \left(\frac{1}{\rho_0} \operatorname{grad}(p') \right) = 0. \tag{2.5}$$

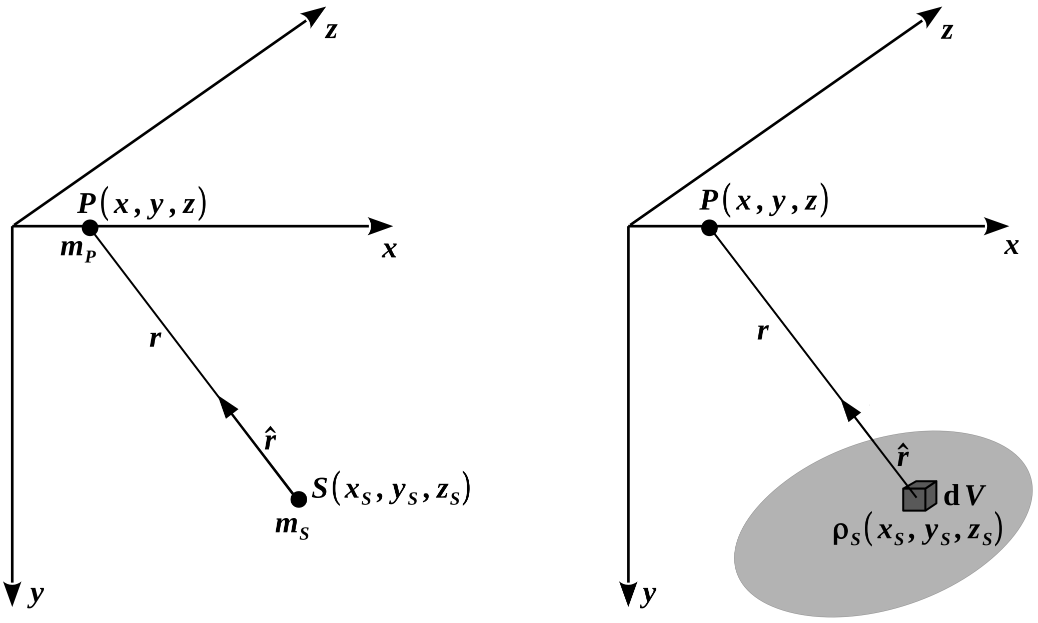
The compressional velocity is given by

$$v_\mathrm{p} = \sqrt{\frac{K}{\rho_0}}. \tag{2.6}$$

2.2. Gravitational field

2.2.1. Newton's law of universal gravitation

The law of universal gravitation was published in 1687 by Isaac Newton in his *Philosophiae Naturalis Principia Mathematica.* It states the existence of a mutual gravitational force $\vec{F}$ between two point masses whose magnitude is proportional to each mass and inversely proportional to the square of their distance r. The force of a point mass m_S centered at source point $S(x_S, y_S, z_S)$ applied on a point mass m_P centered at observation point $P(x, y, z)$ can be formulated in cartesian coordinates as follows (see figure 2.1(a)):



(a)

(b)

Figure 2.1.: Gravitational attraction of a point mass (a) and of a continuously distributed body (b) measured at observation point P

$$\vec{F} = -G\frac{m_S m_P}{r^2}\hat{r} \tag{2.7}$$

with

$$r = \sqrt{(x - x_S)^2 + (y - y_S)^2 + (z - z_S)^2}. \tag{2.8}$$

Here, $G = 6.67384 \frac{\mathrm{m}^3}{\mathrm{kg} \cdot \mathrm{s}^2}$ denotes the gravitational constant. Furthermore, we follow the convention that $\hat{r}$ is directed from the source to the observation point.

2.2.2. Gravitational potential and attraction

If we apply Newton's second law (conservation of momentum), i.e. dividing the force by the mass m_P, we obtain the gravitational attraction (also called gravitational acceleration) $\vec{g}$ at $P(x, y, z)$ caused by mass m_S:

$$\vec{g}(P) = -G\frac{m_S}{r^2}\hat{r}. \tag{2.9}$$

The variable $\vec{g}$ has the SI (*Système Internationale* and International System respectively) unit m/s^2. In the cgs system the unit is cm/s^2 or Gal (named after Galileo Galilei) where 1 Gal equals 1 cm/s^2. In geophysical applications the unit mGal is commonly used with

the conversion 1 mGal = $1 \cdot 10^{-5}$ m/s^2.
In contrast to Newton's original interpretation of gravity as a point attraction, gravity is thought today as a field. In this sense, instead of a direct force between two point masses all point masses together are the cause of the gravitational field. A test particle then moves according to the field strength $\vec{g}$. In general relativity the field is described as the curvature of spacetime. However, in our considerations classical gravity is a sufficient approximation.
Because the gravitational field (equation 2.9) is free of rotation it defines a conservative field and thus can be expressed as the gradient of a scalar potential ϕ:

$$\vec{g}(P) = \nabla\phi(P). \tag{2.10}$$

The gravitational potential ϕ (also called Newtonian potential) represents the work per unit mass which is done by the field if a test particle is moved from its original location to a different one. The total amount of work only depends on the starting point and endpoint. Hence, the gravitational potential can be written as follows:

$$\phi(P) = G\frac{m_S}{r}. \tag{2.11}$$

If the gravitational attraction is defined as $\vec{g} = -\nabla\phi$, the potential ϕ represents the work that is done by the test particle instead of the field.

2.2.3. Gravitational field of continuously distributed bodies

In the field of celestial mechanics it may be feasible to think of a celestial body as a point mass, but in geophysics we are interested to learn about the continuous distribution of matter within the earth. Therefore we replace the point mass or sum of point masses by a continuous density distribution $\rho_S(x_S, y_S, z_S)$ where the infinitesimal mass element $\mathrm{d}m_S$ can be written as $\mathrm{d}m_S = \rho_S(x_S, y_S, z_S)\mathrm{d}V$ (see figure 2.1(b)). Amongst others Kellogg (1929) discussed why the transition from particles to continuously distributed matter is justified. Using the principle of superposition and the integral calculus we now rewrite the Newtonian potential:

$$\phi(P) = G\int_V \frac{\mathrm{d}m_S}{r} = G\int_V \frac{\rho_S(x_S, y_S, z_S)}{r}\mathrm{d}V. \tag{2.12}$$

In fact, in field work not the potential is measured, but the vertical component of the gravitational field g_y. This can be calculated by making use of potential theory (equation 2.10):

$$\begin{aligned} g_y(P) = \frac{\partial\phi(P)}{\partial y} &= \frac{\partial}{\partial y}\left[G\int_V \frac{\rho_S(x_S, y_S, z_S)}{r}\mathrm{d}V\right] \\ &= -G\int_{z_S}\int_{y_S}\int_{x_S} \frac{y - y_S}{r^3}\,\rho_S(x_S, y_S, z_S)\,\mathrm{d}x_S\,\mathrm{d}y_S\,\mathrm{d}z_S. \end{aligned} \tag{2.13}$$

In a similar way gravity gradiometry data can be calculated, i.e. the second derivative of the potential. Today gravity gradiometry data is commonly used in gravity interpretation,

but we will focus only on gravity data in this study.
Moreover, it should be mentioned that gravity is not the only physical theory which incorporates potential theory. It also has high significance in electrostatics or magnetostatics. A comprehensive discussion of potential fields in different physical theories is given by Blakely (1995) and Kellogg (1929).

2.3. Inverse problems

In general, a physical system can be parameterized by a set of model parameters $\vec{m}$ from the model space M and a set of measured data $\vec{d}$ from the data space D. The model parameters and measured data are connected by a physical theory represented by the operator A. Such a theory can be Newton's law of universal gravitation which connects a density model with an observable gravitational field or the acoustic wave equation which connects a density and P-wave velocity model with an observable pressure field.
The process of determining the data by a given set of model parameters and a physical theory is called forward problem. In contrast, the process of determining the model parameters by a given set of data and a physical theory is called inverse problem (cf. Tarantola, 2005). In this respect, nonlinear problems can be denoted as:

$$A(\vec{m}) = \vec{d} \qquad \text{forward problem} \tag{2.14a}$$

$$A^{-1}(\vec{d}) = \vec{m} \qquad \text{inverse problem} \tag{2.14b}$$

Usually it is inevitable to understand the forward problem as it is part of many inversion strategies.

2.3.1. *Ill-posed* problems

When we think about inverse problems the following three questions play an important role:

1. Does a solution exist?
2. Is this solution unique?
3. Is this solution stable?

If the answer to all these questions is positive, the associated problem is called *well-posed* according to Hadamard (1902), if not, it is called *ill-posed*. From a physical point of view it is certain that a solution exists because we study real physical situations and do actual measurements. So the question of existence refers to the mathematical problem formulation. Mathematically it is possible that there is no exact solution because field data always is noisy and the physical theory represented by operator A does not explain the noise. Also, it is possible that we describe the physics incompletely. However, as we investigate

synthetic data without noise a solution should exist.
Especially in gravity inversion the question of uniqueness is of high significance. Because the gravitational field is an integral representation of the density distribution it is intrinsically ambiguous. In other words, an infinite number of density distributions can be added to a density model without changing the externel gravitational field. Such a density distribution can be created by finding two different models with the same external gravity effect and taking their difference (cf. Jacoby and Smilde, 2009, p.274). The seismic inverse problem can be regarded as reasonably unique, at least if the illumination by the waves is sufficient. Although a unique solution should exist, it is not a trivial task to find this solution as it will be explained later in more detail (see section 3.1.2 and 4.3).
Finally, the stability of the solution is also connected to the noise contamination (cf. Zhdanov, 2002). Moreover, instabilities can also occur during the inversion process depending on the chosen inversion strategy.
In this thesis, the goal is to overcome the inherent difficulties of each method by applying a simultaneaous joint inversion of acoustic and gravity data. From the viewpoint of regularization theory it can be said that gravity works as a regularization of acoustic inversion and vice versa.

2.3.2. Solution of the inverse problems

When we think about how we can obtain a solution of the inverse problem, it is crucial to first think about how we define a solution. In the probabilistic approach not one single solution exists, but a set of solutions whose probability is described by a probability function. These solutions and their probabilities incorporate the a priori information, the information associated with the actual data and a physical theory (cf. Tarantola, 2005).
This is a very general approach and has the positive property to take the inherent ambiguity of inverse problems into account. However, as it is computationally expensive we will make use of a deterministic inversion. In this sense, we will look only for one single solution which is called optimum solution. To find the optimum solution we first define a misfit or objective functional E. In our case it is the L_2-norm of the data residuals $\delta\vec{d}$

$$E(\vec{m}) = \frac{1}{2}\delta\vec{d}^T\delta\vec{d} \tag{2.15}$$

with

$$\delta\vec{d} = \vec{d}_{\text{mod}}(\vec{m}) - \vec{d}_{\text{obs}}, \tag{2.16}$$

where $\vec{d}_{\text{mod}}(\vec{m})$ denotes the forward modeled data and $\vec{d}_{\text{obs}}$ the actually observed data. If the misfit functional takes its global minimum (many global minima may exist), the model which is associated with the forward modeled data will be regarded as the optimum solution.

2.3.3. Inversion strategies

The problem of inversion and optimization is a broad field and a large number of inversion strategies exists. The method of choice depends on the concrete problem and also on the computational resources. Extensive discussions are given by Nocedal and Wright (2006), Tarantola (2005), Jacoby and Smilde (2009), and Zhdanov (2002). *Numerical optimization* by Nocedal and Wright (2006) is probably the definitive book if it comes to optimization strategies. Tarantola (2005) focusses on geophysical applications especially on seismics, Jacoby and Smilde (2009) on gravity interpretation and Zhdanov (2002) on regularization in the context of different geophysical applications.

If the relationship between model parameters and data is linear, equations 2.14 can be formulated as matrix vector multiplication and solved by simple matrix inversion. Anyhow, the seismic inverse problem is nonlinear and a different strategy is needed. Generally, nonlinear optimization strategies can be divided into two types: global optimization algorithms and local optimization algorithms. Within global optimization algorithms many forward problems have to be solved, especially if a high number of model parameters is used. Thus, in seismics global optimization algorithms are not applicable with today's computational resources. In contrast, local optimization algorithms can be very efficient but they may get caught in a local minimum. In this regard seismic inversion strategies can be adjusted to reduce the nonlinearity of the inverse problem and avoid the algorithm to get caught in a local minimum as will be explained later. Within local optimization algorithms gradient type algorithms are very prominent. Such a method will be used in this thesis and the basic scheme will be explained in the following.

We already introduced the misfit functional E which depends on the model parameters and on the data. First, we consider the misfit functional of an initial guess $\vec{m}_0$ and regard a small change in the model space Δm. The misfit functional of the new model

$$\vec{m} = \vec{m}_0 + \Delta\vec{m} \tag{2.17}$$

can be written as a second-order Taylor series:

$$\begin{aligned} E(\vec{m}) &= E(\vec{m}_0 + \Delta\vec{m}) \\ &= E(\vec{m}_0) + \Delta\vec{m}\left(\frac{\partial E(\vec{m}_0)}{\partial\vec{m}}\right) + \frac{1}{2}\Delta\vec{m}\left(\frac{\partial^2 E(\vec{m}_0)}{\partial\vec{m}^2}\right)\Delta\vec{m}^T + O(||\Delta\vec{m}||^3). \end{aligned} \tag{2.18}$$

A necessary condition for a minimum is that the first derivative of the misfit functional with respect to $\vec{m}$ vanishes:

$$J_E(\vec{m}) = \frac{\partial E(\vec{m})}{\partial\vec{m}} \approx \frac{\partial E(\vec{m}_0)}{\partial\vec{m}} + \Delta\vec{m}\left(\frac{\partial^2 E(\vec{m}_0)}{\partial\vec{m}^2}\right) \overset{!}{=} 0. \tag{2.19}$$

Rearrangement of equation 2.19 leads to a new expression for the model change:

$$\Delta\vec{m} = -\left(\frac{\partial^2 E(\vec{m}_0)}{\partial\vec{m}^2}\right)^{-1}\left(\frac{\partial E(\vec{m}_0)}{\partial\vec{m}}\right) = -\boldsymbol{H}_E^{-1}(\vec{m}_0)\cdot \boldsymbol{J}_E(\vec{m}_0). \tag{2.20}$$

The variable $\boldsymbol{J}_E$ denotes the Jacobian matrix or gradient of the misfit functional which takes the form of a vector in our case. With N as the number of grid points it can be written as

$$\boldsymbol{J}_E(\vec{m}) = \left(\frac{\partial E}{\partial m_i}(\vec{m})\right)_{i=1,\dots,N} = \left(\tfrac{\partial E}{\partial m_1}(\vec{m}) \quad \tfrac{\partial E}{\partial m_2}(\vec{m}) \quad \cdots \quad \tfrac{\partial E}{\partial m_N}(\vec{m})\right)^T. \tag{2.21}$$

The variable $\boldsymbol{H}_E$ denotes the Hessian matrix which holds information about the curvature of the misfit functional:

$$\boldsymbol{H}_E(\vec{m}) = \left(\frac{\partial^2 E}{\partial m_i \partial m_j}(\vec{m})\right)_{i,j=1,\ldots,N} = \begin{pmatrix} \frac{\partial^2 E}{\partial m_1 \partial m_1}(\vec{m}) & \frac{\partial^2 E}{\partial m_1 \partial m_2}(\vec{m}) & \cdots & \frac{\partial^2 E}{\partial m_1 \partial m_N}(\vec{m}) \\ \frac{\partial^2 E}{\partial m_2 \partial m_1}(\vec{m}) & \frac{\partial^2 E}{\partial m_2 \partial m_2}(\vec{m}) & \cdots & \frac{\partial^2 E}{\partial m_2 \partial m_N}(\vec{m}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial m_N \partial m_1}(\vec{m}) & \frac{\partial^2 E}{\partial m_N \partial m_2}(\vec{m}) & \cdots & \frac{\partial^2 E}{\partial m_N \partial m_N}(\vec{m}) \end{pmatrix}. \tag{2.22}$$

Finally, the basic scheme of a gradient type optimization method can be written as follows:

$$\vec{m}_{n+1} = \vec{m}_n + \Delta\vec{m}_n = \vec{m}_n - \boldsymbol{H}_E^{-1}(\vec{m}_n) \cdot \boldsymbol{J}_E(\vec{m}_n). \tag{2.23}$$

The model at iteration $n + 1$ can be calculated by updating the model at iteration n in the steepest descent direction of the misfit functional $-\boldsymbol{J}_E(\vec{m}_n)$. The Hessian matrix adjusts the steepest descent direction by incorporating information about the curvature of the misfit functional to find a faster way to a local minimum. If the Hessian matrix is neglected the scheme is called gradient method. If it is calculated and used completely the scheme is called Newton method. Lastly, if the Hessian matrix is not calculated explicitly but approximated the scheme is called quasi-Newton method. Such a quasi-Newton method will be used in this study.

3. Implementation

In this chapter, we explain how the forward and inverse problems of seismics and gravity are solved. Moreover, the implementation of the joint inversion is presented. The software that is used in this study is called IFOS2D (*I*nversion of *F*ull *O*bserved *S*eismograms). Its seismic forward code was originally developed by Bohlen (1998). The seismic inversion code is based on the *f*ull-*w*aveform *i*nversion (FWI) code of Köhn (2011). Finally, the gravity forward and inversion code as well as the joint inversion approach was implemented on the basis of a work by Wehner (2015). IFOS2D is maintained and further developed by the Geophysical Institute of the Karlsruhe Institute of Technology (KIT) and available under the GNU General Public License (`http://www.gpi.kit.edu/Software.php`).

3.1. Acoustic code

3.1.1. Forward problem

3.1.1.1. Finite-difference scheme

The acoustic forward problem is solved by a time-domain finite-difference (FD) scheme (Bohlen, 1998). This means that the partial derivatives of the acoustic wave equation are approximated by finite-difference operators and solved numerically.
Therefore, we first discretize the model on an equidistant staggered grid (Virieux, 1986; Levander, 1988) with a grid spacing Δh and calculate the pressure and particle velocity at proceeding time steps Δt:

$$x = i\Delta h, \qquad i = 1, \dots, N_x, \tag{3.1a}$$

$$y = j\Delta h, \qquad j = 1, \dots, N_y, \tag{3.1b}$$

$$t = n\Delta t, \qquad n = 1, \dots, N_t. \tag{3.1c}$$

Here, N_x denotes the number of grid points in x-direction, N_y the number of grid points in y-direction and N_t the number of time steps. The pressure field

$$p := p^{\frac{n}{2}}_{[i,j]} \tag{3.2}$$

is defined on integral grid points in space and half-integral time steps as demonstrated in figure 3.1. The particle velocity field

$$v_x := v^n_{x,[\frac{i}{2},j]} \quad \text{and} \quad v_y := v^n_{y,[i,\frac{j}{2}]} \tag{3.3}$$

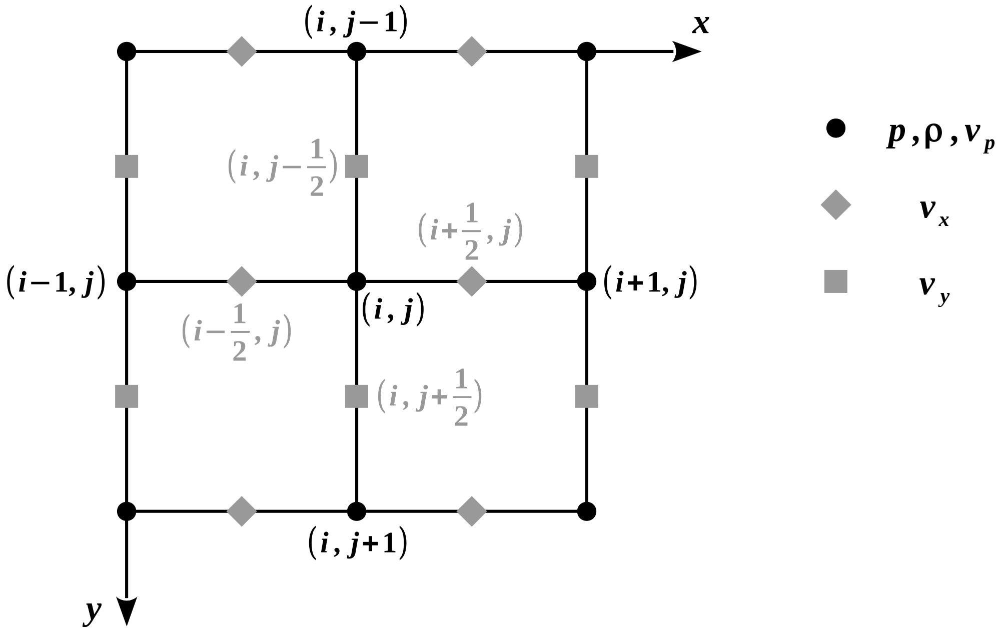


Figure 3.1.: Staggered finite-difference grid. Black dots represent integral grid points, grey squares and diamonds half-integral grid points.

is defined on half-integral gridpoints in space and integral time steps.

To derive the FD operators we consider the second-order Taylor series of a function f with the supporting point x_0:

$$f(x) = \sum_{n=0}^{N} \partial_x^n f(x_0) \frac{1}{n!} (x - x_0)^n = f(x_0) + \partial_x f(x_0)(x - x_0) + O(x^2). \tag{3.4}$$

If we calculate $f(x_0 + \Delta h)$, ignore higher orders and reorganize the equation, we obtain the FD forward operator (equation 3.5a). Evaluating the function at $x_0 - \Delta h$ leads to the FD backward operator (equation 3.5b). The arithmetic mean of the forward and backward operator results in the FD central operator (equation 3.5c). The central operator of the staggered grid is given by equation 3.5d.

$$\frac{\partial f(x_0)}{\partial x} \approx \frac{f(x_0 + \Delta h) - f(x_0)}{\Delta h} = D_+ f(x_0) \tag{3.5a}$$

$$\frac{\partial f(x_0)}{\partial x} \approx \frac{f(x_0) - f(x_0 - \Delta h)}{\Delta h} = D_- f(x_0) \tag{3.5b}$$

$$\frac{\partial f(x_0)}{\partial x} \approx \frac{f(x_0 + \Delta h) - f(x_0 - \Delta h)}{2\Delta h} = D_0 f(x_0) = \frac{1}{2}\left[D_+ f(x_0) + D_- f(x_0)\right] \tag{3.5c}$$

$$\frac{\partial f(x_0)}{\partial x} \approx \frac{f(x_0 + \Delta h/2) - f(x_0 - \Delta h/2)}{\Delta h} = D_{0,\mathrm{sg}} f(x_0) \tag{3.5d}$$

These FD operators are of second order. A better approximation of the partial derivatives can be achieved by expanding the Taylor series, i.e. using higher FD orders. With the central operator of the staggered grid the linearized acoustic wave equation can be discretized. Instead of using the second-order formulation of the acoustic wave equation (equation 2.5), we use the pressure-velocity formulation which is a system of two coupled first-order differential equations (equations 2.3a and 2.4):

$$\frac{v^{[n+1]}_{x,[i+\frac{1}{2},j]} - v^{[n]}_{x,[i+\frac{1}{2},j]}}{\Delta t} = \frac{1}{\bar{\rho}_{[i+\frac{1}{2},j]}} \frac{p^{[n+\frac{1}{2}]}_{[i,j]} - p^{[n+\frac{1}{2}]}_{[i+1,j]}}{\Delta h}, \tag{3.6a}$$

$$\frac{v^{[n+1]}_{y,[i,j+\frac{1}{2}]} - v^{[n]}_{y,[i,j+\frac{1}{2}]}}{\Delta t} = \frac{1}{\bar{\rho}_{[i,j+\frac{1}{2}]}} \frac{p^{[n+\frac{1}{2}]}_{[i,j]} - p^{[n+\frac{1}{2}]}_{[i,j+1]}}{\Delta h}, \tag{3.6b}$$

$$\frac{p^{[n+\frac{1}{2}]}_{[i,j]} - p^{[n-\frac{1}{2}]}_{[i,j]}}{\Delta t} = \rho_{[i,j]} v^2_{p,[i,j]} \frac{v^{[n]}_{x,[i+\frac{1}{2},j]} - v^{[n]}_{x,[i-\frac{1}{2},j]} + v^{[n]}_{y,[i,j+\frac{1}{2}]} - v^{[n]}_{y,[i,j-\frac{1}{2}]}}{\Delta h}. \tag{3.6c}$$

The density on half-integral grid points is averaged as follows:

$$\bar{\rho}_{i+\frac{1}{2},j} = \frac{\rho_{[i,j]} + \rho_{[i+1,j]}}{2} \quad \text{and} \quad \bar{\rho}_{i,j+\frac{1}{2}} = \frac{\rho_{[i,j]} + \rho_{[i,j+1]}}{2}. \tag{3.7}$$

By rearranging equations 3.6 we obtain the final FD scheme:

$$v^{[n+1]}_{x,[i+\frac{1}{2},j]} = v^{[n]}_{x,[i+\frac{1}{2},j]} + \frac{1}{\bar{\rho}_{[i+\frac{1}{2},j]}} \frac{\Delta t}{\Delta h} \left(p^{[n+\frac{1}{2}]}_{[i,j]} - p^{[n+\frac{1}{2}]}_{[i+1,j]}\right), \tag{3.8a}$$

$$v^{[n+1]}_{y,[i,j+\frac{1}{2}]} = v^{[n]}_{y,[i,j+\frac{1}{2}]} + \frac{1}{\bar{\rho}_{[i,j+\frac{1}{2}]}} \frac{\Delta t}{\Delta h} \left(p^{[n+\frac{1}{2}]}_{[i,j]} - p^{[n+\frac{1}{2}]}_{[i,j+1]}\right), \tag{3.8b}$$

$$p^{[n+\frac{1}{2}]}_{[i,j]} = p^{[n-\frac{1}{2}]}_{[i,j]} + \rho_{[i,j]} v^2_{p,[i,j]} \frac{\Delta t}{\Delta h} \left(v^{[n]}_{x,[i+\frac{1}{2},j]} - v^{[n]}_{x,[i-\frac{1}{2},j]} + v^{[n]}_{y,[i,j+\frac{1}{2}]} - v^{[n]}_{y,[i,j-\frac{1}{2}]}\right). \tag{3.8c}$$

Table 3.1.: Number of grid points per minimum wavelength n and factor h for different lengths of the FD Taylor operator.

FD order	n	h
2nd	12	1.0
4th	8	7/6
6th	6	149/120
8th	5	2161/1680

3.1.1.2. Grid dispersion and instability

The spatial and temporal sampling have to fulfill one criterion each to avoid numerical dispersion and instabilities, respectively. The criterion for numerical dispersion is given as

$$\Delta h \leq \frac{\lambda_{\min}}{n} = \frac{v_{\mathrm{p,min}}}{n f_{\max}} \approx \frac{v_{\mathrm{p,min}}}{n 2 f_{\mathrm{c}}}. \tag{3.9}$$

The variable n denotes the number of grid points per minimum wavelength and depends on the FD order (see table 3.1). In the case of second-order FD operators 12 grid points per minimum wavelength should be sufficient. The maximum spatial sampling that avoids numerical disperion can be written in terms of the central frequency f_{c} and the lowest P-wave velocity $v_{\mathrm{p,min}}$ in the model. A sampling according to the Nyquist-Shannon sampling theorem is generally not high enough.

The Courant-Friedrichs-Lewy condition (Courant et al., 1928) states that a minimum temporal sampling Δt has to be chosen to ensure the stability of the modeling. The temporal sampling has to be smaller than the time a wave needs to travel between two adjacent grid points. If this criterion is not fulfilled, the wave amplitude will grow to infinity. In a 2D medium the condition can be written as follows:

$$\Delta t \leq \frac{\Delta h}{h \sqrt{2} v_{\mathrm{p,max}}}. \tag{3.10}$$

The factor h also depends on the FD order (see table 3.1).

3.1.1.3. Initial and boundary conditions

At the beginning, the medium is supposed to be in equilibrium. Therefore the particle velocity and the pressure at every location $\vec{x}$ is given in the following way:

$$\vec{v}(\vec{x}, t=0) = 0, \qquad \dot{\vec{v}}(\vec{x}, t=0) = 0, \qquad p(\vec{x}, t=0) = 0, \qquad \dot{p}(\vec{x}, t=0) = 0. \tag{3.11}$$

At the top of the model a free surface boundary condition is applied using the mirroring technique (Levander, 1988). At the remaining three boundaries a convolutional perfectly matched layer (C-PML) is used (Komatitsch and Martin, 2007). It works as a stretching of the coordinates at the model boundaries by mapping them to complex numbers. As a consequence the waves decay exponentially and artificial reflections from the boundaries can be avoided.

3.1.2. Inverse problem

The inverse problem of acoustic wave propagation is solved by a full-waveform inversion (FWI) algorithm. A comprehensive overview of the FWI was written by Virieux and Operto (2009). In contrast to classical traveltime tomography where only first arrivals are fitted in FWI algorithms whole seismograms are fitted. This provides an increased resolution up to half of the propagated wavelength. The phase information of the wave field is mostly determined by the P-wave velocity whereas the density contributes mainly to the amplitude. As we are especially interested in a high resolution reconstruction of the density the advantage of the FWI should be evident.
Key to the FWI is an efficient and accurate forward modeling engine (see section 3.1.1). As we explained in section 2.3 we will make use of local optimization. Therefore, the convergence to a global minimum cannot be assured. It depends on the accuracy of the initial model, the quality of low frequency information, the noise contamination and the accuracy of the simulation of wave-physics by the forward engine. To mitigate the nonlinearity and *ill-posedness* a hierarchical multi-stage inversion is recommended. This is done by an 8-stage workflow which includes sequential low-pass frequency filtering, offset-dependent trace killing and a single parameter inversion at the beginning (see section 4.3).

3.1.2.1. Misfit functional

The acoustic misfit functional is defined as

$$E_{\mathrm{FWI}}(\vec{m}) = \frac{1}{2}\delta\vec{p}^{T}\delta\vec{p} = \frac{1}{2}\sum_{\text{sources}}\int \mathrm{d}t \sum_{\text{receiver}} ||\delta\vec{p}||^{2} \tag{3.12}$$

with the data residuals

$$\delta\vec{p} = \vec{p}_{\mathrm{mod}}(\vec{m}) - \vec{p}_{\mathrm{obs}}, \tag{3.13}$$

where the modeled pressure $\vec{p}_{\mathrm{mod}} := \vec{p}_{\mathrm{mod}}(\vec{x}_{\mathrm{s}}, \vec{x}_{\mathrm{r}}, t, \vec{m})$ depends explicitly on the source position $\vec{x}_{\mathrm{s}}$, the receiver position $\vec{x}_{\mathrm{r}}$, the time t and the model parameters $\vec{m}$ and the observed pressure $\vec{p}_{\mathrm{obs}} := \vec{p}_{\mathrm{obs}}(\vec{x}_{\mathrm{s}}, \vec{x}_{\mathrm{r}}, t)$ on the the source position $\vec{x}_{\mathrm{s}}$, the receiver position $\vec{x}_{\mathrm{r}}$ and the time t. The explicit dependence on source position, receiver position and time is omitted in all equations for the sake of shortness. Only the dependence on the model parameters of the modeled data and the misfit functional is kept because it is essential for the derivation of the gradient of the misfit functional.

3.1.2.2. Gradient calculation with adjoint state method

For the inversion procedure the gradient of the misfit functional has to be calculated (see section 2.3.3). This is done with the adjoint state method which originated from control theory. It was introduced to geophysical applications by Lailly (1983), Tarantola (1984) and Mora (1987). A more recent overview is given by Plessix (2006). The adjoint state method can be used to calculate the gradient of a misfit functional with respect to the model parameters if the misfit functional does not depend on the model parameters directly, but

through state variables which are solutions of the forward problem. In the acoustic case, the state variables are the modeled pressure data and the model parameters are the P-wave velocities and the densities $\vec{m} = (\vec{\rho}, \vec{v}_{\mathrm{p}})^T$.
The adjoint state method in combination with the minimization of a misfit functional is only one possible inversion strategy. Another widely used strategy is the linearization and iterative solution of the inverse problem. However, within this framework the Fréchet derivative of the data with respect to the model parameters, $\frac{\partial \vec{d}(\vec{m})}{\partial \vec{m}}$, is needed which is represented by a $M \times N$ matrix. Thereby M denotes the number of observations and N the number of model parameters. Thus, the calculation of the Fréchet derivative is computationally expensive and it is not independent of the number of receivers which is the case in the adjoint state method (see equation 3.19). One advantage would be the possibility of a sensitivity analysis where it is necessary to calculate the Fréchet derivative.
In the following we show how the gradient of the misfit functional is derived with the adjoint state method. Generally, the gradient of the misfit functional can be written as

$$\begin{aligned} \frac{\partial E_{\mathrm{FWI}}(\vec{m})}{\partial \vec{m}} &= \frac{1}{2} \sum_{\text{sources}} \int \mathrm{d}t \sum_{\text{receiver}} \frac{\partial ||\vec{p}_{\mathrm{mod}}(\vec{m}) - \vec{p}_{\mathrm{obs}}||^2}{\partial \vec{m}} \\ &= \sum_{\text{sources}} \int \mathrm{d}t \sum_{\text{receiver}} \frac{\partial \vec{p}_{\mathrm{mod}}(\vec{m})}{\partial \vec{m}} \delta \vec{p}. \end{aligned} \tag{3.14}$$

Now, we consider the linear mapping of a small change in the model space $\delta \vec{m}'$ to the data space $\delta \vec{p}'$:

$$\delta \vec{p}' = \int_V \mathrm{d}V \frac{\partial \vec{p}(\vec{m})}{\partial \vec{m}} \delta \vec{m}'. \tag{3.15}$$

For example, this could mean that we observe how the data, i.e. the pressure, at a specific receiver changes if the P-wave velocity is changed at one specific location in the subsurface. The integration over the whole model V sums up the effects of all single changes of the model parameters. Here, $\frac{\partial \vec{p}(\vec{m})}{\partial \vec{m}}$ is the Fréchet derivative. Analogously, we can write the mapping from the data space to the model space:

$$\delta \vec{m}' = \sum_{\text{sources}} \int \mathrm{d}t \sum_{\text{receiver}} \frac{\partial \vec{m}(\vec{p})}{\partial \vec{p}} \delta \vec{p}' = \sum_{\text{sources}} \int \mathrm{d}t \sum_{\text{receiver}} \left[\frac{\partial \vec{p}(\vec{m})}{\partial \vec{m}} \right]^* \delta \vec{p}'. \tag{3.16}$$

The asterisk denotes the adjoint operator. Because the mapping is linear, the Fréchet derivative and its adjoint are identical (Tarantola, 2005):

$$\left[\frac{\partial \vec{p}(\vec{m})}{\partial \vec{m}} \right] = \left[\frac{\partial \vec{p}(\vec{m})}{\partial \vec{m}} \right]^*. \tag{3.17}$$

We observe that equation 3.16 takes the same form as equation 3.14 and therefore the gradient of the misfit functional is equal to the mapping from the data space to the model space:

$$\frac{\partial E_{FWI}(\vec{m})}{\partial \vec{m}} = \delta \vec{m}'. \tag{3.18}$$

If we find a solution to the forward problem which corresponds to equation 3.15, we can identify the Fréchet derivatives and calculate the gradient by using equations 3.16 and 3.17. The small change in the data space $\delta\vec{p}'$ can then be interpreted as the data residuals (equation 3.13).

This is only to give the general idea of the adjoint state method. The whole derivation is not part of this work, so we refer to the works of Tarantola (1984) and Mora (1987) for more details. The actually used gradients are based on the elastic gradients by Köhn (2011). They were adapted to the acoustic case yielding

$$\begin{aligned} \frac{\partial E_{\mathrm{FWI}}}{\partial K} &= -\sum_{\mathrm{sources}} \int \mathrm{d}t \; \frac{p \cdot \tilde{p}}{\rho^2 v_{\mathrm{p}}^4}, \\ \frac{\partial E_{\mathrm{FWI}}}{\partial \rho} &= -\sum_{\mathrm{sources}} \int \mathrm{d}t \left(v_x \tilde{v}_x + v_y \tilde{v}_y \right), \end{aligned} \tag{3.19}$$

where p is the forward propagated pressure field and $\tilde{p}$ the backward propagated pressure field. Moreover, v_x and v_y denote the forward propagated particle velocity fields and $\tilde{v}_x$ and $\tilde{v}_y$ the backward propagated particle velocity fields. Overall, the gradients of the misfit functional are the zero lag cross-correlation of the forward and backward propagated pressure and velocity fields. In addition, it can be seen that they are now independent of the receiver number which is a beneficial feature regarding the computational efficiency. Because the gradients were derived in the parameterization of density ρ and bulk modulus K, but we need the gradients in the parameterization of density ρ and P-wave velocity v_{p}, a final transformation has to be performed:

$$\begin{aligned} \frac{\partial E_{\mathrm{FWI}}}{\partial v_{\mathrm{p}}} &= 2\rho v_{\mathrm{p}} \frac{\partial E_{\mathrm{FWI}}}{\partial K}, \\ \frac{\partial E_{\mathrm{FWI}}}{\partial \rho'} &= v_{\mathrm{p}}^2 \frac{\partial E_{\mathrm{FWI}}}{\partial K} + \frac{\partial E_{\mathrm{FWI}}}{\partial \rho}. \end{aligned} \tag{3.20}$$

3.1.2.3. Gradient preconditioning

As explained in the previous section the gradient is calculated by a zero lag cross-correlation of the forward and backward propagated wave fields. Due to geometrical spreading the amplitude of the wave fields decay with distance resulting in relatively high amplitudes in the vicinity of the sources and adjoint sources, i.e. receivers. Hence, the model update is concentrated on the vicinity of the sources and receivers which is not desired and can even prevent the convergence of the inversion. To overcome this problem and allow model updates to be distributed over the whole model two types of gradient preconditioning are used.

The first one is a local preconditioner which damps the high amplitudes in the vicinity of the sources. It is a circular taper which sets the gradient to zero at the source position and increases logarithmically to one until the boundary of the circle is reached. The preconditioning taper is applied shotwise to preserve information about the source positions from the neighbouring sources.

The second preconditioner is applied globally and also shotwise over the whole model. It represents an approximation of the inverse of the diagonal Hessian matrix elements $\boldsymbol{H}_a^{-1}$ and was developed by Plessix and Mulder (2004):

$$\boldsymbol{H}_{\mathrm{a}}^{-1}(\vec{x}_{\mathrm{s}},\vec{x}) = \left[\epsilon_{\mathrm{FWI}} + \int \mathrm{d}t \; |\vec{u}(\vec{x}_{\mathrm{s}},\vec{x},t)|^2 \left(\operatorname{asinh}\left(\frac{x_{\mathrm{r}}^{\max}(\vec{x}_{\mathrm{s}}) - x}{z}\right) - \operatorname{asinh}\left(\frac{x_{\mathrm{r}}^{\min}(\vec{x}_{\mathrm{s}}) - x}{z}\right)\right)\right]^{-1} . \tag{3.21}$$

The parameter $\vec{u}$ represents the forward modeled wave field, $\vec{x}_{\mathrm{s}}$ the source position and $x_{\mathrm{r}}^{\max}$ and $x_{\mathrm{r}}^{\min}$ the maximum and minimum receiver position of each source, respectively. The hyperbolic functions result from an approximation of the receiver Green's function. This preconditioner does not only damp the gradient in the vicinity of the sources but also increases values in those areas with poor illumination. The parameter ϵ_{FWI} denotes the water level and ensures the stability of the matrix inversion in the case of very small amplitudes of the wave field.

3.1.2.4. Hessian calculation with multi-parameter L-BFGS

The preconditioner of Plessix and Mulder (2004) represents only a rough approximation of the inverse of the true Hessian matrix. Especially, it ignores the off-diagonal elements which are crucial to reduce trade-offs between different model parameters. In the following, we shortly explain the quasi-Newton limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method based on Nocedal and Wright (2006), Brossier (2011), and Wittkamp (2016).

In contrast to the full BFGS algorithm curvature information from only the most recent iterations is used. Because the curvature information from earlier iterations usually has no significant influence on the current Hessian matrix, the convergence of the inversion should not be affected negatively. This approach has the advantage of reduced memory storage demands compared to the full BFGS method and much less computation time compared to the calculation of the full Hessian matrix. For the approximation of the Hessian matrix the gradient differences

$$y_n = \nabla_{\vec{m}} E_{n+1}(\vec{m}) - \nabla_{\vec{m}} E_n(\vec{m}) \tag{3.22}$$

and the model differences

$$s_n = \vec{m}_{n+1} - \vec{m}_n \tag{3.23}$$

of the most recent iterations are needed. Here, n denotes the iteration number. According to Nocedal and Wright (2006) a number of 3 to 20 iterations is sufficient. We use the gradient and model differences of 10 previous iterations in every inversion. Using the gradient and model differences the L-BFGS algorithm directly provides the model update

$$\Delta\vec{m}_n = -\boldsymbol{H}_{\mathrm{a},n}^{-1} \cdot \nabla_{\vec{m}} E_n(\vec{m}). \tag{3.24}$$

Introducing the step length α_n we rewrite the inversion scheme represented by equation 2.23:

$$\vec{m}_{n+1} = \vec{m}_n + \alpha_n \cdot \Delta\vec{m}_n. \tag{3.25}$$

The step length is introduced to make sure that the model update fulfills some kind of convergence condition. The convergence conditions used in this study are explained in the subsequent section. Before the algorithm can be started an initial guess of the Hessian matrix is required. This initial guess consists of the gradient and model differences of the previous iteration. For this reason in the first iteration of every workflow stage (see section 4.3) a classical steepest descent update is performed using a parabolic line search to estimate the step length (Nocedal and Wright, 2006). To ensure the stability of the L-BFGS algorithm it is reset at the beginning of each workflow stage.

A major advantage of the L-BFGS algorithm compared to the steepest descent method is that it yields a model update with physical meaningful units (m/s or kg/m^3, respectively). This is the case because the gradient of the misfit functional (unit: $\frac{\mathrm{Pa}^2}{\mathrm{m/s}}$ and $\frac{\mathrm{mGal}^2}{\mathrm{kg/m}^3}$ respectively) is weighted by the inverse of the approximated Hessian matrix (unit: $\frac{\mathrm{m}^2/\mathrm{s}^2}{\mathrm{Pa}^2}$ and $\frac{\mathrm{kg}^2/\mathrm{m}^6}{\mathrm{mGal}^2}$ respectively).

A small adaption of the algorithm is needed because we use different parameter classes in this study (P-wave velocity and density). Thus, the model and gradient differences show different orders of magnitude and an individual L-BFGS optimization for each parameter class would be necessary. In this case, the approximated Hessian matrix could not be used efficiently to reduce trade-offs between different parameter classes. To overcome this problem we use a dimensionless L-BFGS algorithm with normalized parameter classes. In the end, the model update is denormalized to obtain again physical meaningful values (cf. Brossier, 2011; Wittkamp, 2016).

3.1.2.5. Wolfe line search

After the model update is calculated a line search is executed to find a step length which decreases the misfit functional significantly. In this study we use the Wolfe conditions as decision criterion (Nocedal and Wright, 2006). The first Wolfe condition is called *sufficient decrease condition* and states:

$$E(\vec{m} + \alpha \cdot \Delta\vec{m}) \leq E(\vec{m}) + c_1 \cdot \alpha \cdot \nabla_{\vec{m}} E(\vec{m})^T \cdot \Delta\vec{m}. \tag{3.26}$$

The parameter c_1 is set to zero during every inversion. Consequently, the misfit only has to decrease in general instead of decreasing below the straight line defined by the right hand side of equation 3.26. The second Wolfe condition, named *curvature condition*, is written as follows:

$$\nabla_{\vec{m}} E(\vec{m} + \alpha \cdot \Delta\vec{m})^T \cdot \Delta\vec{m} \geq c_2 \cdot \nabla_{\vec{m}} E(\vec{m})^T \cdot \Delta\vec{m} \tag{3.27}$$

It ensures that too small steps are refused by comparing the curvature of the misfit functional. The slope of the misfit functional after the update has to be greater than c_2 times the initial slope. This indicates that a significant decrease of the misfit functional is possible. The parameter c_2 is set to 0.9 during every inversion.

The step length estimation is done by a line search algorithm for the Wolfe conditions

(Nocedal and Wright, 2006; Métivier, 2014; Wittkamp, 2016). It is initiated with the step length $\alpha = 1$. If the step length does not fulfill the Wolfe conditions, it will be reduced and tested again. Note that as long as the first Wolfe condition is fulfilled and only the second condition is not fulfilled the step length will be increased and tested again. After five unsuccessful tests the step length which reduces the misfit functional most is used. If the misfit functional cannot be reduced the L-BFGS algorithm will be terminated and restarted in the next iteration. The inversion will go on with the next workflow stage and if there is no further workflow stage the inversion is finished.

As additional criterion for the termination of the inversion or the change of a workflow stage we use the following abort criterion AC with n as the iteration number:

$$AC_n \text{ (in \%)} = 100 \cdot \frac{E_{n-2} - E_n}{E_{n-2}} \tag{3.28}$$

It is checked whether the relative reduction of the misfit functional is big enough. The value of the abort criterion is set manually.

3.2. Gravity code

3.2.1. Choice of parameterization

Generally there are two basic strategies to interpret gravity data (cf. Last and Kubik, 1983). From equation 2.13 it can be seen that the gravitational acceleration depends on the one hand on the densities in the subsurface and on the other hand on a geometrical factor which contains the distance to the source and the shape of the source.

The first strategy is to fix the density values and change only the geometry of the sources. In this case a homogeneous body is obtained. A disadvantage is that the relationship between the data and model parameters is nonlinear and the flexibility of the geometry is restricted.

The second approach is to fix the geometry and change only the density values. Here, the relationship between the data and model parameters is linear and a high flexibility is achieved. However, a major disadvantage arises from a strong ambiguity of the corresponding inverse problem that makes the use of regularization necessary.

In this thesis we use the second strategy because it can easily be implemented in our existing finite-difference forward code. If we used the first strategy we would need additional model parameters and an interface between the finite-difference grid and the additional parameters. Although the inverse problem of gravity is then linear, we use a nonlinear quasi-Newton inversion algorithm. Otherwise, we could not use the same inversion code for the acoustic FWI, the gravity inversion and the joint inversion. The increased ambiguity that comes with this parameterization should be restricted by a simultaneous fitting of the pressure seismograms in the joint inversion.

3.2.2. Forward problem

3.2.2.1. Gravitational field by rectangular prisms

In gravity interpretation the forward problem describes the calculation of the gravitational field by a given density distribution. In subsection 2.2.3 we showed how to calculate the vertical component of the gravitational field at a specific observation point analytically using equation 2.13. This formulation is still very general because the integration has not been specified yet. For each observation point the volume integral has to be evaluated which is not a trivial task. For complicated source shapes this integral can become arbitrarily difficult to be calculated computationally. Therefore we will fix the geometry as mentioned above by using rectangular prisms (Blakely, 1995; Nagy et al., 2000). The geometry can be seen in figure 3.2 where the blue prism is one voxel (*vo*lumetric pi*xel*) and all voxels have the same volume.

Now we show how the vertical component of the gravitational field can be calculated with this parameterization. At first, we simplify the integral by shifting the coordinate system such that the origin of the new system coincides with the observation point $P(x' = 0, y' = 0, z' = 0)$:

$$\begin{aligned}
x' &= x - x_S, \\
y' &= y - y_S, \\
z' &= z - z_S, \\
r' &= \sqrt{x'^2 + y'^2 + z'^2}.
\end{aligned} \tag{3.29}$$

The new coordinates x', y' and z' represent the distance between source and observation point and thus are more intuitive when considering the distance dependent gravitational field. Following the principle of superposition we calculate the influence of each prism separately and sum the influence of all N prisms to obtain the complete y-component of the gravitational field at one observation point. Consequently, we substitute the new coordinates into equation 2.13 and perform the integration over the corners of the prism, x_i, y_i and z_i and accordingly x'_i, y'_i and z'_i:

$$\begin{aligned}
g_{y,\text{prism}}(P) &= -G \int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} \frac{y - y_S}{r^3}\, \rho_S(x_S, y_S, z_S)\, \mathrm{d}x_S\, \mathrm{d}y_S\, \mathrm{d}z_S \\
&= G \int_{z'_1}^{z'_2} \int_{y'_1}^{y'_2} \int_{x'_1}^{x'_2} \frac{y'}{r'^3}\, \rho_S(x - x', y - y', z - z')\, \mathrm{d}x'\, \mathrm{d}y'\, \mathrm{d}z' \\
&= G\, \rho_S(x - x', y - y', z - z')\, K,
\end{aligned} \tag{3.30}$$

where K denotes the geometrical kernel. The geometrical kernel is constant for every combination of observation point and subsurface grid point. Therefore, the gravitational field is a simple weighting of the density by the gravitational constant and the geometrical kernel. The solution of this integral is given by Blakely (1995) and Nagy et al. (2000) amongst others. The geometrical kernel can then be written in the following way:

$$
\begin{aligned}
K &= \int_{z_1'=z-z_1}^{z_2'=z-z_2} \int_{y_1'=y-y_1}^{y_2'=y-y_2} \int_{x_1'=x-x_1}^{x_2'=x-x_2} \frac{y'}{r'^3} \,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}z' \\
&= \left(y' \cdot \arctan\left(\frac{x'z'}{y'r'}\right) - x' \cdot \ln(z'+r') - z' \cdot \ln(x'+r') \right) \Bigg|_{x_1'}^{x_2'} \Bigg|_{y_1'}^{y_2'} \Bigg|_{z_1'}^{z_2'} \\
&= \left(x' \cdot \ln(z'+r') + z' \cdot \ln(x'+r') - y' \cdot \arctan\left(\frac{x'z'}{y'r'}\right) \right) \Bigg|_{x_2'}^{x_1'} \Bigg|_{y_2'}^{y_1'} \Bigg|_{z_2'}^{z_1'} .
\end{aligned}
\tag{3.31}
$$

Note that the density has to be constant within each prism, otherwise this approach would not be valid. Also note that we assumed implicitly a flat earth approximation which should be justified in the considered model dimensions (see section 4.1). In a discretized way we can write equation 2.13 as follows:

$$
g_m = G \sum_{n=1}^{N} \rho_n K_{mn}, \tag{3.32}
$$

where m denotes the observation point and n the grid point.

3.2.2.2. Boundary conditions

The implementation of proper boundary conditions is a crucial part of gravity modeling both in respect of accuracy and computational efficiency (cf. Wehner, 2015). If completely neglected the gravitational field would decrease rapidly at the boundaries. As a result, the artificial gravity anomalies could be misinterpreted as a density variation in the subsurface. To avoid such undesired effects the model is elongated in positive and negative x- and z-direction. This is done by a continuation of the boundary grid point density values in the respective direction. The length of the continuation is symmetric and specified by b_x and b_{prism}. Furthermore, we assume that the density model possesses no significant lateral variations in greater depth[1]. Thus, an elongation in the y-direction is not needed because a constant half space has the same influence on every gravimeter station.

The elongation in z-direction is only a multiplication with a constant and consequently has no influence on the calculation time. In contrast, the elongation in x-direction expands the grid and thus has a significant influence on the computation time.

A second way of mitigating the boundary effects is to apply a Bouguer plate reduction. Thereby, a plate with a constant density value is subtracted from the density model before calculating the gravitational field. In this case, only relative gravity variations are considered and not the absolute field strength. Moreover, background density models which are not constant can be used by using a priori information.

In section 4.2.2, the effect of different continuation lengths and the Bouguer plate reduction is demonstrated.

---

[1] This is actually the case (see section 4.1).

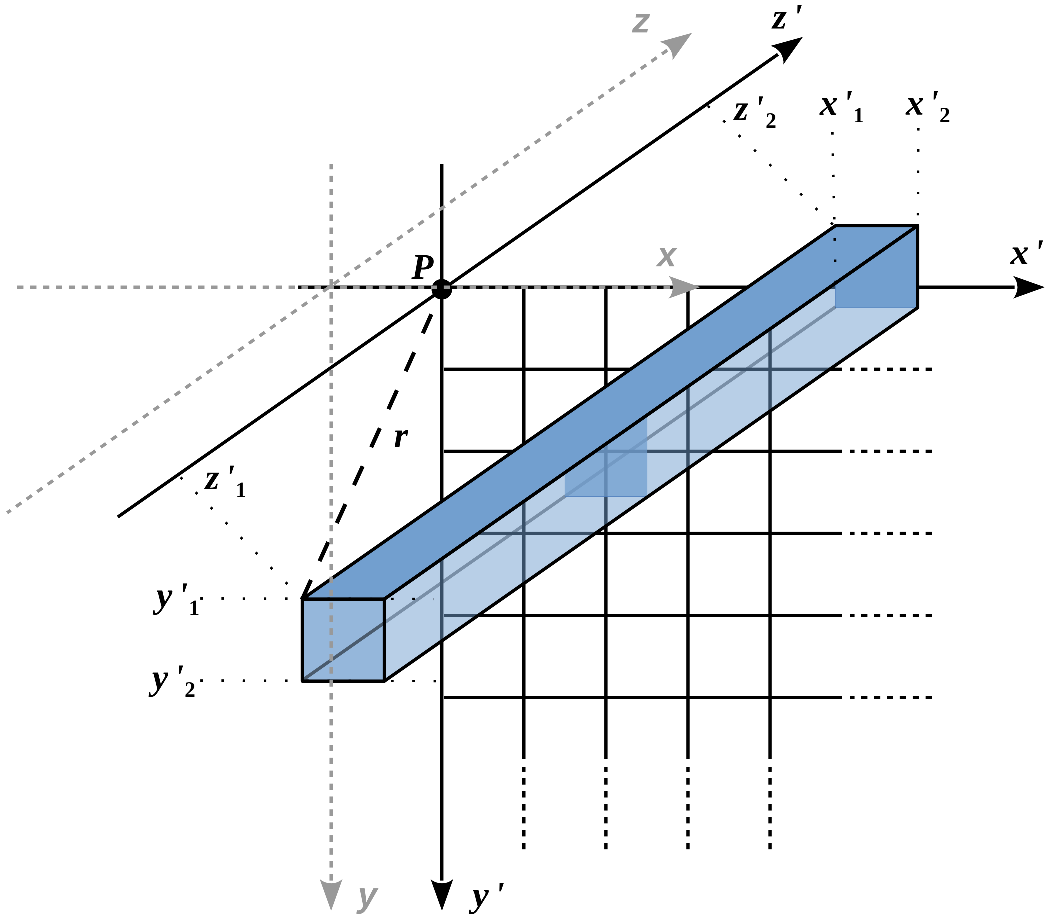


Figure 3.2.: Parameterization with rectangular prisms (blue) which are integrated into the finite-difference grid. The grey coordinate system corresponds to the original finite-difference grid, the black coordinate system is centered around the observation point *P* as origin.

3.2.3. Inverse problem

The inverse problem of gravity is also solved with the quasi-Newton L-BFGS algorithm and a Wolfe line search. Therefore, again the gradient of the misfit functional is needed.

3.2.3.1. Misfit functional

First, we define the misfit functional

$$E_{\mathrm{GRAV}}(\vec{\rho}) = \frac{1}{2}\delta\vec{g}_y^T\delta\vec{g}_y \tag{3.33}$$

with the data residuals

$$\delta\vec{g}_y = \vec{g}_{y,\mathrm{mod}}(\vec{\rho}) - \vec{g}_{y,\mathrm{obs}}. \tag{3.34}$$

The modeled data $\vec{g}_{y,\mathrm{mod}} := \vec{g}_{y,\mathrm{mod}}(\vec{x}_{\mathrm{stat}}, \vec{\rho})$ depends explicitly on the position of the gravimeter station $\vec{x}_{\mathrm{stat}}$ and the density model $\vec{\rho}$. The observed data $\vec{g}_{y,\mathrm{obs}} := \vec{g}_{y,\mathrm{obs}}(\vec{x}_{\mathrm{stat}})$ only depends on the position of the gravimeter station $\vec{x}_{\mathrm{stat}}$. The explicit dependence on the gravimeter station is omitted as in the acoustic case.

3.2.3.2. Gradient calculation

The gradient calculation is straightforward. By making use of equation 3.32 we obtain (Wehner, 2015):

$$\begin{aligned}
J_{E_{\mathrm{GRAV}}} &= \left\{\frac{\partial E_{\mathrm{GRAV}}(\rho_n)}{\partial \rho_n}\right\}_{n=1,\dots,N} && (3.35)\\
&= \left\{\frac{\partial}{\partial \rho_n}\left[\frac{1}{2}\sum_{m=1}^{M}(G\sum_{n=1}^{N}\rho_n K_n - g_{y,\mathrm{obs}})_m^2\right]\right\}_{n=1,\dots,N} && (3.36)\\
&= \left\{\frac{1}{2}\sum_{m=1}^{M}[(GK_n)_m \cdot 2(G\sum_{n=1}^{N}\rho_n K_n - g_{y,\mathrm{obs}})_m]\right\}_{n=1,\dots,N} && (3.37)\\
&= \left\{G\sum_{m=1}^{M}(\delta g_y K_n)_m\right\}_{n=1,\dots,N} \quad . && (3.38)
\end{aligned}$$

Here, n denotes again the grid point, N the total number of grid points, m the observation point and M the total number of observation points. If we interprete the forward problem as a projection of the density values onto the observation plane, we can interprete the gradient as the projection of the data residuals $(\delta g_y)_m$ from the observation plane to the model. Consequently, the value of the gradient at one grid point is the sum of the data residuals weighted by the corresponding geometrical kernel and the gravitational constant.

3.2.3.3. Gradient preconditioning

The gravitational field, more precisely the geometrical kernel, decays rapidly with increasing distance between observation point and source point. Thus, the gradient has naturally relatively high values at the top of the model and low values in the deeper parts. This would restrict the model update to the upper parts of the model which is not desired because it does not reflect the real geological situation in general. To distribute the model update over the whole model with a similar probability we introduce a depth weighting function as preconditioner (Li and Oldenburg, 1998; Boulanger and Chouteau, 2001).
The depth weighting function is an exponential function of the depth y with an adjustable exponent β. It represents an approximation of the diagonal elements of the Hessian matrix and is written as follows:

$$\boldsymbol{H}_{\mathrm{a}}^{-1} = w_{\text{depth}} = (y + \epsilon_{\text{GRAV}})^{\beta} \, . \tag{3.39}$$

The water level, $\epsilon_{\text{GRAV}} = 0.0005$, is introduced for stability reasons. The exponent β is set to 0.8. Tests showed that this value provides good results and it is also within the recommended range (Li and Oldenburg, 1998; Wehner, 2015). An illustrative comparison of the geometrical kernel for one and for multiple observation points and with or without depth weighting, respectively, can be found in the work of Wehner (2015).
It is also possible to calculate the true diagonal elements of the Hessian matrix efficiently:

$$\left\{ \frac{\partial^2 E_{\text{GRAV}}(\rho_n)}{\partial \rho_n^2} \right\}_{n=1,\dots,N} = \left\{ \frac{\partial}{\partial \rho_n} \left(G \sum_{m=1}^{M} (\delta g_z K_n)_m \right) \right\}_{n=1,\dots,N} = \left\{ G^2 \sum_{m=1}^{M} (K_n)_m^2 \right\}_{n=1,\dots,N} . \tag{3.40}$$

Compared to the depth weighting the true diagonal Hessian matrix takes the radial decay into account. Moreover, the high amplitudes in the vicinity of the observation points are considered. However, tests revealed that the gravity inversions with depth weighting and with the true diagonal Hessian matrix as preconditioner, respectively, do not differ significantly. Additionally, further tests showed that in the case of gravity gradiometry data the true diagonal elements of the Hessian matrix are not sufficient to overcome the natural decay of the geometrical kernel. In conclusion, we use only the depth weighting in this study because of its flexibility and generality.

3.3. Joint inversion

A crucial part of this study is the implementation of a simultaneous joint inversion of acoustic and gravimetric data. The goal is to derive a P-wave velocity model and a density model which explain both data sets. The individual inversions are completely independent and have to be connected somehow for this purpose. In ealier studies empirical and analytic relations between different model parameters have been used as well as structural similarities (e.g. Colombo and De Stefano, 2007). However, the validity of empirical relationships and structural similarity is not always justified, especially not in cases of salt or basalt structures. A more general and sophisticated approach is to connect the gradients

of the misfit functionals during the nonlinear local optimization directly. Minimized is in any case a joint misfit functional.
The basic scheme of one iteration within the inversion framework is shown in figure 3.3. The forward modeling, misfit calculation and gradient calculation and preconditioning of both methods are done independently. Afterwards the misfit functionals and gradients are connected. How this is done will be explained in the following. The joint gradients are then smoothed with a median filter to avoid small-scale artifacts. The filter length is approximately the minimum wavelength excited in the model. Finally, the model update is calculated and applied by using the multi-parameter L-BFGS algorithm and the Wolfe line search. This process is repeated until the abort criterion is fulfilled or no step length can be found which reduces the misfit.

3.3.1. Joint misfit functional

The joint misfit functional is the weighted sum of the individual misfit functionals:

$$\begin{aligned} E_{\mathrm{JOINT}}(\vec{v}_{\mathrm{p}},\vec{\rho}) &= E_{\mathrm{FWI}}(\vec{v}_{\mathrm{p}},\vec{\rho}) + \lambda_1 \cdot E_{\mathrm{GRAV}}(\vec{\rho}) \\ &= \frac{1}{2}\cdot \delta\vec{p}^{T}\delta\vec{p} + \lambda_1 \cdot \frac{1}{2}\cdot \delta\vec{g}_y^{T}\delta\vec{g}_y. \end{aligned} \tag{3.41}$$

The weighting is necessary because the magnitude of the individual misfit functionals differs in general. Without a weighting effectively only one misfit functional is minimized. The weighting parameter λ_1 is defined as follows:

$$\lambda_1 = \gamma \frac{E_{\mathrm{FWI}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{E_{\mathrm{GRAV}}(\vec{\rho})}. \tag{3.42}$$

Apparently the gravity misfit functional is normalized to be in the same scale as the seismic misfit functional. The parameter γ is chosen manually and helps to control the influence of each individual method. If γ is set for example to 0.1, the ratio of the individual contributions is 1:10. Note also that λ_1 is always calculated during the first iteration of every workflow stage and is constant during each workflow stage (see section 4.3).

3.3.2. Joint gradients

The gradients are combined in a similar way:

$$\frac{\partial E_{\mathrm{JOINT}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{\partial\vec{\rho}} = \frac{\partial E_{\mathrm{FWI}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{\partial\vec{\rho}} + \lambda_2 \cdot \frac{\partial E_{\mathrm{GRAV}}(\vec{\rho})}{\partial\vec{\rho}} \tag{3.43a}$$

$$\frac{\partial E_{\mathrm{JOINT}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{\partial\vec{v}_{\mathrm{p}}} = \frac{\partial E_{\mathrm{FWI}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{\partial\vec{v}_{\mathrm{p}}}. \tag{3.43b}$$

Because the gravitational acceleration is not sensitive to the P-wave velocity only the gradients with respect to the density have to be added. The weighting parameter λ_2 is then defined as

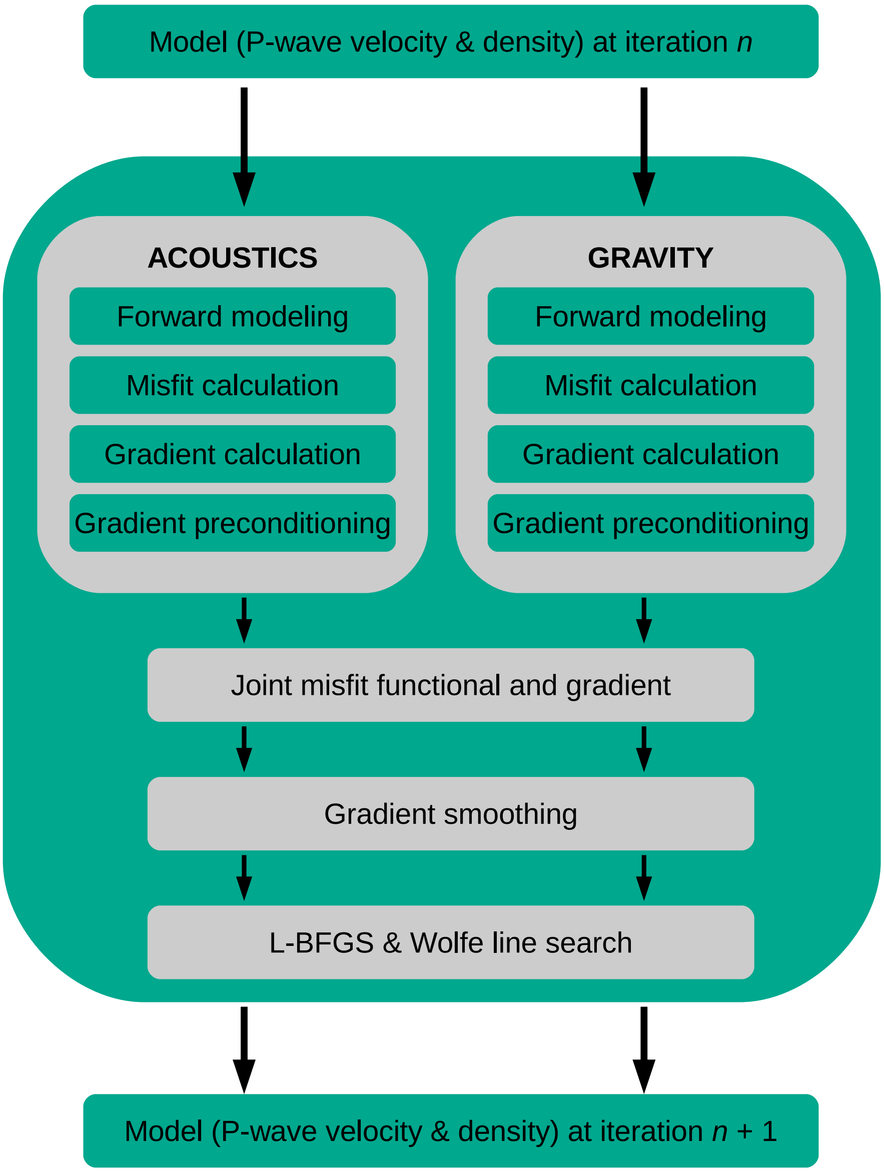

Model (P-wave velocity & density) at iteration *n*
ACOUSTICS
Forward modeling
Misfit calculation
Gradient calculation
Gradient preconditioning
GRAVITY
Forward modeling
Misfit calculation
Gradient calculation
Gradient preconditioning
Joint misfit functional and gradient
Gradient smoothing
L-BFGS & Wolfe line search
Model (P-wave velocity & density) at iteration *n* + 1

Figure 3.3.: Flow chart of one iteration of the inversion algorithm.

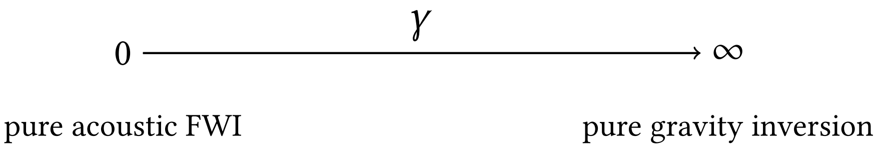


Figure 3.4.: Weighting between acoustic full-waveform and gravity inversion.

$$\lambda_2 = \gamma \frac{\max\left(\left|\frac{\partial E_{\mathrm{FWI}}(\vec{v}_{\mathrm{p}},\vec{\rho})}{\partial \vec{\rho}}\right|\right)}{\max\left(\left|\frac{\partial E_{\mathrm{GRAV}}(\vec{\rho})}{\partial \vec{\rho}}\right|\right)}. \tag{3.44}$$

Again, λ_2 is calculated during the first iteration of every workflow stage and is constant during each workflow stage.

Finally, in figure 3.4 the weighting between the acoustic full-waveform and gravity inversion is illustrated. In chapter 5 the results of a pure FWI ($\gamma = 0$), three different pure gravity inversions ($\gamma = 10^6$) and two joint inversions ($\gamma = 0.1$ and $\gamma = 0.3$) are presented.

4. Model, forward modeling and inversion parameters

In this chapter we describe the salt dome model and the setup of the forward modeling and the inversion. The results shown in the subsequent chapter represent a reconstruction test, i.e. the true model is known ab initio. Based on the true model we calculate pseudo-observed data with the forward engines explained in sections 3.1.1 and 3.2.2. The goal is to derive the true model from an initial guess by fitting the forward modeled data to the pseudo-observed data. Therefore, in the first section we introduce the true P-wave velocity and density model, in the second section the pseudo-observed data and the forward modeling parameters and finally in the third section the inversion parameters.

4.1. Salt dome model

The true model is a typical salt dome structure from Northern Germany and was provided by the company TERRASYS Geophysics. The shape of the salt dome and the density distribution were adapted for our purpose. It was chosen mainly for two reasons. First, gravity is only sensitive to lateral inhomogeneities and thus a model with strong lateral density variations ensures that gravity can provide a significant contribution to the joint inversion. Secondly, a current challenge in exploration geophysics is the interpretation of salt and basalt structures because empirical relationships like Gardner's relation (Gardner et al., 1974) are generally not valid in this context. Additionally, the high reflectivity of salt may cause a poor illumination with seismic waves resulting in difficulties during the inversion or migration.

The true density model is shown in figure 4.1. It consists of different layers of sedimentary rock with a dominant salt dome structure between approximately 7 km and 14 km. The whole model is 16 km wide and 5 km deep. The grid spacing Δh is 12.5 m resulting in a grid size of 1280 times 400 grid points. During the formation of the dome due to the relatively small density of salt the sedimentary layers were pushed upwards resulting in dipping structures close to the dome. At the top of the salt dome there is a thin high density layer, also an artifact of the formation process. Moreover, the right flank of the dome is covered by another high density layer. The different rock types marked in figures 4.1 and 4.2 are listed in the following:

The true P-wave velocity model is derived from the density model by Gardner's relation (Gardner et al., 1974):

$$v_\mathrm{p} = 0.3048 \cdot \left(\frac{\rho}{230}\right)^4 . \tag{4.1}$$

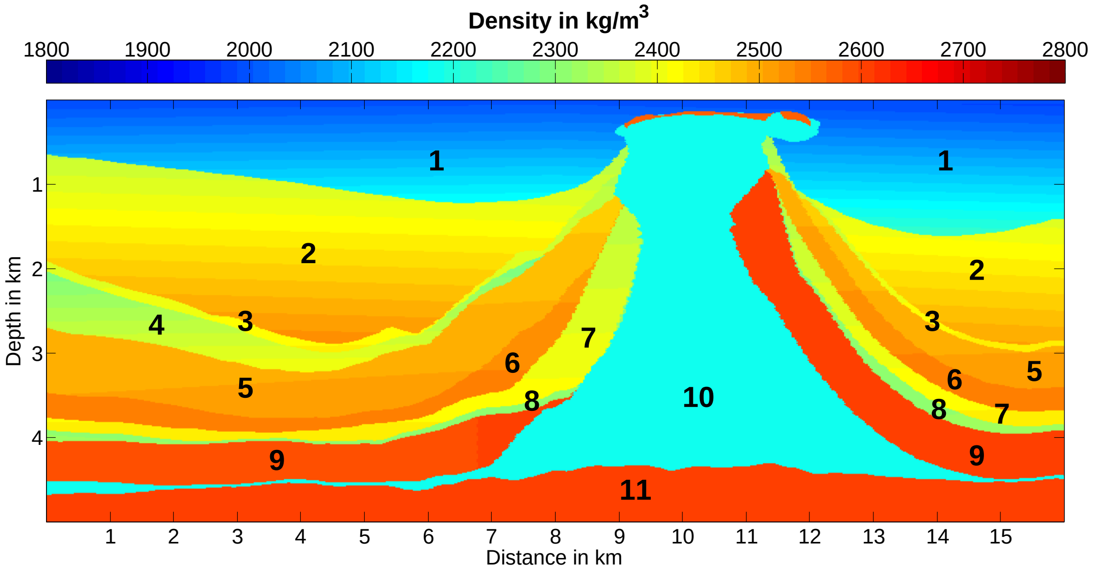


Figure 4.1.: True density model of the salt dome structure. An explanation of the different geological bodies can be found in table 4.1.

Table 4.1.: Explanation of the geological bodies in the true P-wave velocity and density model.

1. Tertiary
2. Upper Cretaceous
3. Lower Cretaceous
4. Jurassic
5. Keuper
6. Muschelkalk
7. Roet 1
8. Roet 2
9. Buntsandstein
10. Zechstein
11. Pre-Zechstein

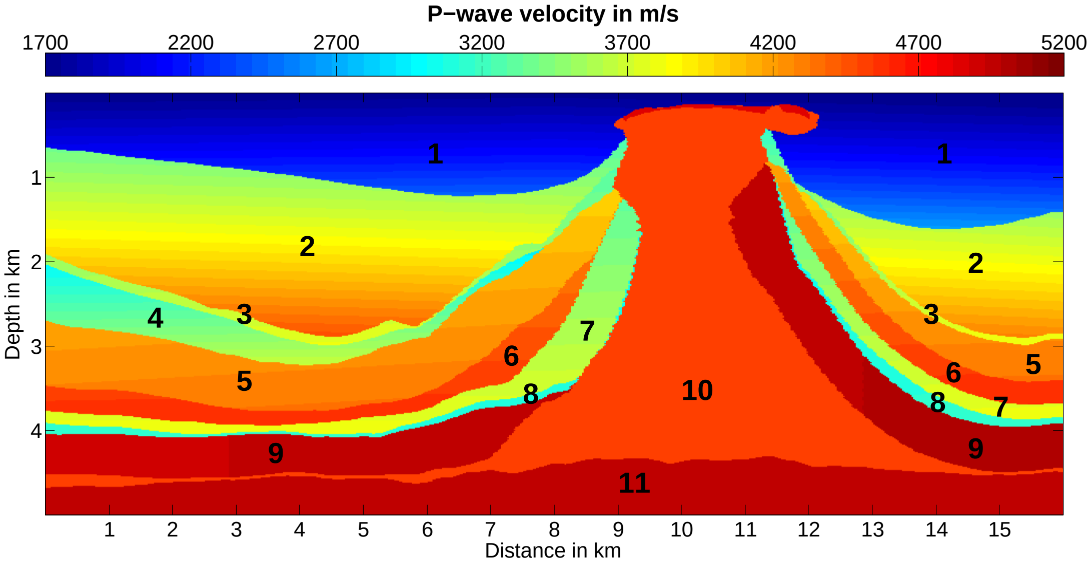


Figure 4.2.: True P-wave velocity model of the salt dome structure. An explanation of the different geological bodies can be found in table 4.1.

Here, the P-wave velocity v_p is given in m/s and the density in kg/m^3. The P-wave velocity values of the salt dome are set manually to 4500 m/s because Gardner's relation is not valid for salt rocks. For the salt density of 2180 kg/m^3 Gardner's relation yields a P-wave velocity of about 2460 m/s. In contrast, in literature (cf. Kearey et al., 2013) values between 4500 m/s and 5000 m/s are mentioned. The true P-wave velocity model is shown in figure 4.2.

4.2. Pseudo-observed data and forward modeling parameters

4.2.1. Acoustics

The acoustic forward modeling parameters are displayed in table 4.2. Both, the pseudo-observed data as well as the forward modeled data during the inversion are calculated with the same parameters.

An exemplary shot of the pseudo-observed data, i.e. the forward modeled data with the true model, is given in figure 4.3 for shot 6 at a distance of 4500 m. Note that the seismogram is trace-normalized for demonstrative reasons whereas the data is not normalized during the inversion. The wavefield is rather complex and can be divided into four groups of wave types: direct wave, reflections from sedimentary layers, salt dome reflections and refracted and scattered waves. Especially complicated is the behaviour of the wavefield when it reaches the salt dome. First, some waves are refracted and guided towards the salt dome. There, some waves travel through the salt structure and reach the receivers at the opposite side. Others are scattered at the sharp edges at about 9 km distance or at the

Table 4.2.: Acoustic forward modeling parameters.

FD order	2
Grid spacing	12.5 m
Grid size	16 × 5 km
	1280 × 400 grid points
Recording time	6 s
Sampling rate	0.0015 s
Time steps	4000
Source type	Explosion
Source shape	Ricker wavelet
Central frequency	5 Hz
Number of shots	21
First shot	1000 m
Last shot	15000 m
Shot spacing	700 m
Shot depth	0 m
Number of receivers	156
First receiver	250 m
Last receiver	15750 m
Receiver spacing	100 m
Receiver depth	0 m
PML width	125 m
PML reference frequency	5 Hz
PML reference velocity	3500 m/s

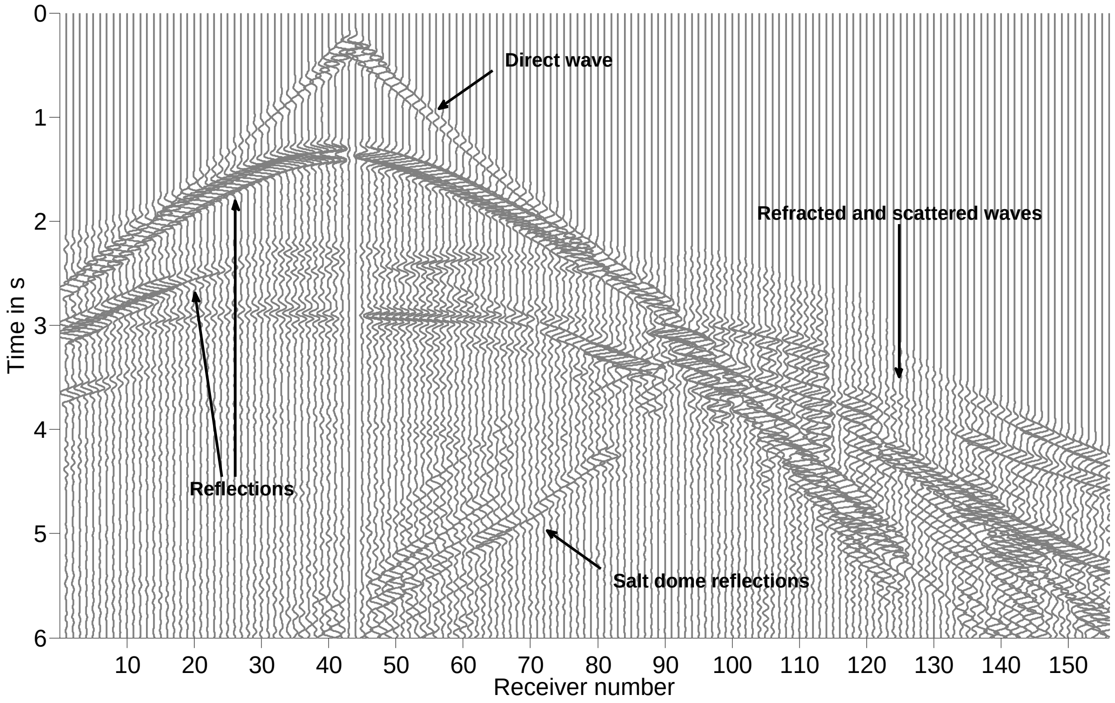


Figure 4.3.: Trace-normalized pseudo-observed data for shot 6.

overhang at 12 km distance, respectively. If the sources are located above the salt dome multiple reflections will be excited with high amplitudes. This will result in high values in the gradients and therefore further complicate the inversion process. The corresponding amplitude-frequency spectrum of shot 6 is shown in figure 4.4. The main frequency content is between 2 Hz and 14 Hz which corresponds to a minimal wavelength of about 120 m and a maximum wavelength of about 2500 m. If we calculate the mean P-wave velocity and assume 8 Hz to be the mean frequency of the signal we will obtain a mean wavelength of about 480 m.

4.2.2. Gravity

The gravimetric forward modeling parameters are shown in table 4.3. As in the acoustic case, the pseudo-observed data as well as the forward modeled data during the inversion are calculated with the same parameters.

The choice of the model extensions b_x and b_{prism} (cf. section 3.2.2.2) is now explained in more detail. Figure 4.5 shows the vertical gravity component for three different cases: variable extension in z-direction with Bouguer plate reduction and fixed extension in x-direction (b), variable extension in x-direction with Bouguer plate reduction and fixed extension in z-direction (c) and variable extension in x-direction without Bouguer plate reduction and with fixed extension in z-direction (d). For the Bouguer plate reduction the mean density of the model, $\bar{\rho} = 2360\,\text{kg/m}^3$, is chosen.

From figure 4.5b it can be seen that the extension in z-direction has no significant influence

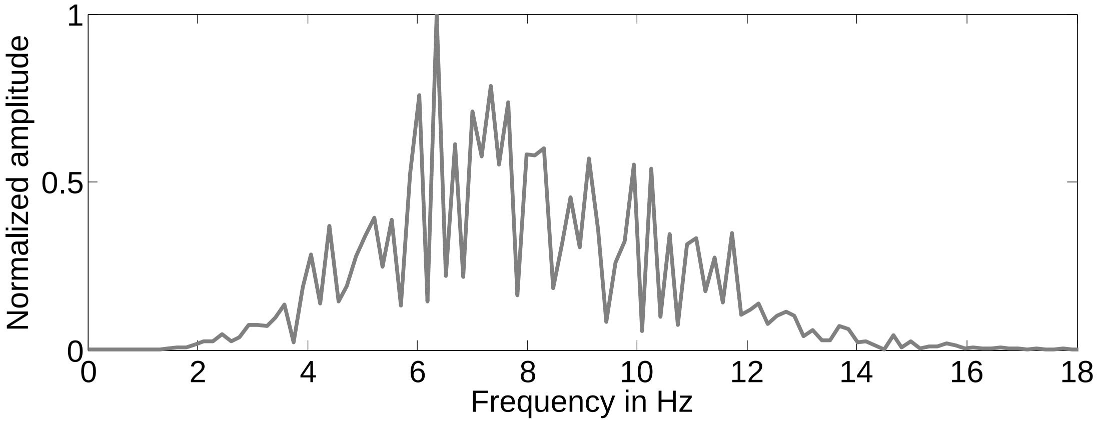


Figure 4.4.: Amplitude-frequency spectrum of the pseudo-observed data for shot 6.

Table 4.3.: Gravimetric forward modeling parameters.

Grid spacing	12.5 m
Grid size	16 × 5 km 1280 × 400 grid points
Number of gravity stations	104
First station	250 m
Last station	15700 m
Station spacing	150 m
Station depth	0 m
b_x	560 grid points (7 km)
b_{prism}	100 km
Bouguer plate	$\bar{\rho} = 2360\,\text{kg/m}^3$

on the shape of the gravity signal. A difference of the shape is only noticeable for the smallest extension of 1 km. With increasing extension the signal is mainly shifted with respect to its amplitude. When the extension reaches 100 km the signal does not change significantly if the model is further extended. In this context note that the black line is hidden beneath the orange line. This is not surprising because the gravitational field is distance dependent and the attraction from masses which are far away from the observation point diminishes. As a result, we choose 100 km as the perpendicular extension of the model as the signal converges at this point.

In figure 4.5c the influence of the extension in x-direction is demonstrated. In this case, the signals show a different shape at the boundaries. If the model is not extended (blue line) there is a steep increase in the signal at the boundaries. This can be explained with the negative amplitude of the gravitational attraction. The abrupt absence of masses causes the signal to be less negative. An extension of the model leads again to a convergence of the signal where the signal is decreasing first (320 grid points) at the left boundary and increasing monotonously later (560 grid points and 800 grid points). We choose an extension of 560 grid points because here the signal is monotonously increasing at both boundaries which is reasonable as the trend of the Upper Cretaceous layer is to be located closer to the observation point and it has a relatively high density compared to the Tertiary layer. Moreover, compared to the extension with 800 grid points the signal does not change significantly and the computation time is smaller.

The last figure (figure 4.5d) shows the gravitational field if no Bouguer plate reduction is applied. Apparently, the modeled signal suffers from strong boundary effects as the amplitude decreases at the boundaries. An extension of the model with 3200 or more grid points is necessary to obtain a monotonously increasing signal at both boundaries. In this case, the size of the model is five times larger and subsequently the computation time increases likewise. The total computation time of a whole joint inversion could then be larger in the order of hours in the worst case.

Finally, the pseudo-observed data that is used during the inversions is presented by the black line in figure 4.5b and 4.5c. The amplitude varies between approximately -4.4 mGal and 3.5 mGal. Note again that we only consider gravity effects relative to the Bouguer plate. The decreasing trend of the amplitude towards the center of the model is due to the salt dome whose density is relatively low and due to the inclination of especially the two uppermost sedimentary layers. The cause of the local maxima at 9.5 km and 11.5 km are the uplifting high density layers close to the salt dome, particularly the Buntsandstein layer at the right flank, and the thin Buntsandstein layer at the top of the salt dome.

Although we do not invert field data, an important practical question is whether these gravity variations could be resolved within actual field surveys. The internal precision of modern gravimeters is of the order of 0.01 mGal (Jacoby and Smilde, 2009). In land surveys, this is the best achievable precision. In marine surveys, the accuracy would be about 1 mGal. Thus, the gravity anomalies in our study are principally resolvable. To give a more detailed discussion commonly applied corrections to the field data are presented in the following based on Jacoby and Smilde (2009). The so-called *complete Bouguer anomaly* includes the following four reductions:

1. normal reduction, $\delta g_n = 0.8316 + 0.0782 \cdot sin^2(\phi)$ mGal,

2. height reduction, $\delta g_h = -0.3086\,\text{mGal/m}$,
3. Bouguer plate reduction, $\delta g_{BP} = 0.1117\,\text{mGal/m}$,
4. terrain reduction, $\delta g_{ter} \approx 0.1\,\text{mGal}$.

Here, ϕ denotes the latitude and for the Bouguer plate reduction the Bouguer density $\rho_{BP} = 2670\,\text{kg/m}^3$ is assumed (Jacoby and Smilde, 2009). These permanent effects take the dependence of the gravity on the latitude, on the height and on the topography into account. Additionally, time-varying effects like tides which are in the order of $\pm 0.15\,\text{mGal}$ have to be considered. It is obvious that all these factors are in a similar order as our target signal. Thus, the resolvability of geological density variations within actual field surveys depends also on the question how accurately these different effects can be corrected.

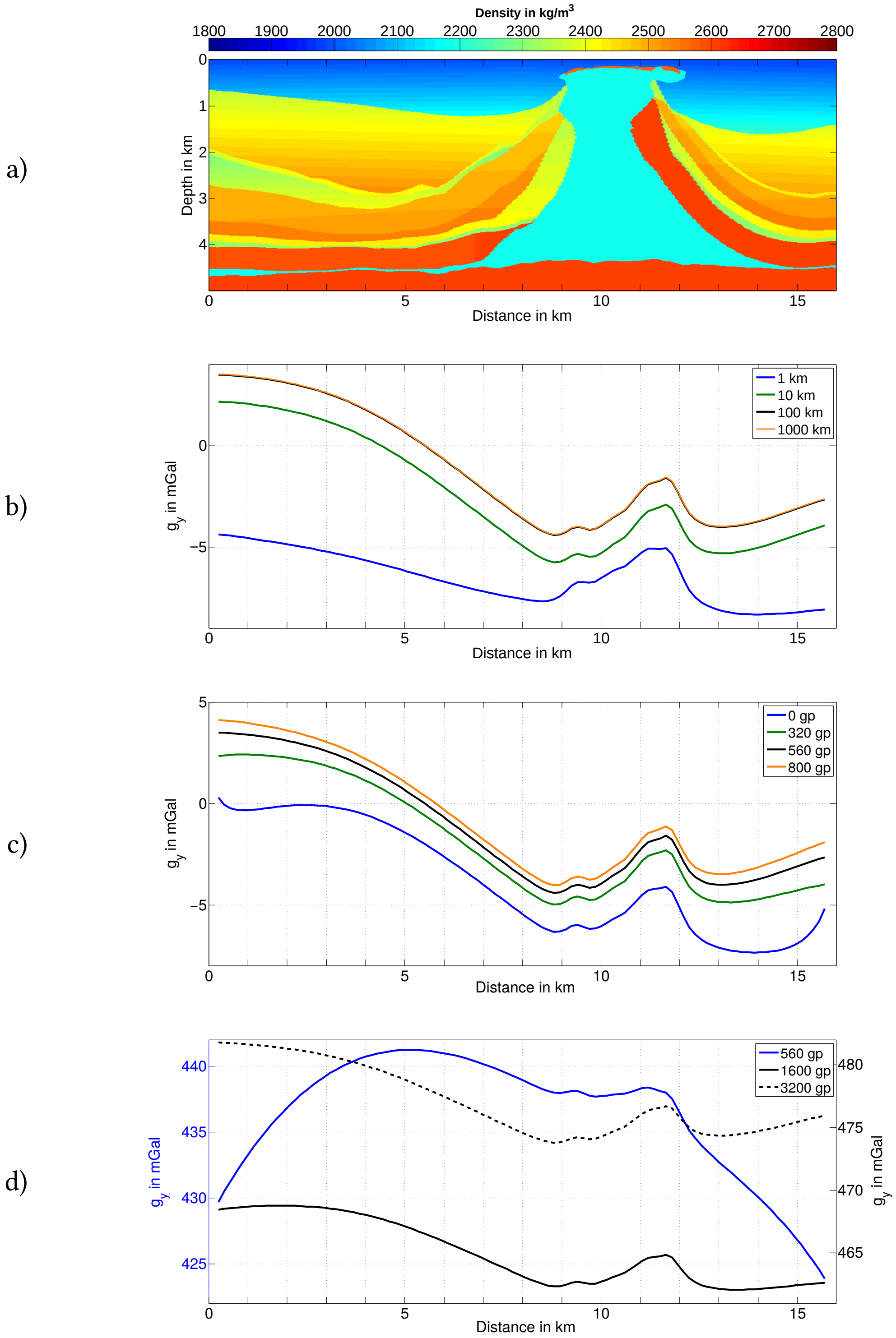


Figure 4.5.: (a) True density model, (b) Bouguer plate reduced vertical gravity component for different extensions in z-direction with an extension in x-direction of ±560 grid points (gp), (c) Bouguer plate reduced vertical gravity component for different extensions in x-direction with an extension in z-direction of ±100 km and (d) vertical gravity component for different extensions in x-direction with an extension in z-direction of ±100 km without Bouguer plate reduction.

4.3. Inversion parameters

In this section we introduce the inversion parameters. Because there are only few specific parameters for the acoustic FWI, the gravity inversion and the joint inversion we explain them together. All parameters are summarized in table 4.4 and some of them have already been explained in the previous chapters. Those parameters which are specific for the acoustic FWI are explained in the upper part of the table, those associated with the gravity inversion in the middle part of the table and those used in both inversions in the lower part. During the joint inversion all parameters are needed. As additional constraint we introduce a lower and upper bound of the P-wave velocity and the density. This is done to exclude unreasonable values or to restrict the model space, respectively.

In sections 3.1.2 and 2.3 we already mentioned the importance of a multi-stage inversion to reduce the nonlinearity and the *ill-posedness* of the acoustic inverse problem. The multi-stage approach (see table 4.5) consists of three properties. The first one is a sequential low-pass (LP) frequency filtering from low frequencies to high frequencies. We use a 4th order Butterworth low-pass filter with three different cutoff frequencies (3 Hz, 6 Hz and 9 Hz). Note that the frequency spectrum is not cut off sharply. At the cutoff frequency the signal is reduced to about 70%, thus higher frequencies are still included to some extent. The second property is an offset-dependent inversion. In section 4.2.1 we explained the complexity of the pressure wavefield. If all traces are inverted simultaneously the inversion will fail. Therefore, we first invert the data from far offsets and afterwards the data from near offsets. The corner distance of the offset-dependent trace killing is set to 2000 m because at this point the direct wave, the sedimentary reflections, the salt dome reflections and the refracted and scattered waves meet. In all inversion steps traces within a range of 160 m from the source are neglected because of their high amplitudes. The third and final property is a single parameter inversion at the beginning. If the seismograms are first roughly fitted by adjusting only the P-wave velocity model, the density model will suffer less from trade-off effects. At this stage the number of iterations is fixed to three.
The joint inversion workflow is almost identical to the acoustic FWI workflow. The only difference is that the weighting parameter γ is set additionally during stages 3 to 8. As mentioned before two joint inversions are presented in the subsequent chapter ($\gamma = 0.1$ and $\gamma = 0.3$).

Table 4.4.: Inversion parameters: specific FWI parameters (top), specific gravity inversion parameters (middle), general parameters (bottom), joint inversion parameters (all).

Low-pass frequency filter	4th order Butterworth
Low-pass frequency filter steps	3 Hz
Median gradient smoothing	100 m (8 grid points)
Circular logarithmic source taper	250 m (20 grid points)
Approx. Hessian water level ϵ_{FWI}	0.005
Minimum/maximum allowed P-wave velocity	1400/5400 m/s
Approx. Hessian water level ϵ_{GRAV}	0.0005
Gravity gradient depth weighting β	0.8
Minimum/maximum allowed density	1600/3000 kg/m^3
L-BFGS model and gradient differences	10 previous iteratons
Wolfe condition parameter c_1	0.0
Wolfe condition parameter c_2	0.9

Table 4.5.: Multi-stage workflow of the acoustic FWI.

Stage	Update		*AC* in %	LP filter in Hz	Offsets
	v_{p}	ρ			
1	yes	no	3 iterations	3	> 2000 m
2	yes	no	3 iterations	3	< 2000 m
3	yes	yes	1	3	> 2000 m
4	yes	yes	1	3	< 2000 m
5	yes	yes	1	6	> 2000 m
6	yes	yes	1	6	< 2000 m
7	yes	yes	1	9	> 2000 m
8	yes	yes	1	9	< 2000 m

5. Inversion results

In this chapter we present the results of the pure acoustic FWI, three gravity inversions with different initial models (constant, gradient and smoothed) and two joint inversions with different weighting parameters ($\gamma = 0.1$ and $\gamma = 0.3$). Finally, we investigate how the density model is resolved in the joint inversion in comparison to both individual inversions.

5.1. Acoustic full-waveform inversion

For the reconstruction test with the acoustic FWI a smoothed version of the true model is used as initial model (see middle row of figure 5.1). The kinematic background information is already included to prevent cycle-skipping or trapping in a local minimum, respectively. The goal of the reconstruction test is consequently the reconstruction of the different sedimentary layers and the shape of the salt dome. The top row of figure 5.1 shows the true model as introduced in the previous chapter. In the lower row the final inversion result after 235 iterations is displayed. The trianlges in the true model mark the receivers used in the seismograms in figures 5.3 and 5.4. The asterisks mark the shots displayed in the same figures. Moreover, the vertical dashed lines show the cross-sections presented in figure 5.5.

First, let us have a look at the final inverted P-wave velocity model. All prominent layers at the left side of the salt dome could clearly be recovered. The two very thin layers (Lower Cretaceous and Roet 2) can be seen only partially. Their thickness of about 100 m is in the order of the maximal resolution and thus this observation is not surprising. The same reason accounts for the poor reconstruction of the very thin salt layer at about 4.5 km depth. Moreover, as there are no significant reflections from this layer there are no strong indications included in the data. The reconstruction of the shape of the salt dome was successful. The high velocity values at the top of the salt dome are caused by strong reflections and scattering effects which correspond to high amplitudes in the wavefield and in the gradients of the misfit functional. The preconditioning is obviously not able to completely reduce these effects. The reconstruction at the right side of the salt dome is comparatively worse. Although most layers are visible the resolution is not as high and the velocity values are less accurate. Also some artifacts are apparent in the uppermost layer and at the overhang of the salt dome. These observations correspond to the degree of seismogram fitting of the modeled data to the pseudo-observed data. The exemplary seismogram at the left side of the salt dome at 4500 m distance (figure 5.3) reveals a high degree of fitting. The exemplary seismogram at the right side of the salt dome at 13600 m distance (figure 5.4) shows are comparatively poor fitting. Main reason for the poor fitting at the right side of the salt dome is that the sedimentary reflections arrive simultaneously with the salt dome reflections and the scattered waves. Hence, the valuable information about the sedimentary layers included in the reflections cannot clearly be projected to the original position in the model.

Next, the reconstruction of the density model is considered. It has to be noted that the initial density model does not include information about the Buntsandstein layer at the right flank of the salt dome as it does in the P-wave velocity model. The reason is the high density contrast between salt and Buntsandstein. Therefore, the Buntsandstein layer vanishes in the smoothing process. The reflectivity of the sedimentary layers could be recovered similarly to the P-wave velocity. In this case even the thin salt layer at the right and left side of the salt dome could be reconstructed. However, the contours of the salt dome are blurred especially in the upper part. Additionally, in the first layer a significant amount of artifacts can be observed. As the pressure field is more sensitive to the P-wave velocity than the density, the density values are sometimes estimated wrongly. For example, at the first layer boundary, at the right flank of the salt dome and at the Buntsandstein layer at the left side of the salt dome the density is overestimated. This observation can be illustrated by looking at the cross-sections in figure 5.5. The reconstructed P-wave velocities stay close to the true model. In some cases the P-wave velocity is underestimated and at the salt dome (distance 10.125 km) the reconstructed velocity oscillates due to artifacts. In contrast, the reconstructed density is generally more unsteady. Many jumps can be observed and it is often overestimated at the boundaries. This corresponds to trade-off effects between the P-wave velocity and the density. In this context it has to be mentioned that the major part of the structural information included in the final density model is possibly caused by crosstalk from the P-wave velocity gradients. To estimate the extent of crosstalk a detailed crosstalk study would be necessary.

At last, in figure 5.2 the evolution of the normalized L_2-misfit is shown. The discrete jumps correspond to changes of the workflow stage. The final relative L_2-misfit at iteration 235 is approximately 8.7% of the initial misfit.

In conclusion, the acoustic FWI succesfully reconstructed the prominent layers of the model and the shape of the salt dome. The very thin layers are only partially visible in the P-wave velocity model and more clearly in the density model. However, the contours of the salt dome are a bit blurred in the density model. The absolute values of the reconstructed P-wave velocity are generally close to the true values whereas the density shows more oscillations and is overestimated sometimes as a result of trade-off effects. A satisfying data fit could be achieved for the shots at the left side of the salt dome, the modeled seismograms for shots at the right side of the salt dome show some deviations from the pseudo-observed data.

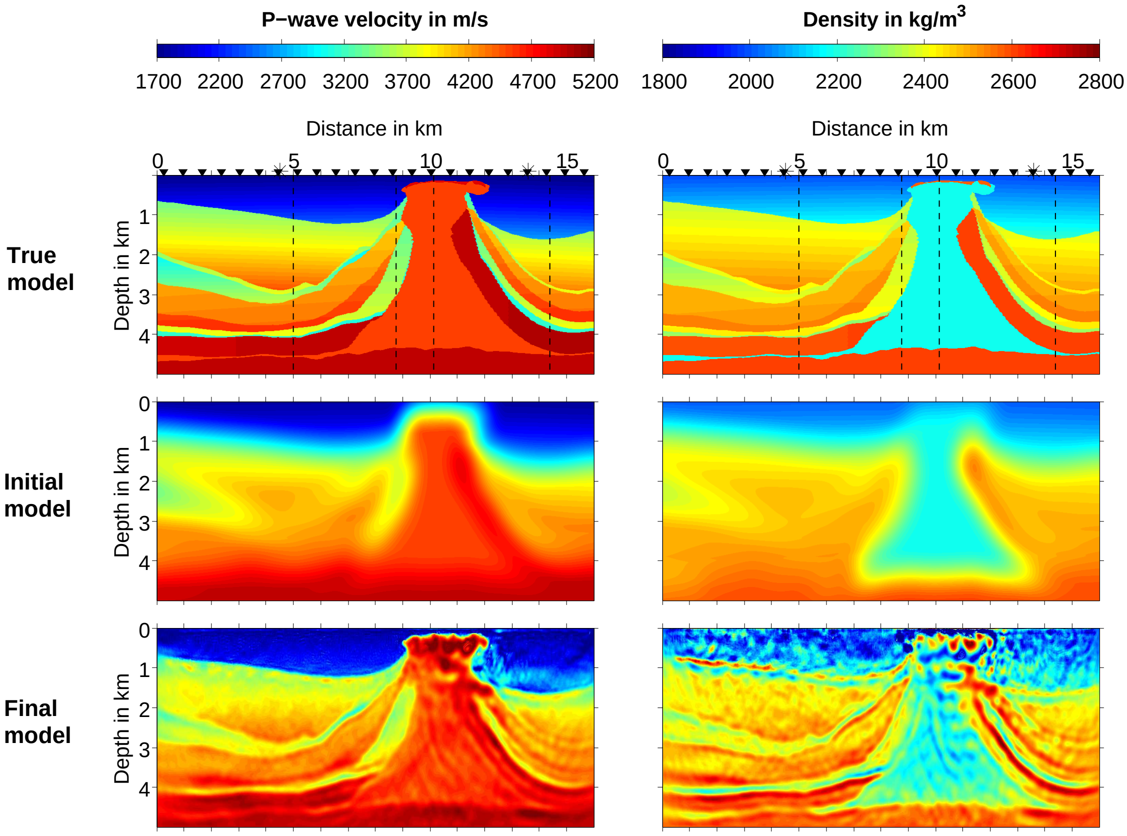


Figure 5.1.: Results of the acoustic FWI. Left column: P-wave velocity model, right column: density model. Top row: true model, middle row: initial model, lower row: final result after 235 iterations. The triangles in the true model correspond to the seismogram traces displayed in figures 5.3 and 5.4. The asterisks mark the shots displayed in the same figures.

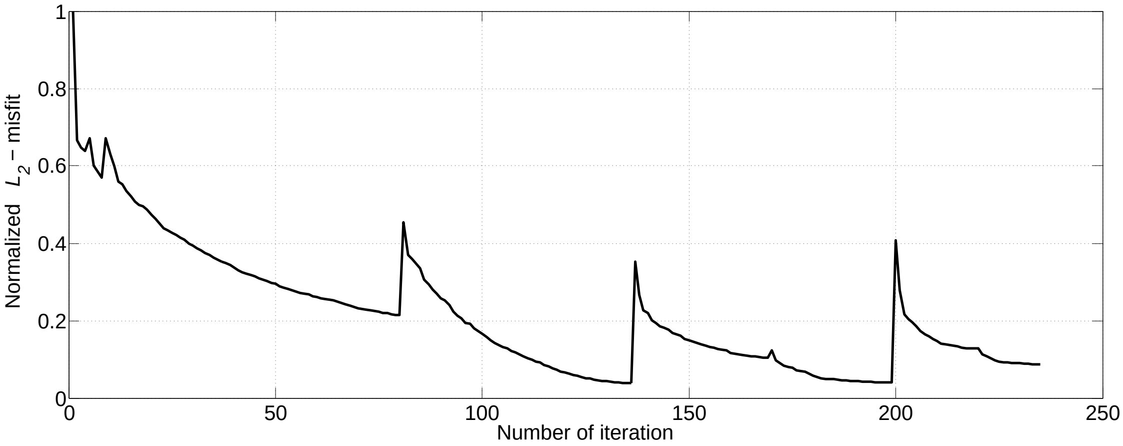


Figure 5.2.: Evolution of the normalized L_2-misfit during the acoustic FWI.

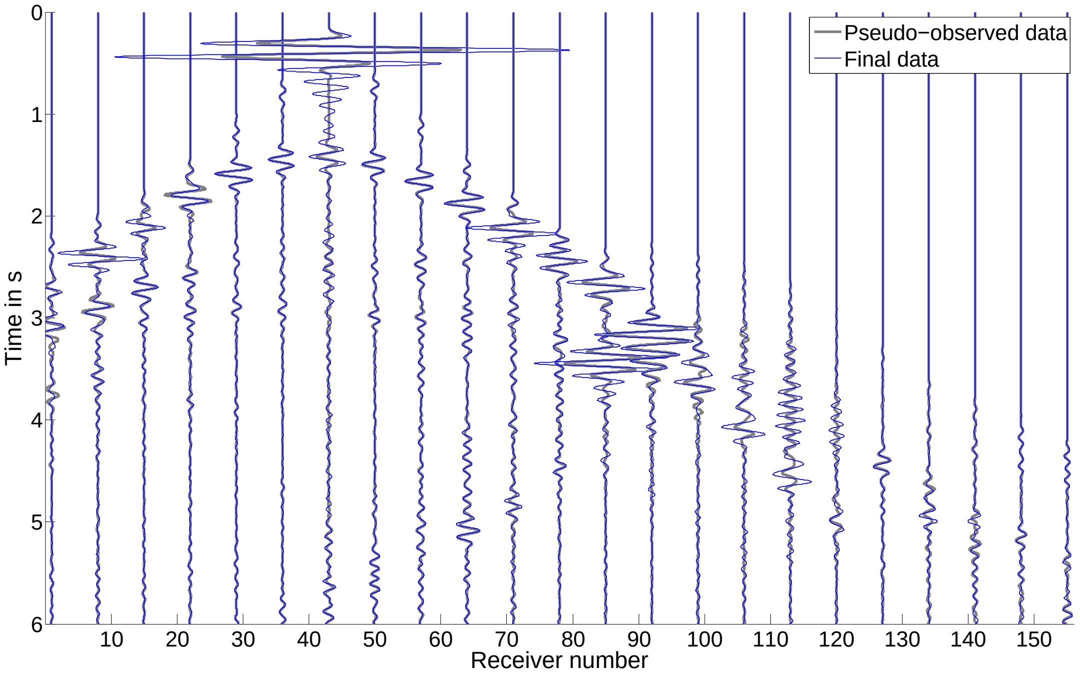


Figure 5.3.: Seismograms of shot 6 at 4500 m distance with the traces marked in figure 5.1. The blue lines correspond to the final inverted model after 235 iterations, the grey lines to the pseudo-observed data.

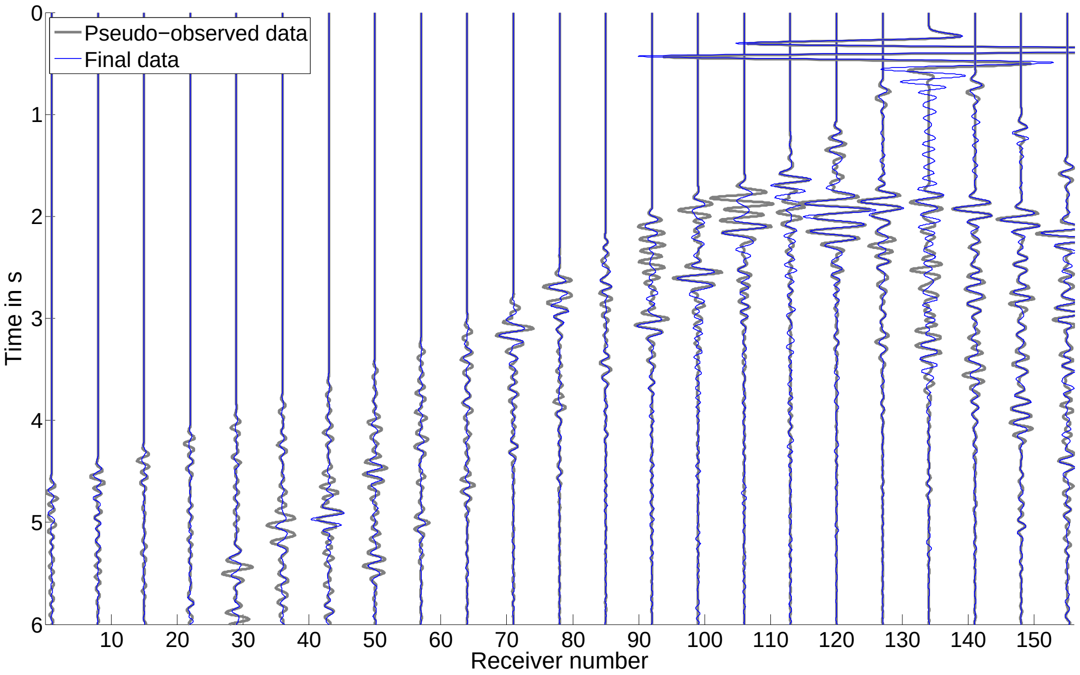


Figure 5.4.: Seismograms of shot 19 at 13600 m distance with the traces marked in figure 5.1. The blue lines correspond to the final inverted model after 235 iterations, the grey lines to the pseudo-observed data.

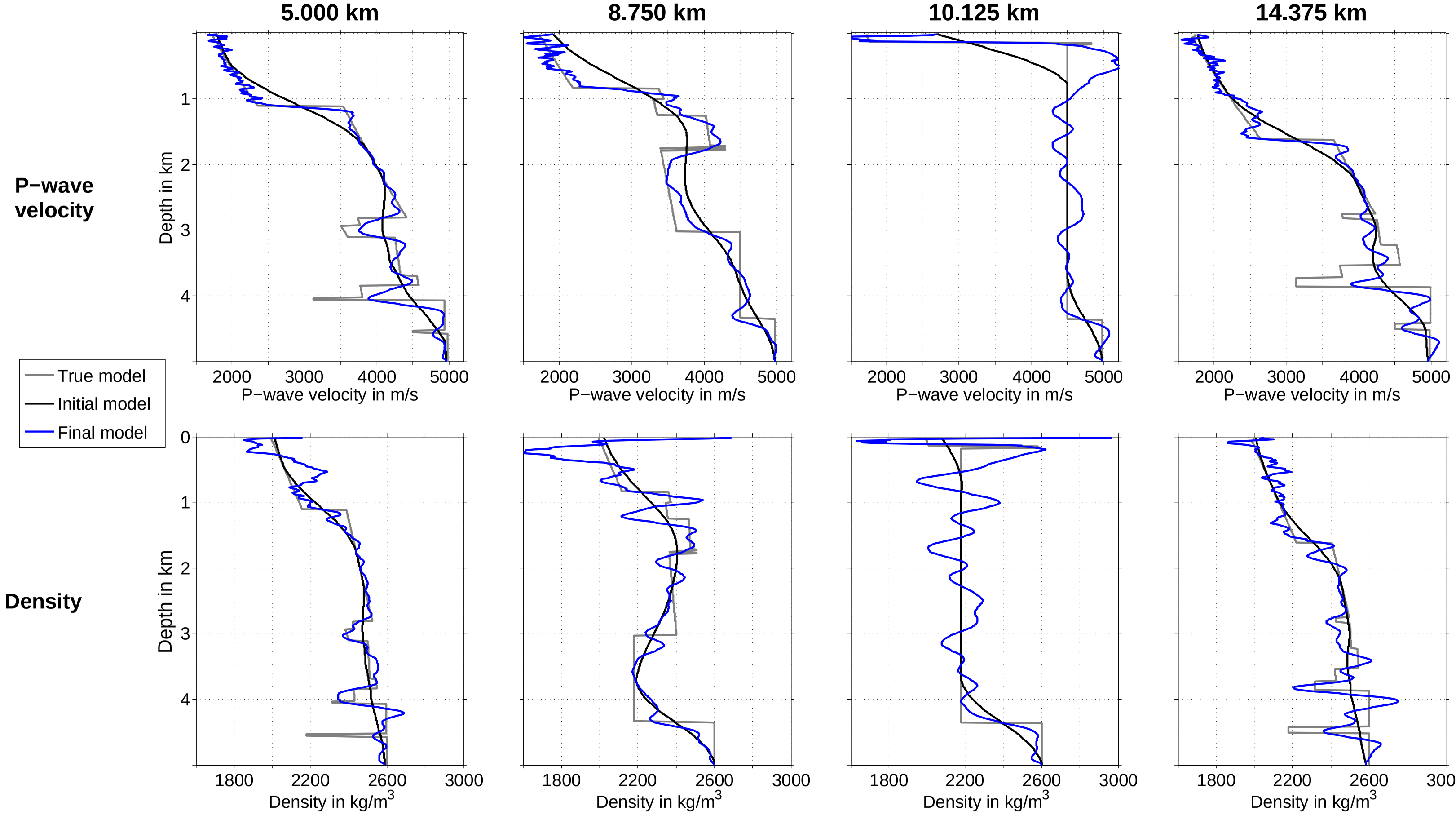


Figure 5.5.: Cross-sections of the P-wave velocity and density models at four different positions (5.0 km, 8.75 km, 10.125 km and 14.375 km) which are marked in figure 5.1. The blue lines correspond to the reconstructed model, the black lines to the initial model and the grey lines to the true model.

5.2. Gravity inversion

Three reconstruction tests were performed with the gravity inversion. In each test a different initial model was used. The results are summarized in figure 5.6. The top row shows the initial model, the middle row the final model and the lower row the gravity field. Each column represents one independent inversion. In the left column the same smoothed initial model was used as in the acoustic FWI. The gravity field of the initial model (black line) is already very close to the pseudo-observed data (grey line) and the final gravity field after 48 iterations (blue line) fits the pseudo-observed data perfectly. The model was updated especially above the salt dome. Some long wavelength updates are there, but hardly visible in the image.
The characteristic long wavelength update of gravity inversion is demonstrated clearly in the next inversion (middle column). Here, an initial model with a linear gradient from 1996 $\mathrm{kg/m^3}$ to 2605 $\mathrm{kg/m^3}$ was used. An update with an approximate wavelength of 10 km is applied over the whole model. At the position of the salt dome updates with smaller wavelengths are also visible. Similar to the first inversion the final model, here after 28 iterations, explains the pseudo-observed data perfectly.
The last inversion is presented in the right column of figure 5.6. In this case a model with a constant density was used initially. The density was set to the mean density of the model with 2360 $\mathrm{kg/m^3}$. Thus, the gravity field of the inital model is zero because the Bouguer plate has the same density as the initial model. The final model after 53 iterations again explains the pseudo-observed data perfectly. Moreover, in the final model the typical elliptical shape of the geometrical kernel can be retrieved.
The perfect data fit of all three inversions is also demonstrated by the evolution of the normalized L_2-misfit (figure 5.7). First, the fast convergence after only a few iterations of all inversions is obvious. This is not surprising because the inverse problem is linear with the chosen parameterization as explained in section 3.2. The final L_2-misfits are: 0.032% (smoothed initial model), 0.002% (gradual initial model) and 0.001% (constant initial model).

Finally, there are two import conclusions regarding the pure gravity inversion. First, the ambiguity is obvious. Even though the final models differ significantly the final gravity field explains the pseudo-observed data perfectly in all three cases. It has to be mentioned that these three models are only a few examples of a generally infinite number of models. Secondly, the model update is restricted to long wavelenghts and thus the gravity inversion is not able to reconstruct the layered structure of the model without additional constraints.

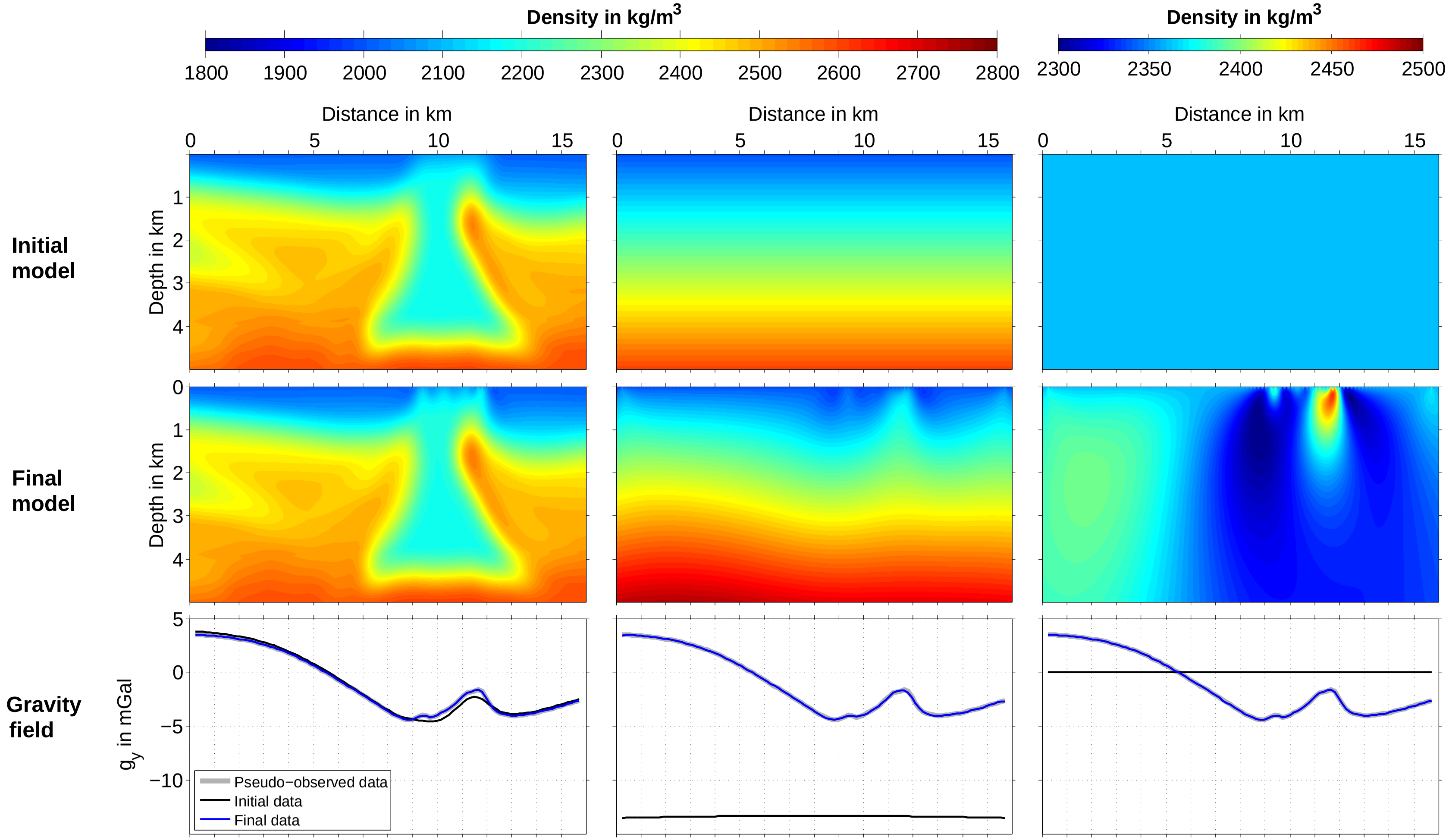


Figure 5.6.: Three gravity inversions with different initial models: smoothed initial model (right column), initial model with a linear gradient (middle column), constant initial model (right column). Top row: initial model, middle row: final model, lower row: gravity field.

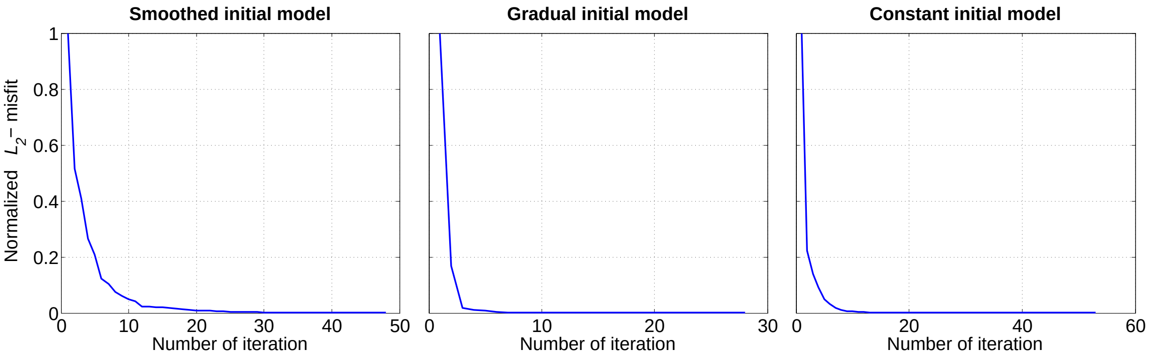


Figure 5.7.: Evolution of the normalized L_2-misfit during the gravity inversions.

5.3. Joint inversion

We performed two different joint inversions with a weighting parameter γ of 0.1 and 0.3. Experiments showed that inversions with parameters smaller than 0.1 disregard gravity so that practically a pure acoustic FWI is executed. Higher values than 0.3 prevent a proper model update so that the joint inversion is terminated quickly.

First, let us have a look at the final P-wave velocity model and the seismograms displayed in figure 5.8. The triangles and asterisks in the true model (top row) denote the traces and shots shown in the seismograms (lower row). The middle row shows the final inverted P-wave velocity model after 159 iterations (left column, $\gamma = 0.1$) and after 62 iterations (right column, $\gamma = 0.3$). Similar to the pure acoustic FWI the prominent layers could be reconstructed. However, the resolution is not as high and artifacts are visible in the uppermost layer and close to the top of the salt dome. The data fit of shot 6 is still satisfying with a weak weight of gravity ($\gamma = 0.1$), but significant residuals are obvious with the stronger coupling ($\gamma = 0.3$). Moreover, it is important to note that the layered structure was already recovered in both inversions after the first two workflow stages where only the P-wave velocity model was updated. Afterwards the joint inversion mainly increased the resolution and decreased the artifacts using only small model updates in each iteration in the range of 0.5 m/s to 10 m/s. During later iterations the average update was less than 1 m/s in both cases.

The results of the density model reconstruction are shown in figure 5.9. Again the top row shows the true model and the middle row the final inverted model. The final gravity field is presented in the lower row. Obviously, both final density models explain the pseudo-observed gravity field almost perfectly and can consequently be regarded as reliable to some extent. In the final density model with the weak coupling the most prominent layers are recovered. However, the layers at the right side of the salt dome are hardly visible. In the final model with the stronger coupling some layers are indicated at the left side of the salt dome. Even though the resolution is decreased compared to the pure acoustic FWI the absolute values are more reliable because trade-off effects could be reduced by incorporating gravity observations. In figure 5.10 this becomes apparent. It shows cross-sections of the density model at the positions marked in figure 5.9. With a higher weight of gravity the density values oscillate less. However, a relative weight of 0.3 seems to be too high because the updates are restricted too much and it might prevent a smooth and fast convergence of the misfit functional. The positive effect of the joint inversion on the reduction of trade-off effects can clearly be demonstrated by looking at the Buntsandstein layer at the left side of the salt dome. In the acoustic FWI the layer boundaries are sharply visible. However, at the upper boundary the density values exceed 2700 kg/m^3 which is more than 100 kg/m^3 too high. At the lower half of the layer the values are mostly about 50 kg/m^3 too low. In contrast, in the joint inversion with $\gamma = 0.3$ the layer is more homogeneous and the density values are overestimated by only 50 kg/m^3.

Finally, the evolution of the L_2 joint misfit is shown in figure 5.11. The top image corresponds to the weighting parameter $\gamma = 0.1$ and the middle image to the weighting parameter $\gamma = 0.3$. In these images the green line represents the normalized joint misfit and the black and the blue line the contribution of the seismic and gravity misfit, respectively. As in the acoustic FWI the jumps mark changes of the workflow stage. Generally the convergence is not as fast as in the pure acoustic and gravity inversion. The reduction of the seismic misfit becomes slower with a higher gravity contribution whereas the reduction of the gravity misfit becomes faster with a higher gravity contribution. Especially noticeable is the misfit evolution from iteration 35 in the case of the high gravity contribution. The gravity contribution raises due to large steepest descent density updates in the first iteration of a new workflow stage and decreases quickly afterwards whereas the seismic misfit can almost not be reduced during every workflow stage (iterations 36-40, 41-45, 46-50, 51-55 and 56-62). In addition to the slower convergence behaviour the misfit reduction is less efficient. With a reduction to 50.5% ($\gamma = 0.1$) and 66.8% ($\gamma = 0.3$) the final seismic misfit is much higher than in the pure acoustic FWI (8.7%). The reduction of the gravity misfit to 0.456% for $\gamma = 0.1$ and 0.254% for $\gamma = 0.3$ (lower image in figure 5.11) is about one order less efficient compared to the pure gravity inversion which reduced the misfit to 0.032%, but still satisfying. The joint misfit could be reduced to 56.1% ($\gamma = 0.1$) and 84.1% ($\gamma = 0.3$).

In conclusion, both joint inversions were able to reconstruct information about the layers and the salt dome in both the P-wave velocity and the density model, however with less resolution. The joint inversion with a relatively small gravity contribution ($\gamma = 0.1$) was able to fit both data sets satisfactorily. In contrast, the joint inversion with a higher gravity contribution ($\gamma = 0.3$) was unable to fit the pressure seismograms significantly whereas the gravity field could also be fitted well. A positive feature of both joint inversions is the increased reliability of the density values. Trade-off effects between P-wave velocity and density could be decreased at least partially. Concerning the inversion process it is especially noticeable that only small updates could be applied and that the convergence speed was decreased compared to both individual inversions. The inverse problem of gravity is linear and thus a fast convergence is possible. In contrast, the acoustic inverse problem is nonlinear and a slower convergence can be expected. Another possible problem could be that only one common step length is used for both the P-wave velocity and the density update. However, different step lengths combined with individual L-BFGS optimizations for each parameter class would lead to increased trade-off effects. Even though the reconstruction of the P-wave velocity model is worse than in the pure acoustic FWI a similar result could be achieved by performing a pure acoustic FWI with the final

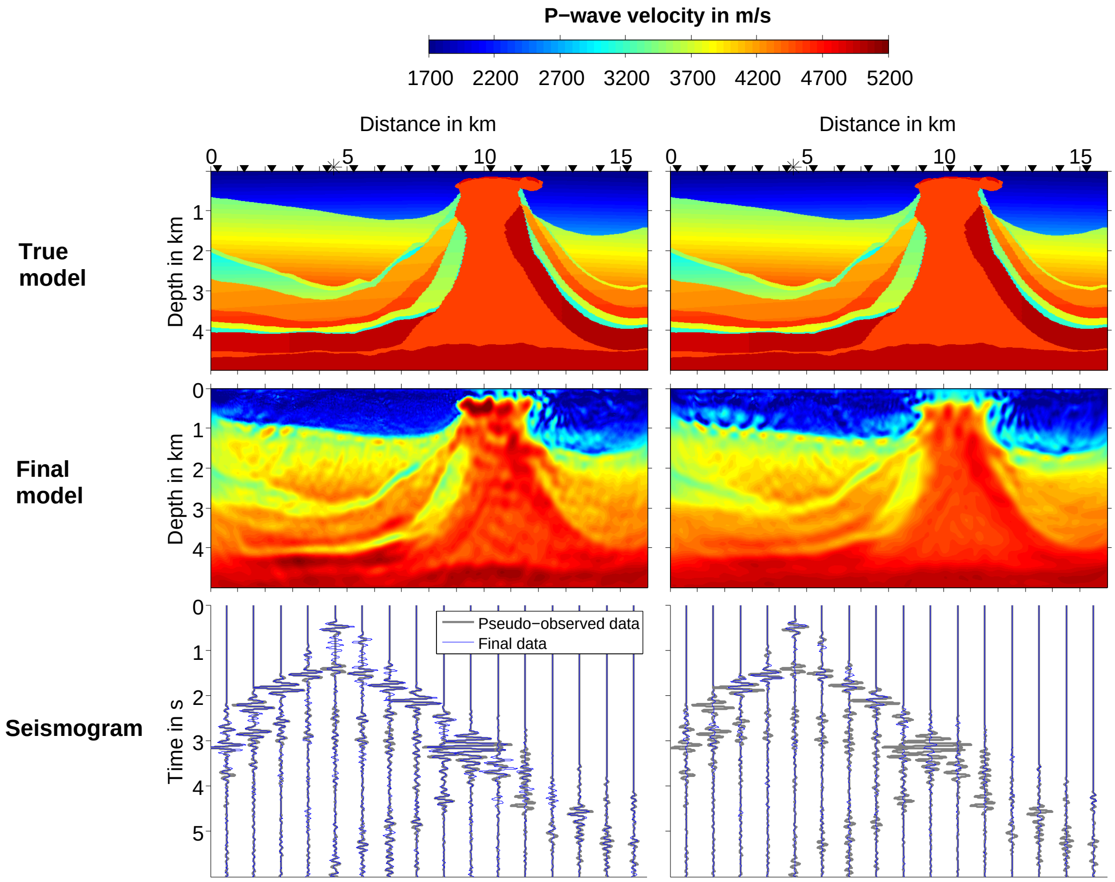


Figure 5.8.: Final P-wave velocity model and seismograms for shot 6 of two joint inversions. Top row: true model, middle row: final model, lower row: seismograms. Left column: $\gamma = 0.1$, right column: $\gamma = 0.3$.

models of the joint inversion as initial model only updating the P-wave velocity model. With this strategy a reliable P-wave velocity and density model could be determined.

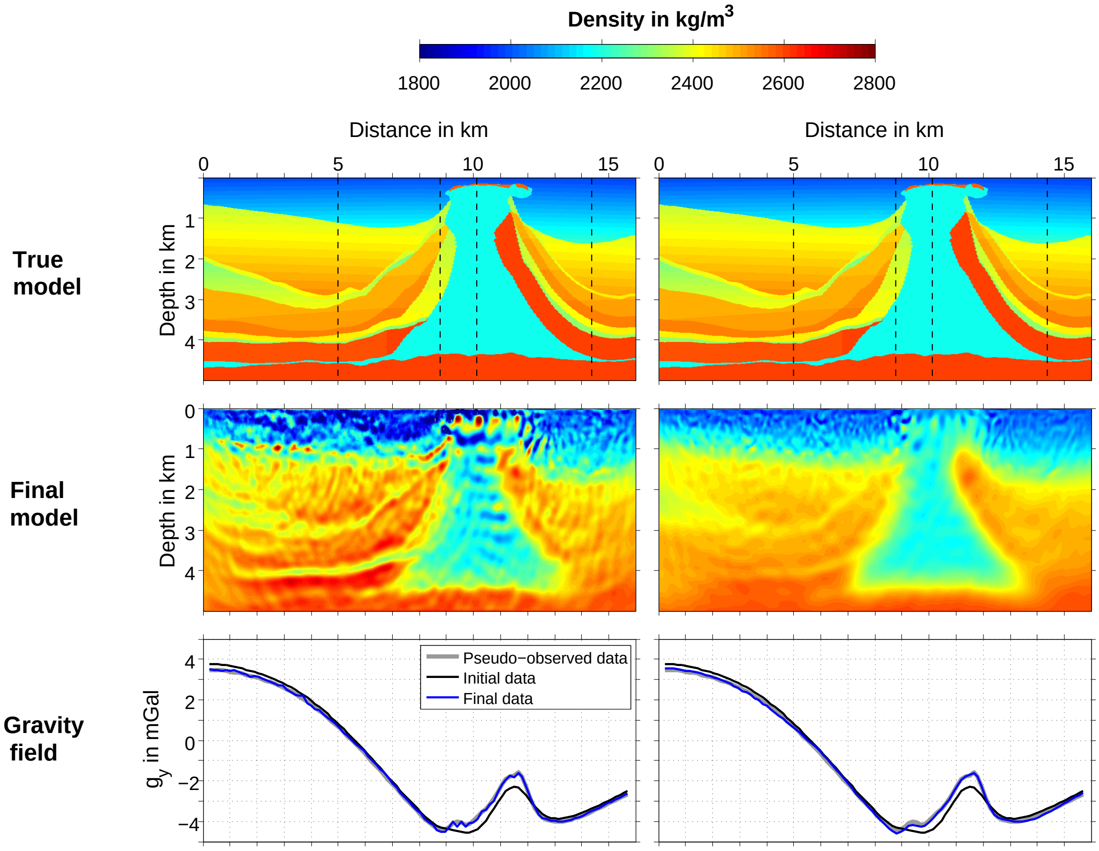


Figure 5.9.: Final density model and gravity field of two joint inversions. Top row: true model, middle row: final model, lower row: gravity field. Left column: $\gamma = 0.1$, right column: $\gamma = 0.3$.

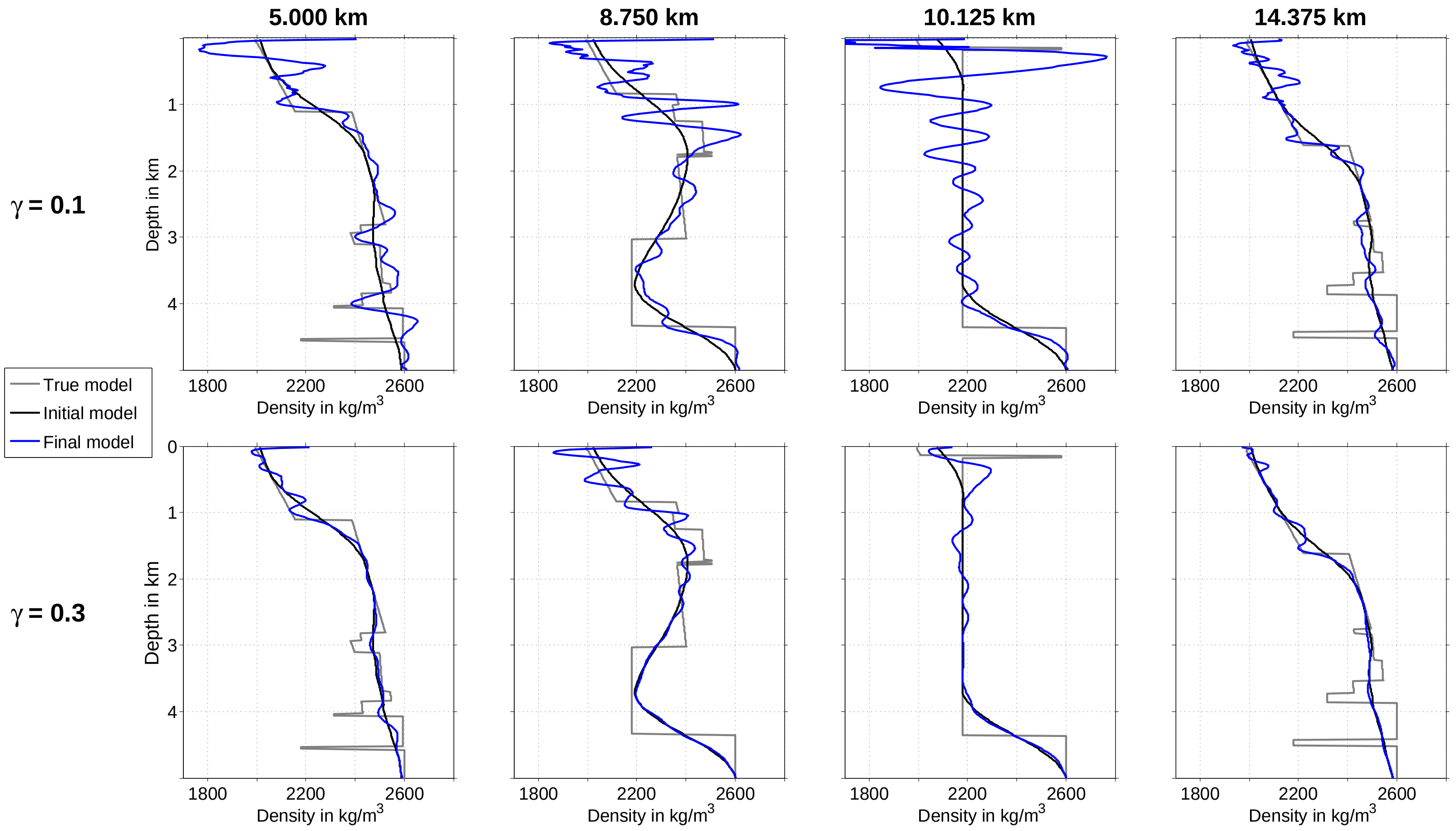


Figure 5.10.: Cross-sections of the density models at four different positions (5.0 km, 8.75 km, 10.125 km and 14.375 km) which are marked in figure 5.9. The blue lines correspond to the reconstructed model, the black lines to the initial model and the grey lines to the true model.

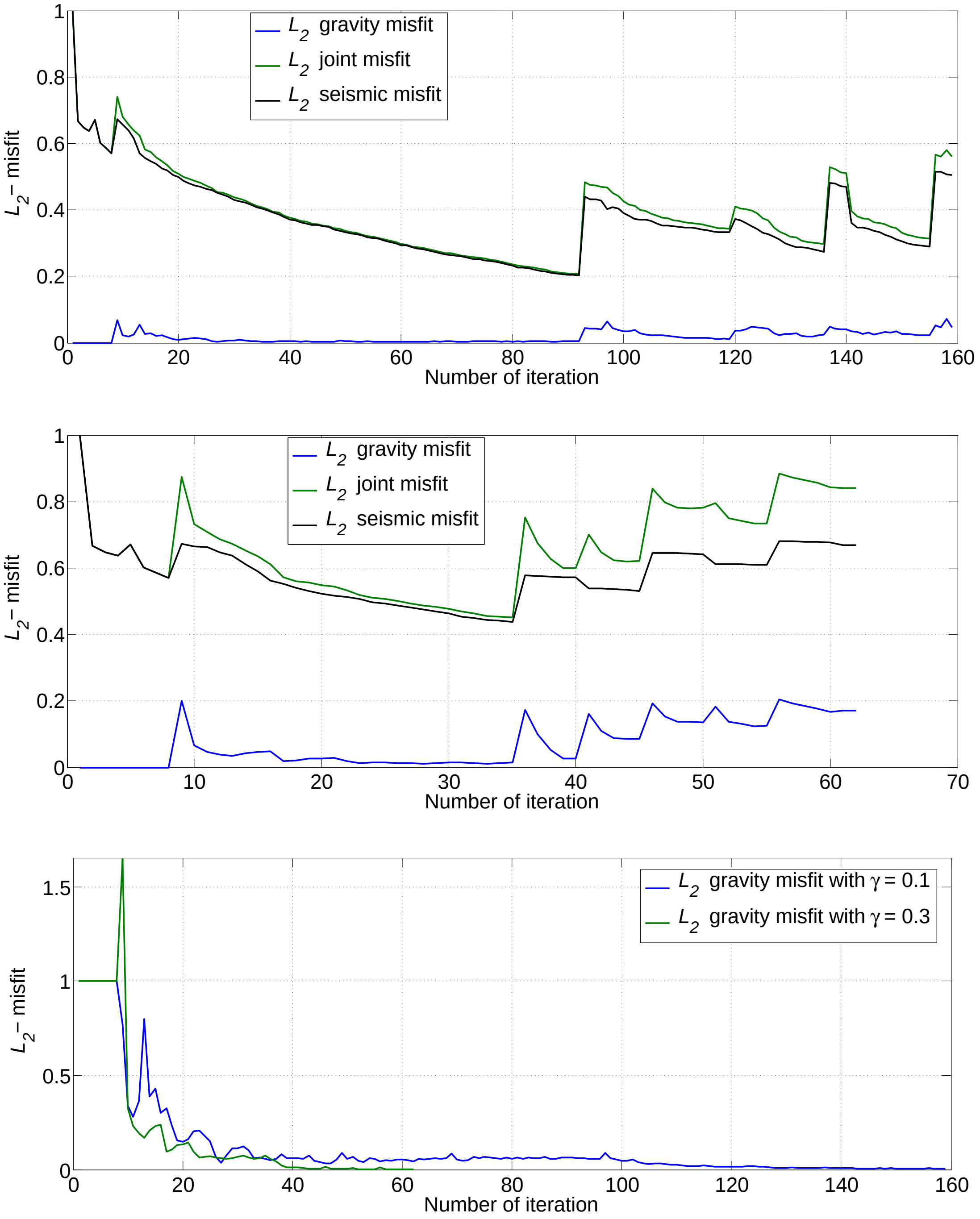


Figure 5.11.: Top: evolution of the normalized joint misfit functional ($\gamma = 0.1$). Middle: evolution of the normalized joint misfit functional ($\gamma = 0.3$). In these two images the green line represents the joint misfit functional, the black line the contribution of the seismic misfit functional to the joint misfit functional and the blue line the contribution of the gravity misfit functional. Bottom: gravity misfit functional normalized with respect to the initial misfit after reduction of the weighting parameter λ_1. In this image the blue line corresponds to $\gamma = 0.1$ and the green line to $\gamma = 0.3$.

5.4. Comparison

At last, the final density models of the pure acoustic FWI, the gravity inversion and the joint inversion are compared. Although the joint inversion results with a low contribution of gravity ($\gamma = 0.1$) show a better data fit and model reconstruction we use the results with the higher weighting ($\gamma = 0.3$) for the comparison. In this case the effect of the joint inversion can be identified more clearly. The final density models and the final gravity fields are displayed in figure 5.12.

The final density model of the joint inversion reveals a better reconstruction than achieved by the gravity inversion. The layered structure of the model is indicated at the left side of the salt dome whereas no signs of layers are visible in the result of the gravity inversion. At the same time the gravity field can be fitted almost identically well. In short the goal of an improved resolution and restriction of the ambiguity of the inverse problem of gravity could be achieved.

The final density model of the acoustic FWI shows a very good resolution. All layers can be identified and the contours of the salt dome are sharpened compared to the initial model. However, as a result of trade-off effects between P-wave velocity and density the density values are overestimated sometimes. Obviously the final density model does not explain the pseudo-observed gravity field which poses doubts about its reliability. The acoustic FWI even worsened the data fit compared to the initial model. In comparison, the joint inversion has less resolution but a higher reliability because the data is fitted almost perfectly. The attenuating effect of the joint inversion on the density update is clearly illustrated in figure 5.13 which shows cross-sections of the density model as marked in figure 5.12. The blue line corresponds to the acoustic FWI and the red line to the joint inversion. The blue line oscillates a lot and overcompensates the deviation from the initial model (black line). In contrast, the red line stays close to the initial model, but often fails to reconstruct the true model (grey line) satisfactorily. Yet, the tendency of the model update at the layer boundaries is correct in most cases.

From these results we conclude that a joint inversion can in fact provide benefits to the density reconstruction of the subsurface. A joint acoustic FWI and gravity inversion uses on the one hand the high resolution obtained by seismic wave propagation and on the other hand the potential of long wavelength gravity effects to reduce trade-off effects between density and P-wave velocity. The long wavelength information mitigates or damps the overcompensation of density values during the acoustic FWI.

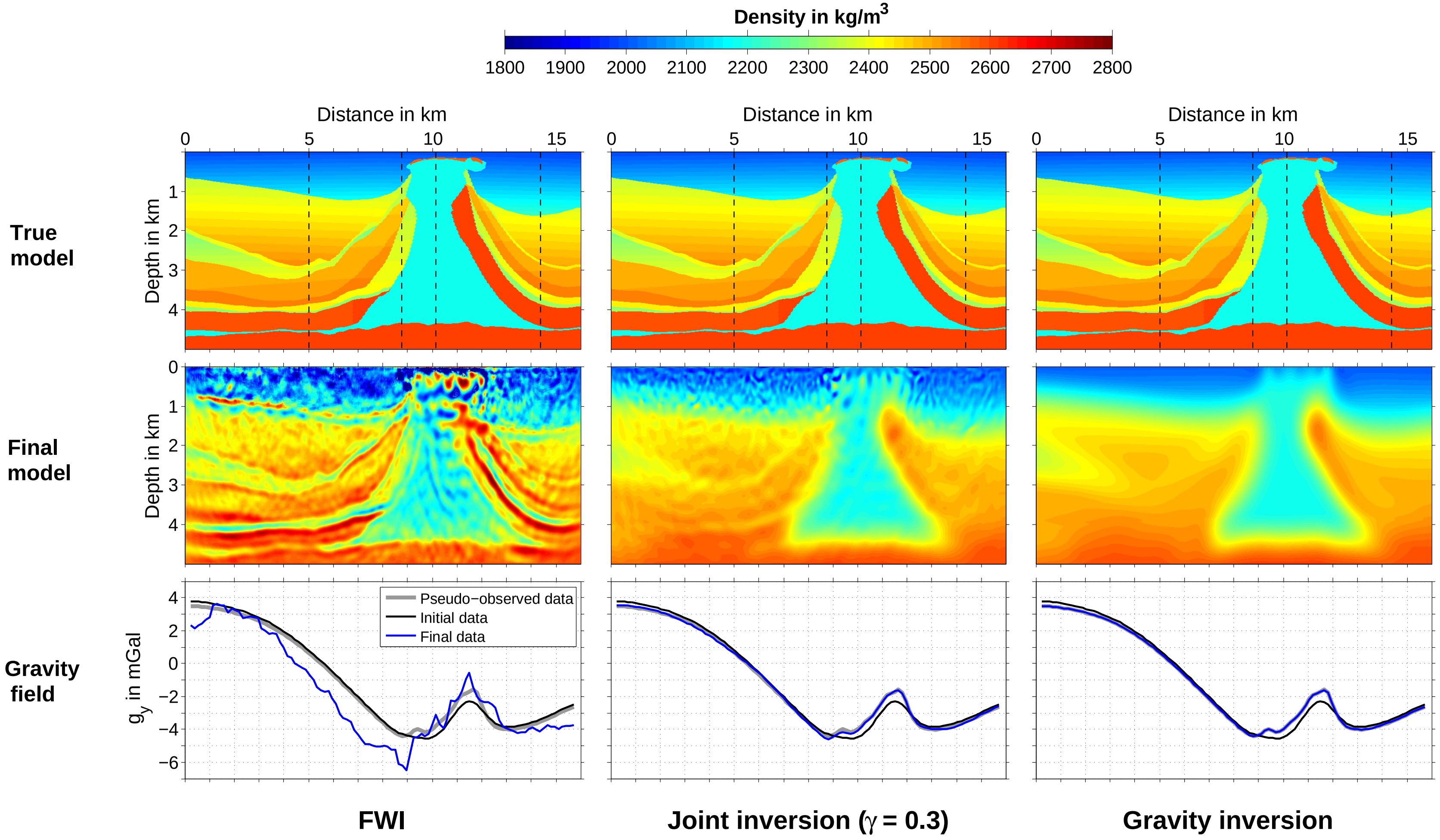


Figure 5.12.: Comparison of the final inverted density model and gravity field between the acoustic FWI, the gravity inversion and the joint inversion ($\gamma = 0.3$).

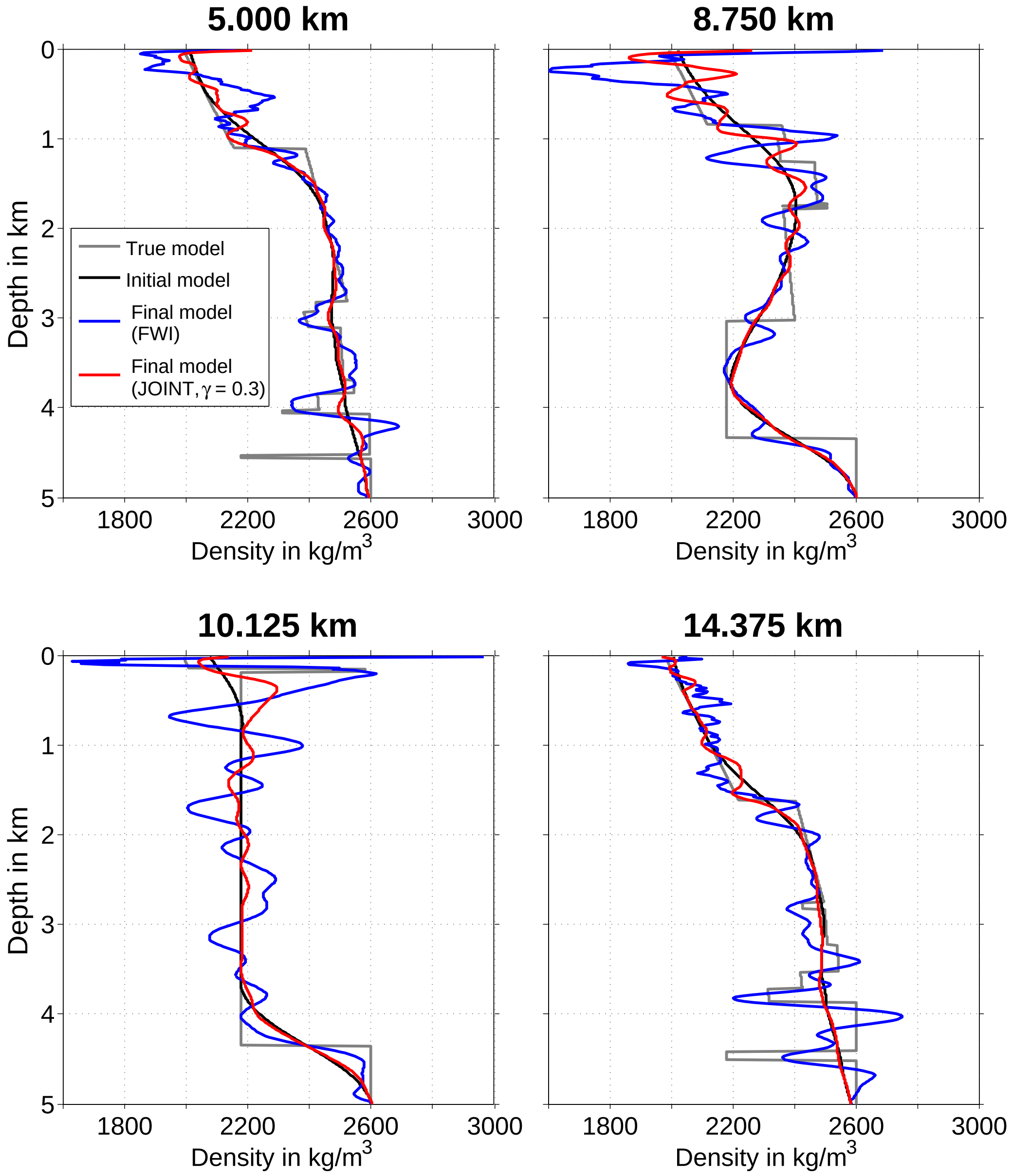


Figure 5.13.: Comparison of the cross-sections of the density model marked in figure 5.12 between the acoustic FWI and the joint inversion ($\gamma = 0.3$).

6. Conclusion and outlook

6.1. Conclusion

In this thesis, we performed a joint acoustic full-waveform and gravity inversion to improve density model building. While the seismic FWI is able to build reliable velocity models it is still very challenging for the density due to trade-off effects and a limited sensitivity of seismic waves to density variations. In contrast, gravity is sensitive to overall density variations without suffering from attenuation effects. However, the latent ambiguity and limited resolution is a major drawback of the gravimetric inverse problem.

For this purpose we implemented a gravity forward modeling, gravity inversion and joint inversion code into our existing full-waveform inversion code. The gravity forward modeling is solved by introducing rectangular prisms to the seismic finite-difference grid. The gravity effect of each prism is calculated by Newton's law of universal gravitation. Based on an initial guess the velocity and density model is updated in order to minimize the pressure seismogram residuals and the gravity field residuals simultaneously. The model update is calculated by preconditioned gradients of the joint misfit functional and the L-BFGS quasi-Newton method in combination with a Wolfe line search algorithm. Three different synthetic reconstruction tests were performed with a salt dome structure which is embedded in sedimentary layers.

The pure acoustic FWI successfully reconstructed the prominent sedimentary layers and the shape of the salt dome in both the P-wave velocity and the density model. Very thin layers are visible partially and are better resolved in the final density model. While the P-wave velocity values are generally close to the true values, the density values are often overestimated at the boundary layers. Trade-off effects can clearly be identified.

Tests with the gravity inversion applied to three different initial models emphasize the latent ambiguity. Although the final models differ significantly the final data fit is almost perfect in all cases. Moreover, the resolution is limited to very long wavelength. Hence, there are no signs of sedimentary layers in the final models.

Two joint inversions with a different relative weight between the acoustic full-waveform and the gravity inversion were applied. Both showed that the drawbacks of each individual inversion can be overcome to some extent. In comparison with the gravity inversion the joint inversion could resolve structures which were not included in the gravity inversion results while the pseudo-observed gravitational field could be explained similarly well. Compared to the acoustic FWI the joint inversion lost resolution, yet trade-off effects could be reduced. Moreover, the final density model of the acoustic FWI is unable to fit the pseudo-observed gravitational field which poses doubts about its reliability.

In conclusion, the joint inversion was able to reduce trade-off effects and improved the reliability of the density model reconstruction. The joint inversion used the high resolution of the FWI while the gravity inversion ensured that the density values stayed within a reasonable range. Although the resolution and the data fit of the seismograms were reduced compared to the pure FWI the potential of the joint inversion to allow a more confidential interpretation of density distributions could clearly be demonstrated.

6.2. Outlook

We achieved promising results with our joint acoustic full-waveform and gravity inversion. However, these results have to be verified by further synthetic and also field data studies. Especially within field data interpretation the application of FWI to derive density models is challenging because usually the field data and the forward modeled data have to be normalized to be comparable and the information about the density is included in the seismic amplitudes. Moreover, attenuation, anisotropy and noise complicate the interpretation of seismic amplitudes. In synthetic studies the true model can be designed arbitrarily and the model can be assumed to continue in the third coordinate direction identical to the 2D plane. In contrast, in field surveys this assumption is in general not justified. Thus, the gravity code should be extended to 3D for this purpose.

As mentioned before the joint inversion uses the high resolution of the FWI to introduce layers to the density model whereas the gravity inversion mitigates the trade-off effects which results in more reliable density values. Yet, it is not clear to which extent the structures in the density model are caused by crosstalk from the P-wave velocity gradients or by an actual sensitivity to density variations. Therefore, a sensitivity analysis of the pressure seismograms to density variations based on Fréchet derivatives could give interesting insights.

The main drawback of gravity inversion which could be mitigated in this study is the latent ambiguity. In addition to seismic data fitting further constraints could be used. In this context for example the following two strategies are possible: Tikhonov regularization and change of the gravimetric modeling parameterization. Tikhonov regularization is applied by introducing a least squares model norm to the data misfit functional which can be weighted if necessary. For gravity a depth weighted model norm or a minimum support stabilizing functional could be promising (Zhdanov, 2002). The depth weighting would lead to a model which is less focussed to the near surface and the minimum support stabilizing functional would enhance the compactness of the objects. However, the necessity of additional weighting parameters complicates the application of regularization during the joint inversion. A more promising approach to further reduce the ambiguity is to change the gravimetric modeling parameterization. In our case, we used 512,000 model parameters, i.e. prisms, and only 104 observations. If the subsurfaces would be parameterized using homogeneous bodies with constant densities instead of a fixed geometry with variable densities the number of model parameters could be reduced significantly. The inversion would then shift the bodies or change its shape. Moreover, in this case homogeneous bodies are achieved which is usually more appropriate in most geological situations.

Bibliography

Blakely, Richard J. (1995). *Potential theory in gravity and magnetic applications.*

Bohlen, Thomas (1998). "Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme". PhD thesis. Universität Kiel.

Boulanger, Olivier and Michel Chouteau (2001). "Constraints in 3D gravity inversion". In: *Geophysical Prospecting* 49.2, pp. 265–280.

Brossier, Romain (2011). "Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance". In: *Computers & Geosciences* 37.4, pp. 444–455.

Chapman, Chris (2004). *Fundamentals of Seismic Wave Propagation.* Cambridge Books Online. Cambridge University Press.

Colombo, D. and M. De Stefano (2007). "Geophysical modeling via simultaneous joint inversion of seismic, gravity, and electromagnetic data: Application to prestack depth imaging". In: *The Leading Edge* 26.3, pp. 326–331.

Courant, R., K. Friedrichs, and H. Lewy (1928). "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Mathematische Annalen* 100.1, pp. 32–74.

Ehrenfried, K. (2004). *Strömungsakustik: Skript zur Vorlesung*. Berliner Hochschulskripte. Mensch-und-Buch-Verlag.

Gallardo, Luis A. and Max A. Meju (2004). "Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints". In: *Journal of Geophysical Research: Solid Earth* 109.B3.

Gardner, G.H.F., L.W. Gardner, and A.R. Gregory (1974). "Formation velocity and density-the diagnostic basics for stratigraphic traps". In: *Geophysics* 39.6, pp. 770–780.

Hadamard, Jacques (1902). "Sur les problèmes aux dérivés partielles et leur signification physique". In: *Princeton University Bulletin* 13, pp. 49–52.

Heincke, Bjoern, Marion Jegen, and Richard Hobbs (2006). "Joint inversion of MT, gravity and seismic data applied to sub-basalt imaging". In: *2006 SEG Annual Meeting*. Society of Exploration Geophysicists.

Heincke, Björn, Marion Jegen, Max Moorkamp, Jin Chen, and Richard W. Hobbs (2010). "Adaptive coupling strategy for simultaneous joint inversions that use petrophysical information as constraints". In: *2010 SEG Annual Meeting*. Society of Exploration Geophysicists.

Jacoby, Wolfgang and Peter L. Smilde (2009). *Gravity Interpretation: Fundamentals and Application of Gravity Inversion and Geological Interpretation.* Earth and Environmental Science. Springer Berlin Heidelberg.

Jeong, Woodon, Ho-Yong Lee, and Dong-Joo Min (2012). “Full waveform inversion strategy for density in the frequency domain”. In: *Geophysical Journal International* 188.3, pp. 1221–1242.

Kearey, Philip, Michael Brooks, and Ian Hill (2013). *An introduction to geophysical exploration.* John Wiley & Sons.

Kellogg, O.D. (1929). *Foundations of Potential Theory.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Dover Publications.

Köhn, Daniel (2011). “Time Domain 2D Elastic Full Waveform Tomography”. PhD thesis. Kiel University.

Komatitsch, Dimitri and Roland Martin (2007). “An unsplit convolutional Perfectly Matched Layer improved at grazing incidence for the seismic wave equation”. In: 72.5, SM155–SM167.

Lailly, Patrick (1983). *The seismic inverse problem as a sequence of before stack migrations.*

Last, B. J. and K. Kubik (1983). “Compact gravity inversion”. In: *Geophysics* 48.6, pp. 713–721.

Levander, Alan R. (1988). “Fourth-order finite-difference P-SV seismograms”. In: *Geophysics* 53.11, pp. 1425–1436.

Li, Yaoguo and Douglas W. Oldenburg (1998). “3-D inversion of gravity data”. In: *Geophysics* 63.1, pp. 109–119.

Métivier, Ludovic (2014). *SEISCOPE OPTIMIZATION TOOLBOX MANUAL.* SEISCOPE consortium.

Mora, Peter (1987). “Nonlinear two-dimensional elastic inversion of multioffset seismic data”. In: *Geophysics* 52.9, pp. 1211–1228.

Nagy, Dezso, Gabor Papp, and Judit Benedek (2000). “The gravitational potential and its derivatives for the prism”. In: *Journal of Geodesy* 74.7-8, pp. 552–560.

Nocedal, Jorge and Stephen Wright (2006). *Numerical optimization.* Springer Science & Business Media.

Plessix, R.-E. (2006). “A review of the adjoint-state method for computing the gradient of a functional with geophysical applications”. In: *Geophysical Journal International* 167.2, pp. 495–503.

Plessix, R.-E. and W. A. Mulder (2004). “Frequency-domain finite-difference amplitude-preserving migration”. In: *Geophysical Journal International* 157.3, pp. 975–987.

Przebindowska, Anna (2013). “Acoustic full waveform inversion of marine reflection seismic data”. PhD thesis. KIT-Bibliothek.

Przebindowska, Anna, André Kurzmann, Daniel Köhn, and Thomas Bohlen (2012). “The role of density in acoustic full waveform inversion of marine reflection seismics”. In: *74th EAGE Conference and Exhibition incorporating EUROPEC 2012*.

Tarantola, Albert (1984). “Inversion of seismic reflection data in the acoustic approximation”. In: *Geophysics* 49.8, pp. 1259–1266.

– (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

Virieux, Jean (1986). “P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method”. In: *Geophysics* 51.4, pp. 889–901.

Virieux, Jean and Stéphane Operto (2009). “An overview of full-waveform inversion in exploration geophysics”. In: *Geophysics* 74.6, WCC1–WCC26.

Wehner, Daniel (2015). “A Combined Elastic Waveform and Gravity Inversion Applied to the Marmousi-II Model”. MA thesis. Kiel University.

Wittkamp, Florian (2016). “Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves”. MA thesis. Karlsruhe Institute of Technology.

Xu, Kun and George A. McMechan (2014). “2D frequency-domain elastic full-waveform inversion using time-domain modeling and a multistep-length gradient approach”. In: *Geophysics* 79.2, R41–R53.

Zhdanov, M.S. (2002). *Geophysical Inverse Theory and Regularization Problems*. Methods in Geochemistry and Geophysics. Elsevier Science.

A. Appendix

A.1. Computational resources

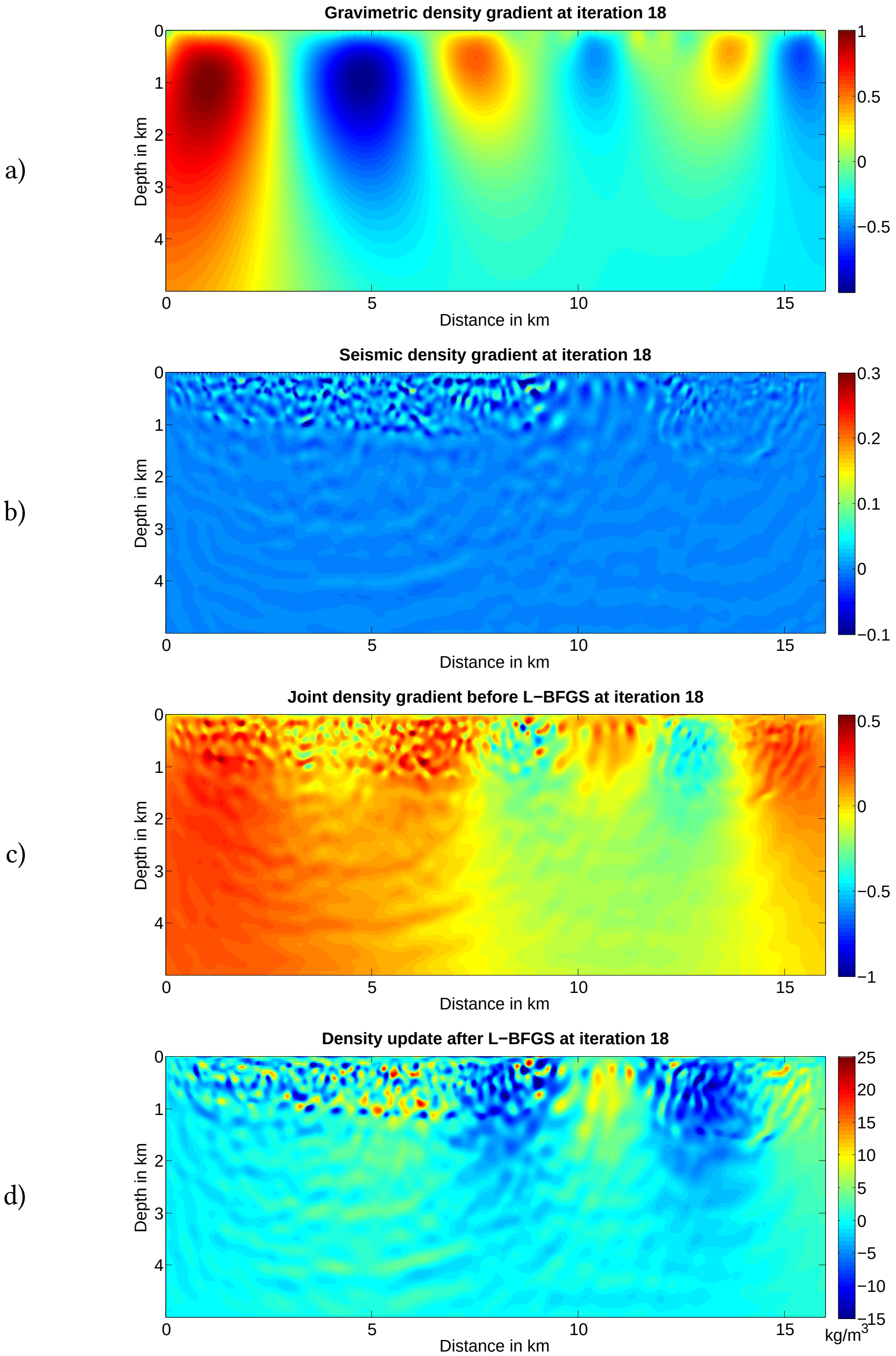
All inversions were calculated on the InstitutsCluster II (IC2) which is maintained by the Steinbuch Centre for Computing (SCC). The IC2 was funded by the Deutsche Forschungsgemeinschaft (DFG; English: German Research Foundation) and established by different institutes of the Karlsruhe Institute of Technology (KIT). Table A.1 shows a list of the used core numbers, the calculation time and the number of iterations of the different inversions shown in this thesis. As C compiler the Intel C Compiler 16.0 was used and for the parallelization the Intel Message Passing Interface (MPI) Compiler 5.1.

A.2. Exemplary gradients and model update

To get a better idea of how the joint approach works we show some exemplary density gradients and a density model update in figure A.1. From the top to the bottom the normalized gravimetric density gradient, the normalized seismic density gradient, the normalized joint density gradient before L-BFGS and the unnormalized model update after L-BFGS are displayed. All images correspond to the joint inversion with $\gamma = 0.3$ and iteration 18. The top three images are shown with a clipped colorbar for demonstrative reasons. Special attention should be paid to the two layers visible between 2 km and 3 km depth and at about 4 km depth in the lower three images. Moreover, the superposition of long wavelength and short wavelength information in the joint gradient and the model update is notable.

Table A.1.: Computational aspects of the inversions shown in this thesis.

Inversion	Number of cores	Calculation time	Iterations
Acoustic	128	3.87 h	235
Joint ($\gamma = 0.1$)	128	3.32 h	159
Joint ($\gamma = 0.3$)	128	1.72 h	62
Gravity (initial model: constant)	64	0.17 h	53
Gravity (initial model: smoothed)	64	0.15 h	48
Gravity (initial model: gradient)	64	0.10 h	28



a) Gravimetric density gradient at iteration 18

b) Seismic density gradient at iteration 18

c) Joint density gradient before L−BFGS at iteration 18

d) Density update after L−BFGS at iteration 18

Figure A.1.: Exemplary gradients and model update of the joint inversion with $\gamma = 0.3$ at iteration 18. The gradients have clipped colorbars for demonstrative reasons while the model update colorbar shows the whole extent. a) Normalized gravimetric density gradient, b) normalized seismic density gradient, c) normalized joint density gradient before L-BFGS and d) unnormalized model update after L-BFGS.

A.3. Acknowledgements/Danksagungen

Abschließend möchte ich all jenen danken, die mir beim Anfertigen dieser Arbeit geholfen haben.

Als Erstes möchte ich mich bei Thomas Bohlen für die Themenfindung und Betreuung der Masterarbeit bedanken. Die Diskussionen und Ratschläge haben immer den Blick für die wesentlichen Aspekte geschärft, wenn es Probleme oder Unklarheiten gab. Außerdem danke ich Joachim Ritter für die Übernahme des Korreferates.

Der gesamten Arbeitsgruppe "Angewandte Geophysik" danke ich für die angenehme Arbeitsatmosphäre, für die interessanten und anregenden Diskussionen und die ein oder andere Ablenkung, die von Zeit zu Zeit wirklich gut getan hat. Besonderer Dank gilt Laura, Niklas und Florian für das gewissenhafte und geduldige Korrekturlesen der Arbeit. Außerdem möchte ich Claudia Payne und Petra Knopf für die organisatorische Betreuung danken.

Des Weiteren möchte ich den Mitarbeitern von TERRASYS Geophysics, Ina Müller, Venke Becker, Markus Krieger und Peter Smilde, danken. Danke für das Bereitstellen des Dichtemodelles, die gute Zusammenarbeit während des vergangenen Jahres und die vielen Ratschläge insbesondere in Bezug auf Gravimetriemodellierung und -inversion.

Ganz besonders möchte ich mich bei meiner Familie für die fortwährende und bedingungslose Unterstützung während und natürlich auch vor meiner Studienzeit herzlich bedanken. Vielen Dank!