RICE UNIVERSITY

Approximate Multi-Parameter Inverse Scattering Using Pseudodifferential Scaling

by

Rami Nammour

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Master of Arts

Approved, Thesis Committee:

Dr. William Symes, Chairman
Professor of Computational and Applied Mathematics

Dr. Liliana Borcea
Professor of Computational and Applied Mathematics

Dr. Alan Levander
Professor of Geology and Geophysics

Dr. Mark Embree
Professor of Computational and Applied Mathematics

Houston, Texas

November, 2010

Abstract

Approximate Multi-Parameter Inverse Scattering Using Pseudodifferential Scaling

by

Rami Nammour

I propose a computationally efficient method to approximate the inverse of the normal operator arising in the multi-parameter linearized inverse problem in reflection seismology.

Solving the inverse problem using direct matrix methods like Gaussian elimination is computationally infeasible. In fact, the application of the normal operator requires solving large scale PDE problems. However, under certain conditions, the normal operator is a matrix of pseudodifferential operators. This proposal shows how to generalize Cramer's rule for matrices to approximate the inverse of a matrix of pseudodifferential operators. The cost of approximating the inverse is a few applications of the normal operator (one for one parameter, two for two parameters, six for three parameters).

I validate the method on the Marmousi model for constant density acoustics for

the one-parameter problem. For the two parameter problem, the inversion of a variable density acoustics layered model: details the various steps of the method, and corroborates its success.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this manuscript, I review a method to approximately solve the linearized inverse problem of constant density acoustics, work I have accomplished as part of my masters thesis. I propose a generalization for variable density acoustics and other multi-parameter inverse problems. Finally, I propose how the method to solve the linearized inverse problem accelerates the convergence of iterative methods aimed at solving the full nonlinear inverse problem (referred to as full waveform inversion methods). Previous work was limited to 2D (two spatial dimensions); however, I also propose extensions to 3D and discuss the similarities between the two approaches and the challenges specific to 3D.

The model problem for this work is the variable density acoustic wave equation, the simplest model describing the reaction of the earth to acoustic excitation (explosions,

air gun ...).

$$\frac{1}{\rho(x)c^2(x)}\frac{\partial^2 p}{\partial t^2}(x,t) - \nabla \cdot \frac{1}{\rho(x)}\nabla p(x,t) = f(x,t), \tag{1.1}$$

where $\rho(x)$ is the density field, $c(x)$ the velocity field, and $p(x,t)$ the pressure field varying as a function of time; $f(x,t)$ represents the source of acoustic energy. Note that the formulation is the same in 2D and 3D.

Assume the earth was at equilibrium before the forcing is put to effect (causal source), the pressure field satisfies:

$$\begin{aligned} p(x,t) &\equiv 0, \quad t \ll 0 \\ f(x,t) &\equiv 0, \quad t \ll 0. \end{aligned} \tag{1.2}$$

The physical setting of the experiment will invariably involve some boundary conditions, at the sea surface for example in the case of a marine geophysical experiment.

It is advantageous to think of equation (1.1) as defining a map that associations the earth properties (density and velocity fields), to the measurements of the pressure at the surface. Group density and velocity parameters together to form the parameter $m = [\rho, c]$. The map I have in mind is:

$$S[m] = p|_{surface}. \tag{1.3}$$

I refer to S as the nonlinear forward map, it maps the model to the measurement of the pressure at the surface $p|_{surface}$. The pressure at the surface is measured using a recording device, like geophones.

The advantage of this abstraction is that it describes any model of the earth in the same way. The generalization to elasticity, for example. falls under the same rubric,

with m standing for the elastic coefficients and density. The methods proposed in this thesis often extend to models of seismic wave propagation other than variable density acoustics (1.1). The behavior of these methods depends not so much on the specific equations chosen to model seismic waves, but rather on properties of the forward map shared by various models.

The inverse problem aims at solving for the model parameters, given the measurements of the pressure at the surface: Given d, solve for m such that $S[m] = d$. Note that though the wave equation itself is linear, the dependence of the solution on the model parameters is nonlinear. The inverse problem is therefore nonlinear.

The linearization of the inverse problem assumes a natural splitting of the model parameters into a background m_0 and a perturbation δm,

$$m = m_0 + \delta m. \tag{1.4}$$

The background is given and it is required to solve for the the perturbation δm. The formal derivative of the nonlinear forward map at m_0 maps the perturbation δm to δp:

$$F[m_0]\delta m = \delta p. \tag{1.5}$$

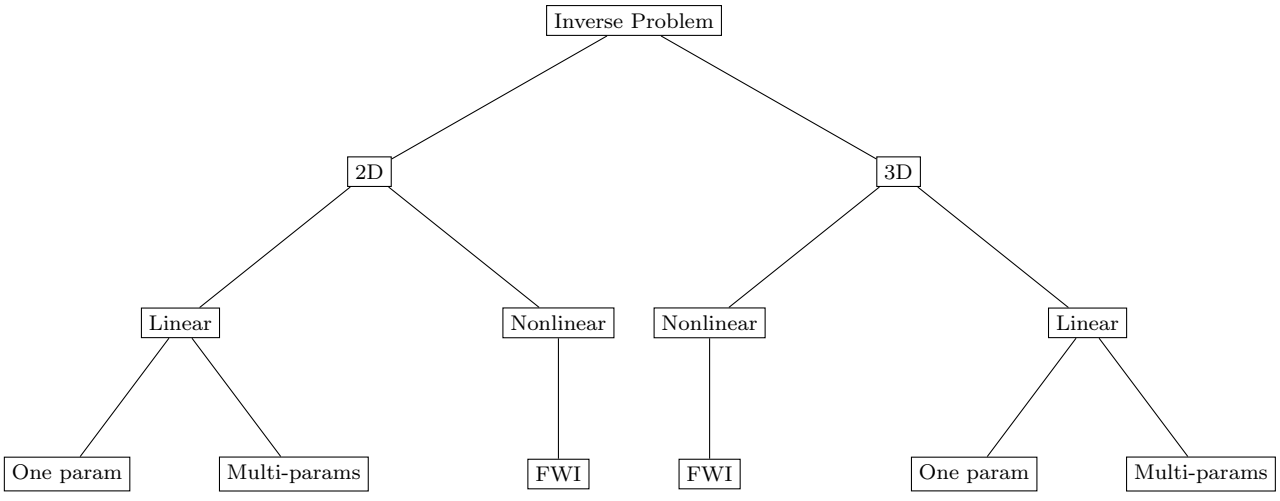
The linearized forward map F is referred to as Born Modeling. The linearized forward problem is: given δd, find δm such that $F\delta m = \delta d$. I discuss the details of the linearized inverse problem in the chapter devoted to linearization.

The acoustic wave equation models the dependence of the response of the earth on two parameters (density and velocity). One simplification assumes a constant

density field, constant density acoustics. The constant density approximation is valid when the acoustic waves are reflected from a region that exhibits negligible density contrast relative to velocity contrasts. Allowing for multi-parameter descriptions of the earth models allows for a better description of the underlying physics. However, it comes at the expense of introducing ill-conditioning in the inverse problem thus making it more difficult to solve accurately. This manuscript begins by describing one parameter inversions and then multi-parameter extensions.

Finally, an approximate solver for the linearized inverse problem may yield one way to accelerate the convergence of iterative methods for the nonlinear inverse problem (referred to as full waveform inversion methods or FWI in short). I explain this in the chapter on full waveform inversion.

I summarize the structure of this proposal in the following diagram:



Chapter 2

Theory and Literature Review

2.1 Introduction

This chapter describes the linearization of the inverse problem in the abstract setting, with emphasis on the case of variable density acoustics. Solving the linearized inverse problem requires solving the normal equations. The efficient approximation of the inverse of the normal operator relies on its pseudodifferential nature. For one parameter inversion, the application of the normal operator on one input vector suffices to infer its action and represent its approximate inverse from a class of pseudodifferential operators. The extension to p-parameters generalizes this result and requires p applications of the normal operator.

2.2 Linearization of the Inverse Problem

The linearization of the inverse problem splits the model parameters into a smooth part m_0 and a rough perturbation δm,

$$m = m_0 + \delta m. \tag{2.1}$$

we assume that the background reference model m_0 is given and inverse problem is to recover δm.

The nonlinear Forward map is thus approximated by,

$$S[m] = S[m_0 + \delta m] \approx S[m_0] + F[m_0]\delta m. \tag{2.2}$$

The linear operator F is known as born modeling. Formally F is the derivative of the nonlinear at m_0, it measure the sensitivity of S to small variations in the model. The conditions under which the right hand side of equation (2.2) provides an good approximation to its left hand side, are discussed in Stolk (2000). The fact that the background velocity is smooth (and some suitable conditions on m_0) implies that F is generically asymptotically invertible (Stolk, 2000), in 2D (an invertible Fourier integral operator). The same result is conjectured for 3D. The smooth part of the velocity field models the kinematics in the problem, it controls the large scale behavior of the propagation of the wave: travel times, positioning of reflectors .... The rough part describes the nature of the reflection decided by the the amplitude and nature of the discontinuities in the earth parameters.

The linearized inverse subproblem: Given d, m_0 find δm so that,

$$F[m_0]\delta m \approx d - S[m_0] := \delta d. \tag{2.3}$$

The linearized subproblem is an approximation due to the linearization process, and equation (2.3) is interpreted in a least squares sense to the arrive the the normal equations:

$$F^*[m_0]F[m_0]\delta m = F^*[m_0]\delta d. \tag{2.4}$$

The operator F^* is adjoint to F, and is known as the migration operator. The operator F^*F is called the normal operator or the Hessian. The right hand side of (2.4) is the migrated image $m_{mig} = F^*\delta d$.

An explicit linearization of the acoustic wave equation for example yields:

$$\begin{gathered} \frac{1}{\rho_0 c_0^2}\frac{\partial^2 \delta p}{\partial t^2} - \nabla \cdot \frac{1}{\rho_0}\nabla \delta p = \frac{2\delta c}{\rho_0 c_0^3}\frac{\partial^2 p}{\partial t^2} - \frac{1}{\rho_0}\nabla\frac{\delta\rho}{\rho_0}\cdot\nabla p_0 \\ \delta p \equiv 0, \quad t \ll 0, \end{gathered} \tag{2.5}$$

where ρ_0 and c_0 are the background density and velocity fields, respectively. The first order perturbations to ρ_0 and c_0 are $\delta\rho$ and δc, respectively.

The linear forward map is therefore,

$$F\delta m = F\left(\frac{\delta c}{c}, \frac{\delta\rho}{\rho}\right)^T = \delta p, \tag{2.6}$$

where δp is obtained by solving (2.5). Note that the solution δp is linear in the model perturbation δm.

Solution of the linearized inverse problem requires solving the normal equations (2.4), equivalently inverting the normal operator. In reality, the application of the

normal operator as seen from equation (2.5) and its adjoint requires the solution of large scale PDE problems. The process of applying the normal operator amounts to modeling followed by a migration. These processes typically require computations that can take days or weeks on computer clusters. Also the problem is large scale: in 2D the fields are length $\approx 10^6$, making the normal operator a $10^6 \times 10^6$ in matrix representation. These numbers prohibit explicitly storing the normal operator to invert it using direct matrix methods like Gaussian elimination. Krylov subspace methods are used to solve (2.4), but the expensive application of the F^*F limits the number of affordable iterations since these methods require at least one application of the normal operator per iteration.

The properties of the normal operator have been studied in the literature on the subject (Beylkin, 1985; Rakesh, 1988). For one parameter inversion (constant density acoustics), the normal operator is a pseudodifferential operator under specific conditions when the background velocity field is smooth. Stolk (2000) proves that the normal operator is the sum of a pseudodifferential operator and a non-microlocal part. He discusses the conditions under which the non-microlocal part is a Fourier integral operator and can be analyzed as such. Stolk (2000) concludes by proving that the normal operator is generically a pseudodifferential operator plus a smoother Fourier integral operator correction in 2D. The smoother error is of lower frequency order in high frequency asymptotics. In the more general case, the normal operator is a $p \times p$ matrix of pseudodifferential operators for p-parameter inversion, when scattering

preserves polarization as in P to P or S to S scattering (Beylkin and Burridge, 1990). The limitation to polarization preserving scattering is necessary, otherwise the normal operator maps one reflector to multiple reflectors in different places (one for each polarization). The normal operator would not preserve singularities in such case, and is therefore not a pseudodifferential operator. Symes (1998) provides an explicit proof that the normal operator is a two by two matrix of pseudodifferential operators for the variable density acoutics case.

Pseudodifferential operators provide a generalization of differential operators, they are defined by their action on a function $u(x)$:

$$Q_m u(x) = \int q_m(x,\xi)\hat{u}(\xi)e^{ix.\xi}\, d\xi, \tag{2.7}$$

where $q_m(x,\xi)$ is the symbol of the pseudodifferential operator, and $\hat{u} = \mathcal{F}[u]$ is the Fourier transform of u.

Pseudodifferential operators are defined in terms of their symbols $q_m(\mathbf{x},\xi)$,

$$q_m(\mathbf{x},\xi) : \Omega \times \mathbb{R}^n \backslash \{\mathbf{0}\} \to \mathbb{R},$$

where $\Omega \subset \mathbb{R}^n$ is an open set and $n = 2$ or 3 (the dimension of the space).

The symbols of interest $q_m(\mathbf{x},\xi)$ are smooth and homogeneous of order m, and for any compact set $K \subset \mathbb{R}^n$, and real α,β, there exists constants $C_{K,\alpha,\beta}$, such that

$$|D_x^\alpha D_\xi^\beta q_m(\mathbf{x},\xi)| \leq C_{K,\alpha,\beta}(1+|\xi|)^{m-|\beta|}, \tag{2.8}$$

for all $\mathbf{x} \in K$ and $\xi \in \mathbb{R}^n$. Homogeneity means that, given $r \in \mathbb{R}$,

$$q_m(\mathbf{x}, r\xi) = r^m q_m(\mathbf{x},\xi). \tag{2.9}$$

Homogeneous symbols satisfy (2.8); however, it should be noted that (2.8) is satisfied by a more general class of symbols not treated in this thesis. By allowing such general classes of symbols pseudodifferential operators generalize differential operators.

In fact, we can allow for more general symbols that admit a polyhomogeneous expansion of the form:

$$q \sim \sum_{j\geq 0} q_{m-j} \quad \text{, in the sense,} \quad q - \sum_{j=0}^{N-1} q_{m-j} \quad \text{is a symbol of order } m-N \tag{2.10}$$

Pseudodifferential operators of order m, generalize a differential operator of order m as operators between Sobolev spaces:

$$Q_m : H^s_{loc}(\Omega) \to H^{s-m}_{loc}(\Omega).$$

The first term in the polyhomogeneous expansion q_m is called the principal symbol of q, and the remainder of the expansion maps H^s to a smoother space H^{s-m+1}. Properties of pseudodifferential operators are thus defined up to smoother error, which in the frequency domain corresponds to a lower order in frequency, at high frequency.

In this thesis, pseudodifferential operators are represented by their principal symbol to capture the most singular part.

For a complete account on pseudodifferential operators and their applications in solutions of PDEs, please consult (Taylor, 1981).

Pseudodifferential operators act in phase-space, they are determined by their principal symbol $q_m(x, \xi)$ depending on both the spatial variable x and the momentum or Fourier variable ξ. The action of pseudodifferential operators is even more specific,

the singularities in u are preserved in Q_m, however the amplitude of these singularities will be modified depending on:

- Spatial position of the singularity,
- Orientation of the singularity (referred to as dip),
- The order m of the pseudodifferential operator.

We say that pseudodifferential operators act by *scaling* the input vector. To understand this behavior it is enough to compare δm and F^*d from equation (2.4) (Figures 2.1 and 2.2), for a specific model called the Marmousi model. Note that the inverse and the migrated image are related by the normal operator. It is obvious that the positions of the reflectors or discontinuities is preserved, the normal operators acts by scaling the amplitudes of δm.

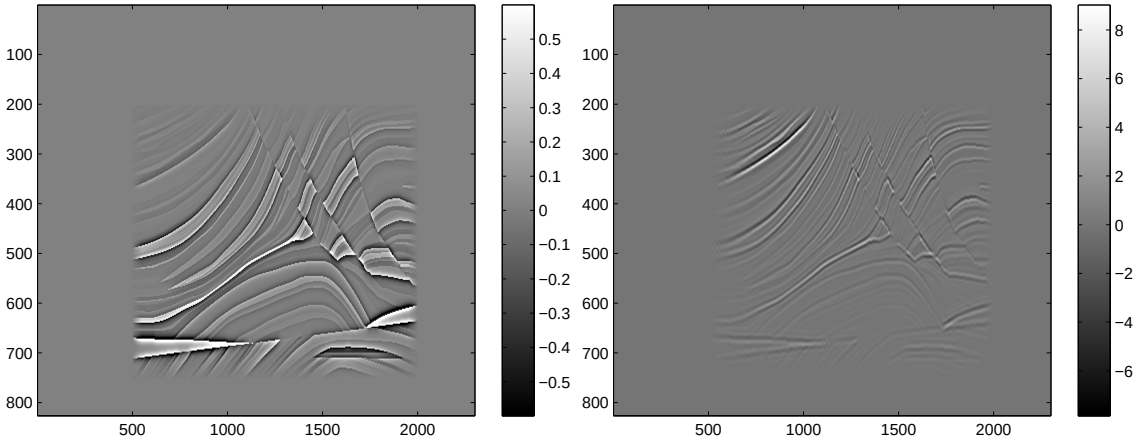


Figure 2.1: δm

Figure 2.2: $m_{mig} = F^*\delta d = (F^*F)\delta m$

The sense in which the normal operator acts by locally scaling the amplitudes of the discontinuities while preserving their position, is made precise using the asymptotic expansion lemma for pseudodifferential operators (Taylor, 1981). Let $\chi(x)$ be a smooth function compactly supported inside a ball, and $\Psi(x)$ a smooth function with non-vanishing gradient inside the same ball. I call a function of the form $\chi(x)e^{i\omega\Psi(x)}$ a *localized monochromatic pulse*. Then

$$F^*F\chi(x)e^{i\omega\Psi(x)} = q_m(x, \omega\nabla\Psi(x))\chi(x)e^{i\omega\Psi(x)} + O(\omega^{m-\beta}), \tag{2.11}$$

where $\beta > 0$, ω is the frequency and $q_m(x, \omega\nabla\Psi(x))$ is the principal symbol of the normal operator positively homogeneous of order m:

$$q_m(x, \omega\xi) = \omega^m q_m(x, \xi), \qquad \omega > 0 \tag{2.12}$$

We can even see how the scaling is related to the symbol of the pseudodifferential operator from (2.11).

In seismic images, reflectors are interfaces between two regions of space that exhibit discontinuities or high contrasts in physical parameters, these discontinuity account for the high frequency components (rapid changes). The relevant functions are of the form $a(x)f(\Psi(x))$, an amplitude modulated function of the phase. The high frequency components of these functions correspond to the discontinuities/reflectors. The vector $\nabla\Psi(x)$ is normal to the level sets of the phase function and represents dip as it points to a direction normal to the reflector identified with the level set of the phase function. This normal vector fails to be well defined in multiple dip events (fault, point reflector ...).

2.3 Scaling Methods

The realization that the normal operator acts by scaling the amplitudes of the right hand side of equation called the migrated image (2.4), lead to the idea that the action of the normal operator may be approximated from its application to a single input vector. In the same way diagonal matrices scale all vectors in the same way, the normal operator is approximately diagonal in a basis of localized monochromatic pulses (justified by equation (2.11)). I call methods that rely on one application of the normal operator to approximate it or to approximate its inverse *scaling methods*, and I refer to the approximations they yield as *scaling factors*. The choice of the migrated image as the input vector was suggested by Symes to Claerbout and Nichols as they developed an early scaling method in 1994. The choice of this input vector is motivated by the fact that the migrated image contains all relevant directions or reflectors.

Claerbout and Nichols (1994) propose approximating the normal operator and its inverse as a multiplication by a smooth function, using the migrated image as input vector. The method was then refined by Rickett (2003).

Guitton (2004) proposes a more general near diagonal approximation of the normal operator: near diagonal integral operators that are not completely specified by Guitton.

Symes (2008) proposes a correction to the Claerbout and Nichols method, he proves that the normal operator is approximated by multiplication by a smooth func-

tion after application of a Laplacian filter of a specific power, the power is predicted by the underlying theory and the filter is completely specified in contrast to Guitton's method. However, the method cannot approximate the inverse of the normal operator in places of the image that admit multiple dip events (faults, point reflectors ...). The method requires dip to be well defined in all parts of the image thus limiting its applicability. This method therefore fails locally in places that admit a fault or a point reflector for example.

Herrmann et al. (2008b) approximately diagonalize the normal operator in a frame of approximate localized monochromatic pulses, namely *curvelets*. They rely on the asymptotic expansion lemma (2.11), to justify their method. This method is capable of resolving multiple dip events.

I propose a generalization to the method introduced by Symes, a scaling method that resolves multiple dip events and skips the explicit diagonalization of the normal operator and the use of curvelets altogether. The method relies on a truncated spherical harmonics expansion of the symbol to approximate its action. The method reduces to the method proposed by Symes, when the expansion of the symbol consists of the first term. I present this method in more detail in the next chapter.

2.4 Amplitude Versus Offset (AVO)

The Zoeppritz equations specify how waves are transmitted and reflected at an interface. The study is conducted for the elastic wave equation, and the results specify the

reflection and transmission coefficients in terms of the offset angle (Aki and Richards, 1980). These equations are nonlinear in the physical parameter of the earth. Nonetheless, they can be solved if one is willing to invest the computational cost.

The complexity of the Zoeppritz equations led to attempts to simplify these equations while preserving the qualitative predictions of the full Zoeppritz equations. Usually these simplifications are linearizations in the relative differences of the physical parameters from the two sides of the interface. Aki and Richards (1980) present one of these linearizations (p. 153). Another simplification widely used in the AVO study is that presented by Shuey (1985). Shuey presents a formulation that demarcates the behavior for normal incidence, from intermediate angles (about 30 degrees), to wide angles (approach to critical angle) (Shuey, 1985).

While zero-offset reflection contains information about the acoustic impedance only, the variation of the reflection coefficients with offset angle contain information about all the elastic parameters (Lörtzer and Berkhout, 1989). In AVO, the Zoepritz equations and their various simplifications are used in an inverse problem sense to infer information about the elastic parameters from the variation of the amplitudes of the reflection coefficients as a function of offset (or offset angle).

Rutherford and Williams (1989) classifies different reflectors according to the qualitative variation of the reflection coefficient as a function of offset angles. This classification consists of three classes of reflectors that exhibit qualitatively different AVO behavior. Rutherford and Williams (1989) deals with gas sands encountered in ex-

ploration and splits them into:

1. Class 1: High impedance sands,

2. Class 2: Near zero impedance contrast sands, and

3. Class 3: Low impedance sands.

The use of AVO data to approximate the material parameters abounds in the literature. Lörtzer and Berkhout (1989) presents a statistically Baysesian based approach to predict a combination of physical parameters that predicts the AVO variation. The paper stresses the need for multi-component data to alleviate the ill-conditioning of the problem and pin down certain parameter combinations that are difficult to resolve (Lörtzer and Berkhout, 1989). It also quantifies how the resolution of different parameters varies as a function of the angle range, and how sensitive the recovery of parameters is to calibration parameters used in the statistical approach.

The method presented in this manuscript is an alternative to AVO analysis in that the inverse contains all the info that an AVO study can possibly yield about the different parameters. In fact, after approximating the Zoeppritz equations and taking into account the uncertainty in the data, AVO analysis yields information about anomalies in the physical parameters (usually the Poisson ratio) rather than quantitative measures of these parameters. Linearized inversion aims at recovering the physical parameters quantitatively.

2.5 Linearized Multi-Parameter Inversion

The attempts for multi-parameter inversion are limited in the literature. Bourgeois et al. (1989) study the linearized multi-parameter inversion and conclude that the success of this procedure relies on an accurate background velocity model. With an accurate background model, Bourgeois et al. (1989) conclude that the recovery of the impedance is possible in that it inversion yields significant corrections to seismic images produced by migration. The results of this paper confirm the well known fact that the recovery of the acoustic impedance is a the a well conditioned problem for variable density acoustics. The recovery of the other parameters remains challenging, an aspect of this inverse problem that is alluded to in the subsequent chapters.

Santosa and Symes (1988) present a study of the inverse problem for a layered acoustic fluid, they parametrize the problem in terms of the density and the incompressibility. They conclude that the problem is well conditioned away from critical angles of reflection and with enough aperture, in the absence of low velocity zones. In the presence of low velocity zones, the degradation of the conditioning of the inverse problem is unavoidable (Santosa and Symes, 1988). The study of the conditioning in this paper resembles the conditioning study I present on the normal operator in this manuscript. The inverse problem studied in this paper is effectively one dimensional (layered), the method presented here is not limited to 1D in fact the formulation is independent of dimension. The conditioning study, in both works, is restricted to layered models; these models allow for an analytical study of the conditioning of the

inverse problem.

Multi-parameter linearized inversion constitutes part of Minkoff (1995). In fact, she shows that the success of linearized multi-parameter inversion relies on pinning down other aspects of the inverse problem accurately: source estimation, background field approximation, modeling the physics accurately (including elasticity, attenuation ...). Minkoff (1995) tries, and succeeds, in fitting the data by including all the parameters mentioned above. Linearized inversion constitutes one of these steps but relies heavily on all the others. In this thesis, I assumed the source known, the background model given and the data was generated using the model used to fit it. These steps ensure that linearized multi-parameter inversion has the potential of succeeding.

Charara et al. (1996) invert for P-velocity, S-velocity and density in the linear elasticity inverse problem. Though the inversion is nonlinear, the estimation of these parameters is done in the linear regime (when the background fields are accurate enough). Charara et al. (1996) incorporates constraints on both the data space and the model space and uses a least squares inversion approach by formulating these constraints through covariance matrices as advocated in (Tarantola, 1987). Some of these covariance matrices rely on a priori knowledge like well data. The inversion relies on starting at a good initial model, stressing the role of the accurate background model in successful inversion.

The works presented above are all known as wave equation methods, they rely

on a numerical solution of the wave equation to obtain seismic images and data, and inversion is formulated in terms of these quantities. Another approach derives explicit formulas for the inverse of the linearized forward map, which involves calculating geometric optics quantities derived under an asymptotic regime. This work follows the approach of Beylkin (1985). These computations usually involve conditions on the medium of propagation for the geometric optics quantities to be well defined, and this reflects as an instability in computing these quantities in complex media. The method proposed in this manuscript is closer to wave equation methods, in that it does not require any computation of geometric optics quantities.

Foss et al. (2005) go through the necessary computations to derive the asymptotic inverse to the linearized map for anisotropic elastic media. The authors present a numerical example on the recovery of a linear combination of the density and two other elastic parameters, they claim that their "framework applies with decent accuracy" (Foss et al., 2005). The paper does not show synthetic examples so it is difficult to judge the accuracy of the framework.

Virieux et al. (1992) take a mixed approach to invert for the P and S impedance in linear elasticity. They use geometric optics computations to calculate the forward map and an approximation of the Hessian, and a Gauss-Newton iterative method for the inversion. The approximation of the Hessian accelerates the convergence of the inversion. Virieux et al. (1992) also study the conditioning of the inverse problem, and conclude that it is ill conditioned for single component data, multi-component data

is necessary for a successful inversion. The authors note that this is in accordance with the study of Santosa and Symes (1988) in the special case of a layered fluid. The inversion for P and S impedances is successful in the linear regime, the authors stress the importance of an accurate background model (Virieux et al., 1992).

The attempts for linear multi-parameter inversion in the literature, that use wave equation methods, rely on an iterative approach to minimize the least square misfit between the measured data and the data predicted by linearized modeling. The method proposed here, to use a few applications of the normal operator to produce an approximate inverse, is novel.

Chapter 3

Methods

3.1 Introduction

This chapter presents the scaling method to approximate the inverse of the normal operator efficiently for one parameter inversion, and its generalization to multi-parameter inversion. The inverse of the normal operator is represented from a class pseudodifferential operators defined by a truncated Fourier expansion of their symbol. The efficient approximation relies on an algorithm derived by Bao and Symes (1996) to approximate the action pseudodifferential operators, I refer to this algorithm as the PsiDO algorithm. This algorithm was derived in 2D but extends to 3D. The presentation of the method is independent of the space dimensions, the availability of the PsiDO code in 2D limits the applications to 2D only. I discuss the generalization to 3D briefly.

3.2 One Parameter Inversion: Pseudodifferential Scaling

Recall that the aim of this manuscript is to solve,

$$N\delta m = b, \tag{3.1}$$

where $N = F^*F$ is the normal operator and $b = F^*\delta d \in Range(N)$ is the migrated image.

Given b and Nb (we refer to Nb as the remigrated image). We seek a scaling factor c that minimizes the following objective function:

$$c = \underset{c\in\Psi DO}{argmin}\|b - cNb\|^2 \tag{3.2}$$

The advantage of obtaining c lies in the ability to derive an approximate inverse δm_{inv} given c:

$$\delta m = N^\dagger b \approx N^\dagger cNb \approx cN^\dagger Nb = cb := \delta m_{inv}. \tag{3.3}$$

The first equation expresses the solution of equation (3.1), where $N^\dagger$ is the pseudoinverse (regularized inverse) of N. The second approximation relies on the quality of the fit in equation (3.2). The third uses the property that pseudodifferential operators approximately commute. The successive approximations thus yield an approximate inverse $\delta m_{inv} = cb$.

The sense in which the scaling factor c approximates the inverse of the normal operator is specifically in the sense of (3.3): c scales the amplitudes of the migrated

image b in the same way the inverse of the normal operator $N^\dagger$ does. The scaling factor approximates the action of the inverse of the normal operator on one right hand side of the normal equations, namely the migrated image.

3.3 Multi-Parameter Inversion: Cramer's Rule for Pseudodifferential Operators

The generalization to p-parameters formally tries to solve the same problem:

$$N\delta m = b. \tag{3.4}$$

The model m is a collection of p-parameters, and the normal operator is therefore a $p \times p$ matrix of pseudodifferential operators in polarization preserving scattering. The p migrated images are contained in b.

The theory of pseudodifferential operators introduces a powerful concept: The algebraic relationships between symbols and matrices of symbols can be mapped to asymptotic properties for pseudodifferential operators. This concept is powerful since matrices of symbols are matrices of scalar functions, with a plethora of identities and theorems from linear algebra to choose from. This work uses a version of Cramer's rule to devise an inversion scheme for multi-parameter inversion.

For this end, we recall the definition of the adjugate of a matrix A, denoted by $Adj(A)$. Defined as the transpose of the matrix of cofactors of A. In our case, the transpose may be ignored, as the matrix A and thus its adjugate, are symmetric

positive definite (since N is too). When the matrix A is invertible the adjugate may be defined as

$$Adj(A) := det(A) * A^{-1}, \tag{3.5}$$

where A^{-1} is the inverse of the matrix A and $det(A)$ is the determinant of the matrix A. More generally, the adjugate is defined to be the matrix which satisfies:

$$Adj(A) * A = A * Adj(A) = det(A) * I, \tag{3.6}$$

where I is the identity matrix. If we define the adjugate of N to be $Ajd(N) = op(Adj(A))$, with slight abuse of notation. We can map the property (3.6) on matrices to a property on matrices of pseudodifferential operators:

$$Adj(N) * N \approx N * Adj(N) \approx det(N) * I. \tag{3.7}$$

The equation above features another abuse of notation, with $det(N) := op(det(A))$. The power of (3.6) is revealed when applied to (3.1).

$$Adj(N)\, b = Adj(N) * N\, \delta m \approx det(N) * \delta m. \tag{3.8}$$

Equation (3.8) recovers the inverse up to the pseudodifferential operator $det(N)$, after the application of the adjugate. The restriction of this approach to 1-parameter inversion and 2-parameter inversion is particularly simple and elegant. The extension to general p-parameters is more involved as shown by the case p=3, which we discuss in the following section.

The problem of recovering δm, is not solved yet. While δm is recovered up to the pseudodifferential operator $det(N)$, the inverse of this factor needs to be approximated

to complete the inversion scheme. For this end, we resort to a method similar to the one we previously developed for p=1.

First apply the normal operator again, to form:

$$N\, det(N) * \delta m \approx det(N) * N * \delta m = det(N) * b. \tag{3.9}$$

Where we have used the fact that scalar pseudodifferential operators approximately commute with matrices of pseudodifferential operators, to commute N and $det(N)$.

Now, given b and $det(N) * b$, approximate the scaling factor c:

$$c = \underset{c \in \Psi DO}{argmin} \|b - c * det(N) * b\|^2. \tag{3.10}$$

And approximate the solution of normal equations by:

$$\begin{aligned} \delta m = N^{-1} b \approx N^{-1}\, c * det(N) * b \approx c * det(N) * N^{-1} b \\ \approx c * det(N) * \delta m := \delta m_{inv} \end{aligned} \tag{3.11}$$

Thus the scaling factor c is an approximation of the inverse of $det(N)$ in that it is applied to $det(N) * \delta m$ (obtained previously), to approximate δm.

It is also straightforward to see that the one-parameter case is a restriction of this general approach to $p = 1$.

The efficiency of this approach relies on the efficiency of the optimization in (3.10). Any optimization scheme will require the application of the pseudodifferential scaling factors at each iteration. It is therefore pivotal to utilize an algorithm that applies pseudodifferential operators efficiently. Bao and Symes (1996) develop an algorithm to efficiently approximate the action of pseudodifferential operators, which relies on

a spherical harmonics expansion of their symbols. We refer to this algorithm as the PsiDO algorithm.

The advantage of using the PsiDO algorithm is twofold: first its efficiency, and second the ability to represent pseudodifferential operators that act by spacial, frequency and dip dependent scaling capable of resolving multiple dip events.

3.4 The PsiDO Algorithm

I present the PsiDO algorithm that allows the efficient representation and approximation of the scaling factors. As developed by Bao and Symes (1996), the algorithm is presented explicitly in 2D. The extension of the PsiDO code to 3D constitutes part of my future work.

This discussion is restricted to 2D, so we may write $\mathbf{x} = (x, z)$. Recall that a pseudodifferential operator is characterized by its symbol and defined by

$$Q_m u(x,z) = \int\int q_m(x,z,\xi,\eta)\hat{u}(\xi,\eta)e^{i(x\xi+z\eta)}\,d\xi\,d\eta, \tag{3.12}$$

where $q_m(x,z,\xi,\eta)$ is the principal symbol, homogeneous of degree m, and $\hat{u} = \mathcal{F}[u]$ is the Fourier transform of u.

Thus writing $\xi = \omega\cos\theta$, $\eta = \omega\sin\theta$, and using the homogeneity of q_m, we have

$$q_m(x,z,\xi,\eta) = \omega^m \tilde{q}_m(x,z,\theta). \tag{3.13}$$

Notice that $\tilde{q}_m(x,z,\theta) = q_m(x,z,\cos\theta,\sin\theta)$ is periodic and smooth in θ, and hence it admits a rapidly converging Fourier expansion. We thus truncate the Fourier

series, approximating the symbol by its first $K+1$ Fourier modes:

$$\tilde{q}_m(x,z,\theta) \approx \sum_{l=-K/2}^{l=K/2} c_l(x,z)e^{il\theta} = \sum_{l=-K/2}^{l=K/2} \omega^{-l} c_l(x,z)(\xi+i\eta)^l. \tag{3.14}$$

Plugging (3.14) into (3.12) we obtain

$$Q_m u(x,z) \approx \sum_{l=-K/2}^{l=K/2} c_l(x,z)\mathcal{F}^{-1}[\omega^{m-l}(\xi+i\eta)^l \hat{u}(\xi,\eta)]. \tag{3.15}$$

Fourier transform theory identifies ω^{m-l} as the symbol of $(-\nabla)^{\frac{m-l}{2}}$, and ξ and η are respectively the symbols of $D_x = -i\partial_x$ and $D_z = -i\partial_z$.

Sampling the field $u(x,z)$ and the symbol $\tilde{q}_m(x,z,\theta)$,

$$U_{ij} = u(x_0 + (i-1)\Delta x, z_0 + (j-1)\Delta z),$$

$$Q_{ijk} = \tilde{q}_m(x_0 + (i-1)\Delta x, z_0 + (j-1)\Delta z, k\Delta\theta),$$

$$i = 1, \cdots, M,\ j = 1, \cdots, N,\ k = -K/2, \cdots, K/2.$$

Choosing $\Delta\xi = \frac{1}{(M-1)\Delta x}$ and $\Delta\eta = \frac{1}{(N-1)\Delta z}$ yields the unaliased discretizations of the symbols of the square root of the negative Laplacian, D_x and D_z

$$\Omega_{pr} = 2\pi\sqrt{(p\Delta\xi)^2 + (r\Delta\eta)^2}$$

$$\Xi_{pr} = 2\pi p\Delta\xi$$

$$Z_{pr} = 2\pi r\Delta\eta$$

$$p = -M/2, \cdots, M/2, \quad r = -N/2, \cdots, N/2$$

Equation (3.15) suggests the following algorithm to estimate $Q_m u$ (Bao and Symes, 1996). All Fourier transforms refer to a discrete Fourier transform.

1. Compute $\hat{U}_{pr} = \mathcal{F}[U_{ij}]$.

2. For each $i \in [1, M]$ and $j \in [1, N]$,

   compute $\hat{Q}_{ij} = \{\hat{Q}_{ijl}\}_{l=-K/2}^{K/2}$ the discrete Fourier transform of $Q_{ij} = \{Q_{ijk}\}_{k=-K/2}^{K/2}$.

3. Initialize $(QU)_{ij} = 0$, for $i \in [1, M]$, $j \in [1, N]$,

   For $l = -K/2 : K/2$

   (a) compute $\{R_{ij}^l\}_{i=1,j=1}^{M,N} = \mathcal{F}^{-1}[\Omega_{pr}^{m-l}(\Xi_{pr} + iZ_{pr})^l \hat{U}_{pr}]$

   for $p = -M/2, \cdots, M/2$ and $r = -N/2, \cdots, N/2$

   (b) accumulate

   $$(QU)_{ij} = (QU)_{ij} + \hat{Q}_{ijl} R_{ij}^l$$

   End

A straightforward discretization of (3.12) has a computational complexity of $O(N^4 \log(N))$. The algorithm described above uses FFT (Fast Fourier Transform), and thus exhibits a complexity of $O(KN^2(\log(N) + \log(K)))$. The appeal of this approach is that K is independent of N. In fact, applications to reflection seismology require that the symbol be smooth and slowly varying in θ, thus may be captured accurately by a modest number of Fourier modes or, more explicitly, a small K.

The dependence on dip is captured in the angle variable θ and the method allows us to capture multiple dip events by increasing $K > 1$.

The extension of this algorithm to 3D requires the replacing the Fourier expansions of the symbol with spherical harmonics expansion to split the spacial and Fourier

dependence of the symbol.

3.5 Summary

I summarize the procedure that constitutes the scaling method.

To solve $N\delta m = b$,

- Apply $Adj(N)$ on b to form: $Adj(N) * b \approx det(N) * x$
- Apply N to the result to get: $N * det(N) * x \approx det(N) * b$
- Represent the scaling factors using the PsiDO algorithm : $c_i = Q_m[q_i]$
- Compute the scaling factor c:

$$c = \underset{c \in \Psi DO}{argmin} \|b - c * det(N) * b\|^2.$$

- Approximate the inverse: $x_{inv} := c * det(N) * x \approx x$

The method applies in 2D and 3D, the explicit representation of the scaling factors is limited to 2D by the PsiDO algorithm. Extension of the PsiDo algorithm to 3D constitutes a part of my future work and is presented in the proposal chapter.

Chapter 4

Results

4.1 Introduction

In this chapter I summarize the results of the masters thesis concerning constant density acoustics inversion. I validate the scaling method for one parameter inversion on the Marmousi benchmark model. I then discuss the challenges of multi-parameter inversion: the ill conditioning of the recovery problem for the impedance and density for variable density acoustics. Preconditioning the inverse problem alleviates the ill conditioning and permits the simultaneous recovery of impedance and density.

4.2 One parameter inversion: Constant density acoustics

In this case the adjugate is particularly easy, $Adj(N) = I$, and $det(N) \equiv N$. And this approach boils down to approximating c s.t.

$$c = \underset{c \in \Psi DO}{argmin} \|b - c * N * b\|^2. \tag{4.1}$$

I validate one parameter inversion on the 2D Marmousi synthetic benchmark model (Versteeg and Grau, 1991). The model is smoothed to construct the perturbation $\mathbf{m}_{true}$ (Figure 4.1). The images are windowed and tapered to the window of interest. The true model is born modeled and then migrated to obtain the migrated image and the process is repeated to obtain the remigrated image (Figures 4.2 and 4.3). These images show the amplitude distortion resulting from the application of the normal operator. It is obvious that the amplitudes in the deeper part of the image are attenuated, making these regions invisible without amplitude correction. The pseudodifferential scaling method with $K = 1$ and $K = 5$ yields scaling factors that I apply to the migrated image to obtain the approximate inverses (Figures 4.4 and 4.5). The amplitudes are recovered to the right order of magnitude. Moreover, the amplitudes are uniform in depth and compare better to those of the real image. Both these results are successful approximate inversions.

The scaling method with $K = 1$ cannot resolve multiple dip events, in contrast with $K = 5$. I plot the difference between the two inversion results in figure 4.6 to

study this feature. It is apparent that the amplitude difference is greatest at the locations of multiple dip events, where two reflectors intersect (faults). The high amplitude difference appears as brighter or dimmer spots in figure 4.6.

Please consult my masters thesis, where I present a better test of the ability of the scaling method to resolve multiple dip events.

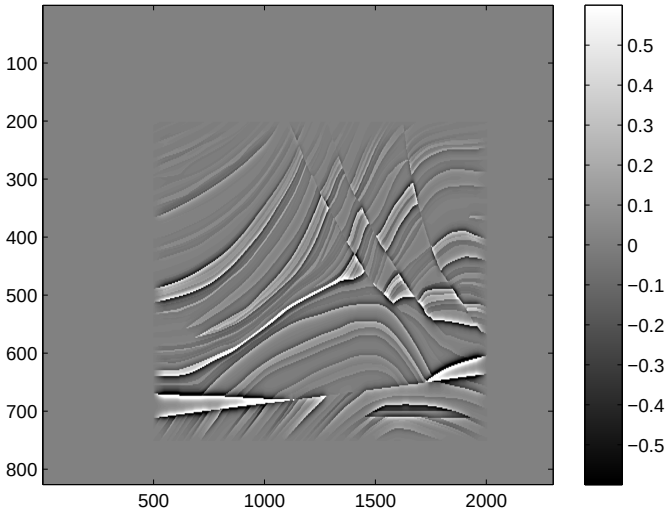


Figure 4.1: $\mathbf{m}_{true}$

4.3 Two-parameter case: p=2

The restriction two two-parameters is appealing because it is the first instance of multi-parameter inversion. It turns out that it is also particularly simple and elegant.

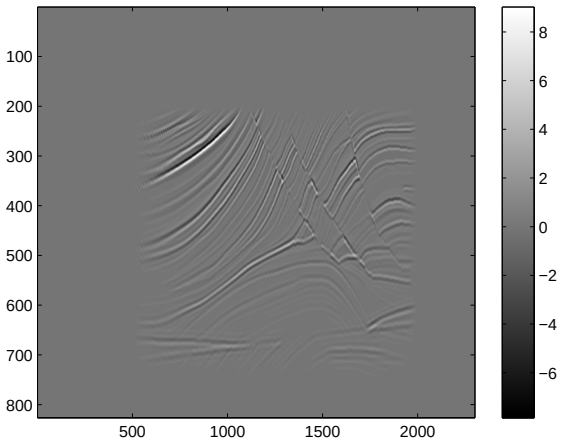


Figure 4.2: $\mathbf{m}_{mig} = F^* d$

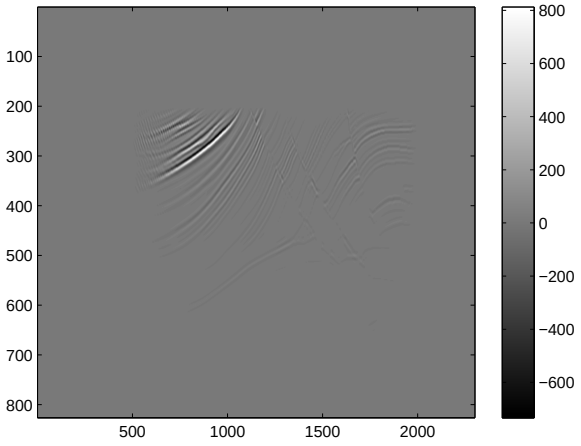


Figure 4.3: $\mathbf{m}_{remig} = F^* F \mathbf{m}_{mig}$

In this case, denoting,

$$N = \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}.$$

Its Adjugate is then given by,

$$Adj(N) = \begin{pmatrix} N_{22} & -N_{12} \\ -N_{12} & N_{11} \end{pmatrix}.$$

And

$$Adj(N) * b = \begin{pmatrix} N_{22}b_1 - N_{12}b_2 \\ -N_{12}b_1 + N_{11}b_2 \end{pmatrix}. \tag{4.2}$$

The aim is to express (4.2) as a combination of permutations of indices of b and applications of N. We will here introduce a notation that allows the manipulation of expressions like (4.2), while it seems too involved for $p = 2$, it will simplify the manipulation for $p \geq 3$, it is mostly an exercise in representation theory.

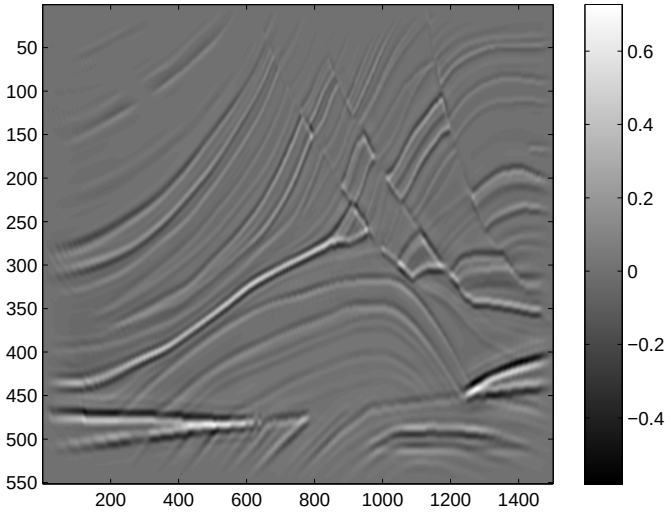


Figure 4.4: Scaling with $K = 1$

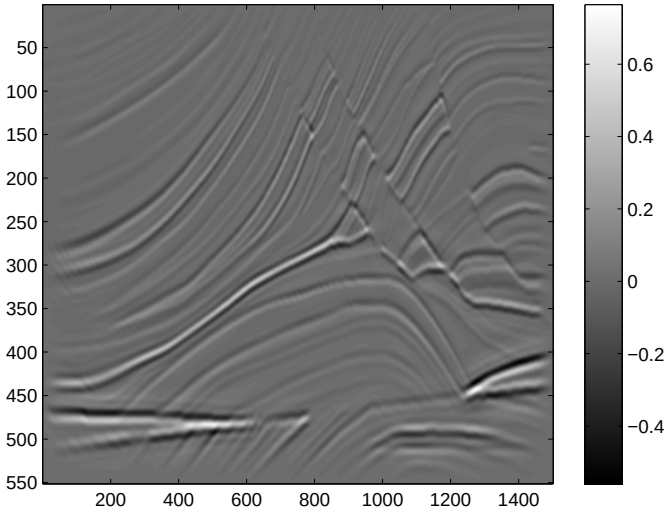


Figure 4.5: Scaling with $K = 5$

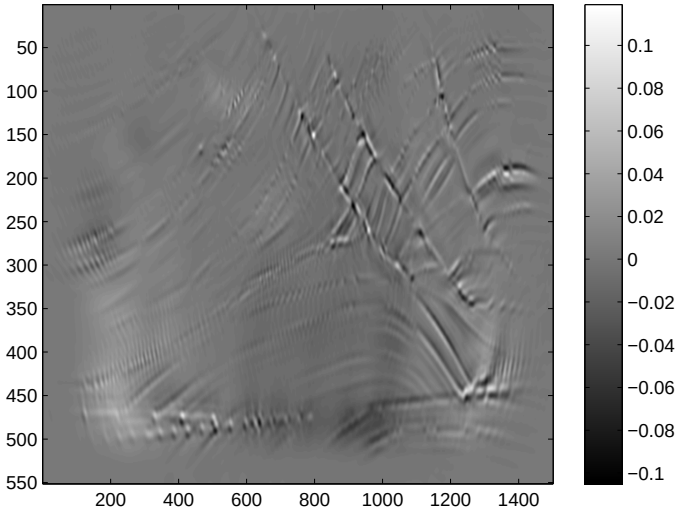


Figure 4.6: Difference between scaling with $K = 5$ and $K = 1$

Denote,

$$N_{ij}b_k e_l = e_l^T e_i N e_j^T e_k b := lijk. \tag{4.3}$$

Note that the symmetry of N allows us to commute i and j if needed.

The power of this representation becomes apparent in the following manipulation:

$$\begin{aligned} Adj(N) * b &:= 1221 - 1122 - 2121 + 2112 \\ &= 1221 - 1212 - 2121 + 2112 \\ &= (12 - 21)(21) + (21 - 12)(12) \\ &= (12 - 21)(21 - 12) \end{aligned} \tag{4.4}$$

In the first step we use commute indices, and the afterwards we are using matrix factorization. The end result is interpreted as:

$$\begin{aligned} Adj(N) * b &:= (12 - 21)(21 - 12) \\ &= (e_1^T e_2 - e_2^T e_1) N (e_2^T e_1 - e_1^T e_2) b \\ &= -JNJ = J^T N J\, b. \end{aligned} \tag{4.5}$$

Where,

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Equation (4.5), implies that the application of the adjugate on the migrated image requires one application of N and another to approximate the scaling factor. Which brings the cost of approximating the inverse for $p = 2$ to two applications of the normal operator.

4.3.1 Application: Layered variable density acoustics

As a first application to the two parameter inversion, we construct a variable density acoustics model perturbation consisting of velocity layer, and a density layer in a different place (See Figures 4.7 and 4.8). The background model is homogeneous, with $vp = 2\,km/s$ and $dn = 2000\,kg/m^3$.

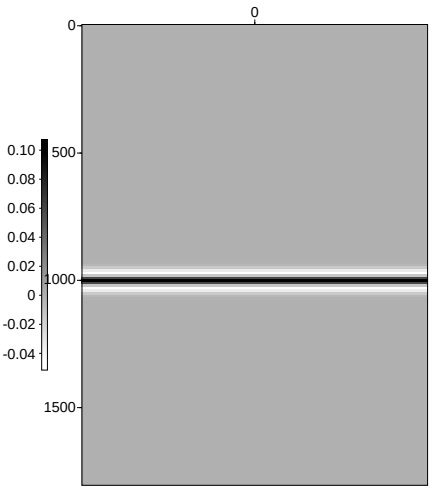


Figure 4.7: vp

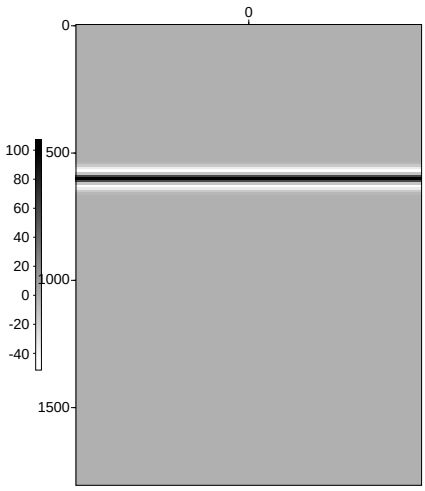


Figure 4.8: dn

Migrating the model perturbation shows how migration mixes the effects of the two models in the two components of the migrated images (Figures 4.10 and 4.9. We shall refer to the migrated images as b_1 and b_2, to remain consistent with our notation where the vector of migrated images is b. This example, albeit simple, stresses a new challenge of multi-parameter inversion: For one parameter inversion, the events in the migrated image corresponded to events in the true model. In multi-parameter

inversion, events in the migrated images may correspond to an event in one or more of the components of the model. It is virtually impossible to tell that these migrated images correspond to a model with separate events for velocity and density without successful inversion. Applying the scheme outlined above, we form

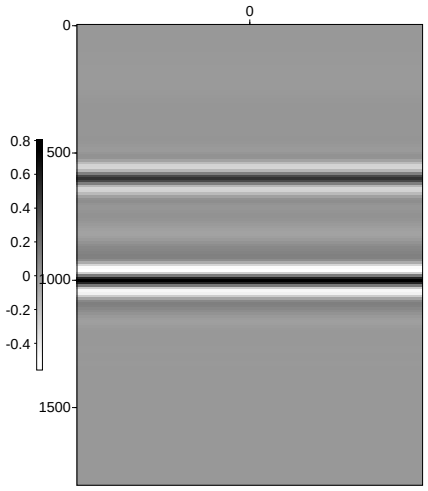


Figure 4.9: b_1

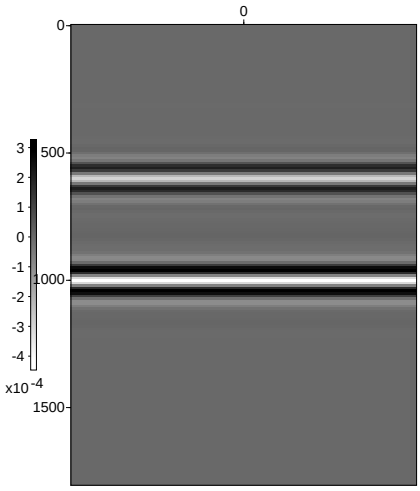
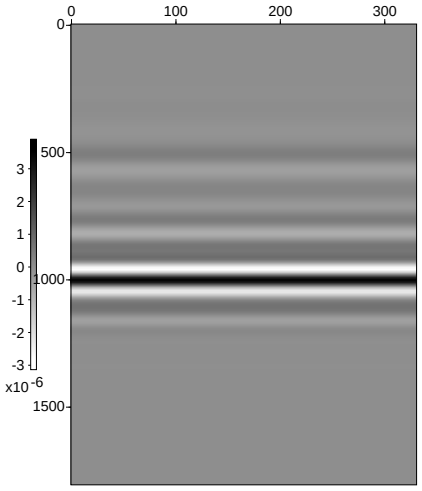


Figure 4.10: b_2

$$J^T N J\, b \approx det(N) * x.$$

The result is shown in Figures 4.11 and 4.12, and shows how one application of the normal operator effectively separated the contributions of the velocity and density events. It remains to effect an amplitude correction, by approximating an inverse to $det(N)$. For this end, we are required to form $N * det(N) * x \approx det(N) * b$, shown in Figure 4.13 and 4.14.



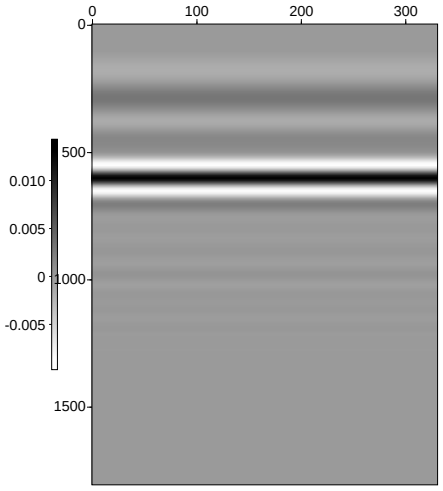


Figure 4.11: $(J^T N J\, b)_1 \approx det(N) * x_1$ Figure 4.12: $(J^T N J\, b)_2 \approx det(N) * x_2$

The final step corrects the amplitudes of $det(N) * x$ by applying undoing the effect of $det(N)$, which yields an approximate inverse. This final step complements the separation we obtained earlier with an amplitude correction, Figure 4.15 and 4.16 shows that the approximate inverse compares favorably with true model. An interesting observation on this result is the fact that the velocity model is better recovered than the density model: traces of the velocity event in the density model are more apparent than that the density event in the velocity model. This observation is in accordance with the theoretical fact that the recovery of velocity in variable density acoustics is better conditioned than the recovery of density. The following section explains this fact in more detail.

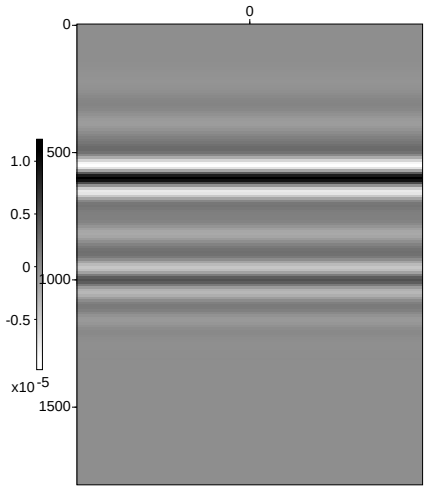


Figure 4.13: $det(N) * b_1$

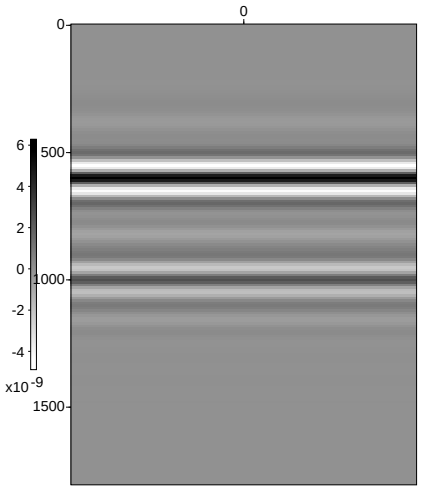


Figure 4.14: $det(N) * b_2$

4.4 Conditioning of the normal operator

While the results of the previous section prove to be successful in one instance and for a particular geometry, a successful inversion method requires more general preconditioners. We thus turn to the general form of the symbol of the normal operator for variable density acoustics in 2D derived in (Symes, 1998).

$$A = f(\theta) \begin{pmatrix} 1 & \sin^2(\frac{\theta}{2}) \\ \sin^2(\frac{\theta}{2}) & \sin^4(\frac{\theta}{2}) \end{pmatrix} |\xi| \tag{4.6}$$

The opening angle θ depends on source position x_s, receiver position x_r and spatial position x. ξ is the Fourier variable. I suppress the explicit form of the function $f(\theta)$ as the ill-conditioning of the matrix is due to the matrix part of equation (4.6).

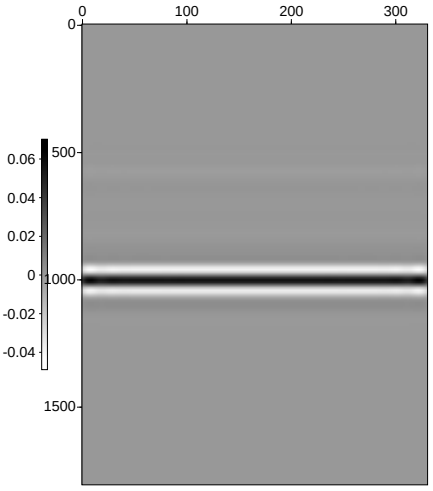


Figure 4.15: inv_{vp}

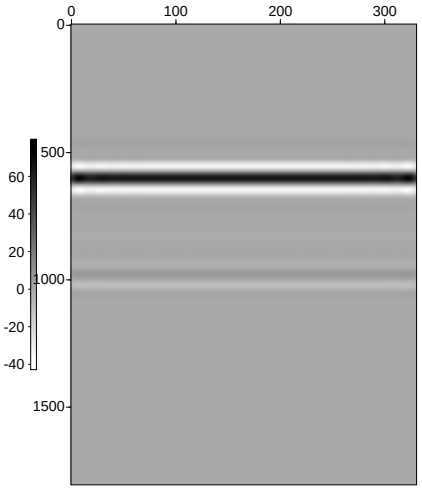


Figure 4.16: inv_{dn}

In what follows, we will search for the weight function $f(\theta)$ that minimizes the condition number of the normal operator. This approach yields an optimal weight $f(\theta)$ that renders the normal operator better conditioned.

To be explicit, the study of the conditioning of matrices of the form:

$$N = \int_0^{\theta_{max}} d\theta\, f(\theta) \begin{pmatrix} 1 & \sin^2(\frac{\theta}{2}) \\ \sin^2(\frac{\theta}{2}) & \sin^4(\frac{\theta}{2}) \end{pmatrix}.$$

Denote the eigenvalues of N by $0 < \lambda_{min} \leq \lambda_{max}$ (since the matrix is positive definite). Minimize the condition number:

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}}, \quad s.t.\ f \geq 0,\ \int_0^{\theta_{max}} f(\theta)\, d\theta = 1.$$

If the condition number is parametrized in terms of $S = \lambda_{max} + \lambda_{min} = trace(N)$

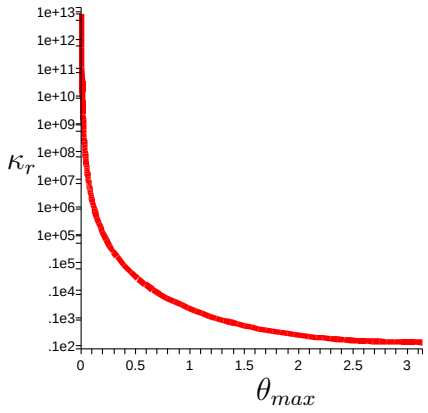


Figure 4.17: Condition number as a function of θ_{max}

and $P = \lambda_{max}\lambda_{min} = det(N)$, then

$$\kappa = \frac{S + \sqrt{S^2 - 4P}}{S - \sqrt{S^2 - 4P}} \tag{4.7}$$

As a reference, we study the condition number κ_r as a function of θ_{max} for $f(\theta) = \frac{1}{\theta_{max}}$ (correct normalization). Figure 4.17 is a logarithmic plot of the condition number of N as a function of θ_{max}, it shows how the condition number increases as the maximum offset angle decreases. The main source of ill-conditioning in this problem is the maximum offset angle.

A candidate weighting to ameliorate the condition number of N is one that amounts to a low offset/large offset stack:

$$f(\theta) = (1 - \alpha)\delta(\theta) + \alpha\delta(\theta - \theta_{max}), \quad 0 \leq \alpha \leq 1. \tag{4.8}$$

This type of stack that puts emphasis on large offset and small offset separately has

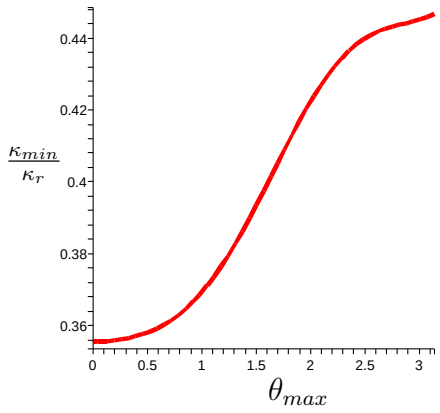


Figure 4.18: Ratio of optimal condition number to reference

been used since it is known that the different offset ranges give different information about the underlying physical parameters.

Minimizing κ with the weight given in (4.8), gives

$$\alpha = \frac{1}{2+\beta},$$

$$\kappa_{min} = \frac{\beta + 1 + \sqrt{1+\beta}}{\beta + 1 - \sqrt{1+\beta}},$$

where $\beta = \sin^4(\frac{\theta_{max}}{2})$.

It is interesting to note that that the result depicted above predicts that for large offset ($\theta_{max} \to \pi$), small offsets are weighted double.

To compare the condition number obtained with this weight to the reference case, I plot the ratio $\frac{\kappa_{min}}{\kappa_r}$ as a function of θ_{max} in Figure 4.18. The reduction in condition number is about 33% of its reference value for low offsets, and 45% for large offsets.

The reduction in the condition number is not dramatic which warrants a closer look at the asymptotics of the condition number. For small θ_{max}:

- Reference case:
  - $\lambda_{max} = 1 + \mathcal{O}(\theta_{max}^4)$
  - $\lambda_{min} = \frac{\theta_{max}^4}{180} + \mathcal{O}(\theta_{max}^6)$
  - $\kappa_r = \frac{180}{\theta_{max}^4} + \mathcal{O}(\theta_{max}^{-2})$
- Optimal stacks:
  - $\lambda_{max} = 1 + \mathcal{O}(\theta_{max}^4)$
  - $\lambda_{min} = \frac{\theta_{max}^4}{64} + \mathcal{O}(\theta_{max}^8)$
  - $\kappa_{min} = \frac{64}{\theta_{max}^4} + \mathcal{O}(1)$

The two cases exhibit the same asymptotics, which explains why the reduction in the condition number is not in orders of magnitude. Interestingly, the result above also predicts that an adequately weighed small offset/large offset stack, is better conditioned than using the entirety of the offset range!

The first order conditions for this problem turn out to emphasize an interesting property of this optimization problem. The first order conditions are,

$$\delta\kappa = 0 \Rightarrow 2\frac{\delta S}{S} = \frac{\delta J}{J}.$$

The solution to this equation yields. a different parametrization of the optimization

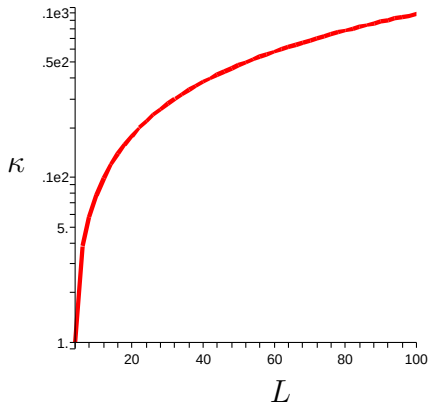


Figure 4.19: Condition number as a function of $L = \frac{S^2}{P}$

problem in terms of L, with $L \geq 4$:

$$S^2 = LP \Rightarrow \frac{S^2}{P} = L.$$

Which we can note also algebraically,

$$\kappa = \frac{S + \sqrt{S^2 - 4P}}{S - \sqrt{S^2 - 4P}} = \frac{1 + \sqrt{1 - \frac{4}{L}}}{1 - \sqrt{1 - \frac{4}{L}}}.$$

Minimizing the condition number amounts to minimizing $L = \frac{S^2}{P}$, to move along the curve that describes the condition number as a function of L, see Figure 4.19.

Another important piece of information that we can obtain from the reference case is the eigenvector with the largest eigenvalue, which we can calculate analytically. The result is not shown here, and the calculations are done using MAPLE. However, this calculation reveals the conditioning of the recovery of the interesting physical parameters: velocity, density and impedance. The angle that the vector representing

each of these parameters makes with the eigenvector corresponding to the the largest eigenvalue, describes the conditioning of the recovery of the parameter at hand. The results are shown in Figures (4.20), (4.21) and (4.22). In fact these figures explain the assertion made earlier about the recovery of the density being more ill conditioned compared to the recovery of the velocity. Figure (4.21) shows that, for small offset angle, the density is almost perpendicular to the optimal eigendirection, and therefore is aligned with the eigenvector with the smallest eigenvalue. The impedance is the best conditioned physical parameter for recovery, as it is aligned with the optimal eigendirection for small coverage angle. The velocity is in between, starting off with an angle of 45 degrees, and remaining in the mid-range. It is in this sense that the recovery of the density is the most ill conditioned of the three physical parameters chosen ususally.

4.5 Multi-component data

Another way to ameliorate the conditioning of the normal operator is measuring multi-component data. The data considered classically is the measurement of the pressure field at the surface. When additional data is available, it may be used to better pinpoint the material parameters sought in the inverse problem and accordingly improve the condition number of the normal operator.

We consider the case where the linearized forward map F mas the relative perturbations perturbations in the impedance and density to pressure perturbations and

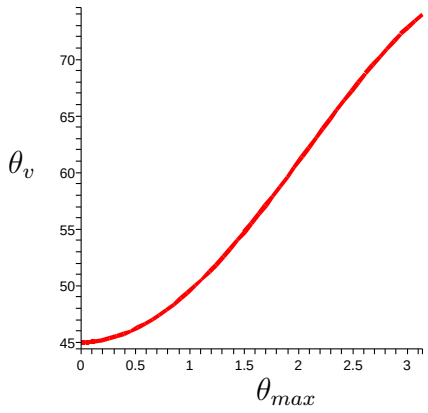


Figure 4.20: Angle in degrees that the velocity vector makes with the eigenvector corresponding to the largest eigenvalue as a function of maximum offset angle

averaged vertical derivatives of such perturbations. The second set of measurement is available and is known as over under cable data, a practical method to measure the vertical gradient of the pressure field at the surface.

$$F : \begin{pmatrix} \frac{\delta\sigma}{\sigma}(x) \\ \\ \frac{\delta\rho}{\rho}(x) \end{pmatrix} \rightarrow \begin{pmatrix} \delta p(x_s, x_r, t_r) \\ \\ c_0 \int_0^{t_r} \frac{\partial p}{\partial z_r}(x_s, x_r, t'_r)\, dt'_r \end{pmatrix}. \tag{4.9}$$

As usual x_s, x_r, x are respectively the source, receiver, and spatial positions. t_r is time sampled at the receiver location. z_r is the vertical component of x_r. A procedure similar to the one used in (Symes, 1998) yields an expression for the symbol of the normal operator for this specific case of multicomponent data. Surprisingly, it takes

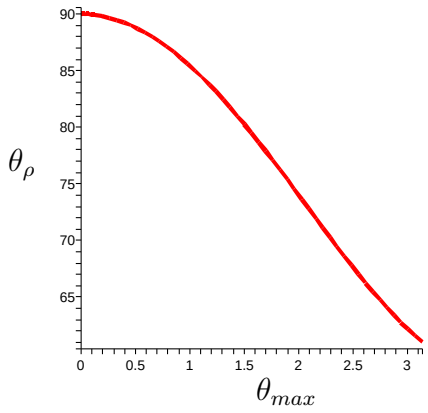


Figure 4.21: Angle in degrees that the density vector makes with the eigenvector corresponding to the largest eigenvalue as a function of maximum offset angle

the form:

$$A' = f(\theta) \begin{pmatrix} a(\theta)^2 & -a(\theta)^2 \sin^2(\frac{\theta}{2}) \\ -a(\theta)^2 \sin^2(\frac{\theta}{2}) & a(\theta)^2 \sin^4(\frac{\theta}{2}) \end{pmatrix}. \tag{4.10}$$

We can write out $a(\theta) > 0$ explicitly. However, it is more interesting to note that the effect of the multi-component data improves the condition number by changing the weight in front of the matrix studied in the previous section (the trace is scaled by a^2 and the determinant is scaled by a^4 , hence the eigenvalues are scaled by a^2 and the condition number remains unchanged). In the case of variable density acoustics, and for this type of multicomponent data, we can improve the conditioning by optimally weighting the matrix as discussed in the previous section. In fact, adequately weighting the matrix emulates multi-component data with single-component data! And the

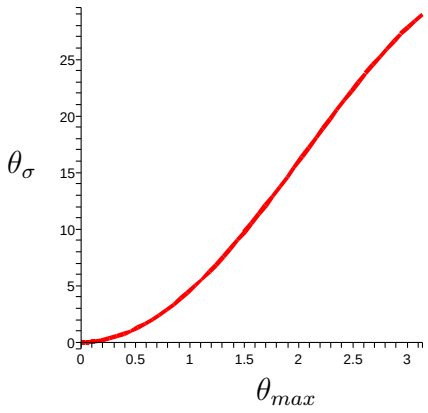


Figure 4.22: Angle in degrees that the impedance vector makes with the eigenvector corresponding to the largest eigenvalue as a function of maximum offset angle

study in this section turns out to lead to the previous section again.

Chapter 5

The Proposal

5.1 Introduction

This chapter presents the parts that constitute my future work and relates them to the parts of the thesis previously done.

5.2 Extension to 3D

The only part that restricts the application of the method to 2D is the PsiDO algorithm written explicitly for 2D. An extension of the PsiDO algorithm to 3D effectively opens up the possibility of applying the approximate inversion method to 3D models.

The 3D extension proceeds in the same fashion as Bao and Symes (1996), first write the symbol as:

$$q_m(x, y, z, \xi, \eta) = \omega^m \tilde{q}_m(x, y, z, \theta, \phi). \tag{5.1}$$

The truncated spherical harmonics expansion of $\tilde{q}$ is then given by:

$$\tilde{q}(x,y,z,\theta,\phi) = \sum_{l=0}^{K} \sum_{n=-l}^{l} c_{ln}(x,y,z) Y_l^n(\theta,\phi), \tag{5.2}$$

where Y_l^n are the spherical harmonics basis function. The coefficients can be calculated as,

$$c_{ln}(x,y,z) = \int_{\Omega} \tilde{q}(x,y,z,\theta,\phi) Y_l^{n*}\, d\Omega = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta sin(\theta) \tilde{q}(x,y,z,\theta,\phi) Y_l^{n*}. \tag{5.3}$$

The relationship between ω, ϕ and θ to ξ, ζ, and η is given my the spherical coordinates transformation:

$$\begin{aligned} \xi &= \omega cos(\phi) sin(\theta) \\ \zeta &= \omega sin(\phi) sin(\theta) \\ \eta &= \omega cos(\theta) \end{aligned} \tag{5.4}$$

In order to evaluate the action of the PsiDO algorithm, we need to express $Y_l^n(\theta,\phi) = Y_l^n(\xi,\zeta,\eta)$. In fact,

$$Y_l^n(\theta,\phi) = \sqrt{\frac{(2l+1)(l-n)!}{4\pi(l+n)!}} P_l^n(cos(\theta)) e^{in\phi}. \tag{5.5}$$

@here P_l^n are the associated Legendre polynomials, they may be calculated by recursion formulas or using an identity known as the Rodrigues' formula. Thus,

$$\begin{aligned} \omega &= \sqrt{\xi^2 + \zeta^2 + \eta^2} \\ cos(\theta) &= \frac{\eta}{\omega} \\ e^{in\phi} &= e^{in \tan^{-1}\left(\frac{\zeta}{\xi}\right)} \end{aligned} \tag{5.6}$$

Plugging into the action of a pseudodifferential operator,

$$Q_m\, u(x,y,z) = \sum_{l=0}^{K} \sum_{n=-l}^{l} c_l(x,y,z) \mathcal{F}^{-1} \left\{ \omega^m Y_l^n(\xi,\zeta,\eta) \hat{u}(\xi,\zeta,\eta) \right\} \tag{5.7}$$

The cost of this algorithm is $(K+1)^2$ applications of the inverse Fourier transform, if we use FFT, the cost would be $(K+1)^2 N^3 \log(N)$.

If we only use this algorithm to approximate scaling factors as discussed in the methods section, then we only need to represent the associated Legendre polynomials explicitly. The symbol is then parametrized by $c_l(x,y,z)$. In fact we do not require any spherical harmonics transforms, such transforms are needed if we are given a symbol and asked to apply its action, which is not the problem at hand since we never have access to the symbol of the normal operator.

5.3 Full Waveform Inversion

This section shows how the scaling method I developed in the setting of the linearized inverse problem might be used to accelerate the convergence of iterative schemes for the nonlinear problem. These methods are known as full waveform inversion methods. The iterative schemes require solutions of large scale PDE problems at each iteration, accelerating the convergence of these methods cuts down their time/resources consumption.

In this section we take a step back to the original nonlinear inverse problem. I point out how the pseudodifferential scaling method developed for the linearized

inverse problem accelerates the convergence of the optimization for the model m. I present the method in the case of one parameter inversion, the generalization to multi-parameters follows suit.

I developed this work with the help of Dr Fuchun Gao in the setting of his FWI code.

One way to recover the model m without linearizing around an a-priori known background model is trying to optimize for m by trying to fit the data through the nonlinear forward map. The objective function is given by

$$J = \frac{1}{2} \left\| S[m] - d \right\|^2 . \tag{5.8}$$

The gradient of the objective function is,

$$g = F^*(S[m] - d). \tag{5.9}$$

And the Hessian is,

$$H = F^*F + \frac{\partial F^*}{\partial m}(S[m] - d). \tag{5.10}$$

Newton's method to minimize the objective function in an effort to recover m will have the following updates:

$$m_{k+1} = m_k - H^{-1}g. \tag{5.11}$$

Gauss-Newton's method neglects the second term in equation (5.10), $H \approx F^*F$ under assumptions of small residual or mild nonlinearity. Even after this approximation, it is too expensive to invert the Hessian. The approximation to the inverse of the normal operator obtained by the pseudodifferential scaling method serves as

a substitute for H^{-1}. The inverse of the Hessian is usually replaced by a constant and the Gauss-Newton method reduces to steepest descent with line search, the scaling factor preconditions the problem and introduces curvature information about the objective function and is expected to accelerate the convergence of the optimization.

Full waveform inversion does not split the model into a smooth and rough component. The approximation of good scaling factor however relies on the fact that the velocity field is split into these two parts. We therefore proceed at each step with splitting the velocity,

$$m_k = m_{k0} + \delta m_k.$$

Apply the normal operator on the rough part to obtain a scaling factor c,

$$c = \underset{c \in \Psi DO}{\operatorname{argmin}} \|\delta m_k - c\, F^*[m_{k0}] F[m_{k0}] \delta m_k\|^2. \tag{5.12}$$

And use the scaling factor thus derived to approximate the inverse of the Hessian, $H^{-1} \approx c$:

$$m_{k+1} = m_k - \alpha_k c\, g. \tag{5.13}$$

where α_k is a line search parameter to ensure decrease in the objective function.

This development is independent of the space dimension.

The justification for using the scaling factor as an approximation to the inverse Hessian is somewhat a posteriori. The approximation of the scaling factor requires the application of the normal operator and is thus comparable to the expense of one iteration of the FWI. The use of the scaling factor is justified if the overall

number of applications of the modeling and migration operators is decreased when using the scaling factor as opposed to simple steepest descent. Also, the scaling factor is dependent on the smooth part of the model, and we may be able to skip its approximation at each step if the smooth part remains constant. This last point constitutes part of the experiments that we need to conduct as part of the future work.

Also, the scaling factor is an approximation of the inverse only in the part of the data that is explained by the linearized theory (no multiples, ...). It remains to be seen how well of an approximation it is on the data that cannot be explained by the linearized part exclusively.

Herrmann et al. (2008a) applies part of this program on the linearized least squares problem with satisfactory results on the acceleration of the inversion process. Jang et al. (2009) also use the scaling factor they derive from the method of virtual sources to precondition full waveform inversion and accelerate the conversion of the nonlinear optimization problem.

5.4 Variable density acoustics: $p = 2$

I presented preliminary results on variable density acoustics for a layered model. The method presented in this manuscript succeeds at separating the effects of different physical parameters from the migrated images, and corrects for the amplitudes afterwards: a successful approximate inversion. The simple model was chosen to gauge the

success of the different steps of the method and an easier interpretation of the results. I propose to apply the method to more complex models, exhibiting multi-dips events. One possible application would be a variable density acoustics Marmousi model.

Dealing with large models requires parallelization of the finite difference code that is used to apply the normal operator (modeling, migration). This endeavor constitutes part of my future work.

I also propose to extend the conditioning study of presented for this case, to derive approximate indicators about the local condition number of normal operator. Alternatively, approximate projectors of seismic images onto the eigenvector of the normal operator with the largest eigenvalue provide an index to which physical parameters are better conditioned for recovery. I will also attempt to formulate such projectors.

5.5 Three-parameter case: $p = 3$

When formulating the method explicitly for more than two parameters, the algebra becomes a bit more involved, this is why we only show the case $p = 3$. However the procedure is generalizable to any p, albeit tediously.

$$N = \begin{pmatrix} N_{11} & N_{12} & N_{13} \\ N_{12} & N_{22} & N_{23} \\ N_{13} & N_{23} & N_{33} \end{pmatrix}.$$

Its adjugate is,

$$Adj(N) = \begin{pmatrix} (N_{22}N_{33} - N_{23}^2) & -(N_{12}N_{33} - N_{13}N23) & (N_{12}N_{23} - N_{13}N22) \\ -(N_{12}N_{33} - N_{23}N13) & (N_{11}N_{33} - N_{13}^2) & -(N_{11}N_{23} - N_{13}N12) \\ (N_{12}N_{23} - N_{22}N13) & -(N_{11}N_{23} - N_{13}N12) & (N_{11}N_{22} - N_{12}^2) \end{pmatrix}.$$

We will again introduce special notation that will facilitate subsequent manipulations so we can write $Adj(N) * b$ as a series of swap operations on the entries of b followed by applications of N. The entries of $Adj(N) * b$ are of the form:

$$N_{ij}Ni'j'b_k e_l = e_l^T e_i N e_j^T e_{i'} N e_{j'}^T e_k b := lijj'i'k. \tag{5.14}$$

Where again the adopted notation only shows the indices. Symmetry of N and approximate commutativity of its entries, allows us to deduce identities like $liji'j'k = ljii'j'k = lijj'i'k = li'j'ijk$. Using this notation and the previous identities we can write:

$$
\begin{aligned}
Adj(N) * b &= 122331 - 123231 + 123132 - 121332 + 121233 - 122133 \\
&\quad + 231231 - 233121 + 233112 - 231132 + 231123 - 232113 \\
&\quad + 312231 - 313221 + 313122 - 311232 + 311223 - 312123 \\
&\quad = (12)[2331 - 3231 + 3132 - 1332 + 1233 - 2133] \\
&\quad + (21)[3231 - 2331 + 1332 - 3132 + 2133 - 1233] \\
&\quad + (31)[2231 - 3221 + 3122 - 1232 + 1223 - 2123] \\
&\quad = (-12 + 21)[33(21 - 12) + 23(13 - 31) + 13(32 - 23)] \\
&\quad + (31)[22(31 - 13) + 32(12 - 21) + 12(23 - 32)]
\end{aligned} \tag{5.15}
$$

Equation (5.15) is interpreted as follows:

1. Form $N(e_2^T e_1 - e_1^T e_2)b$, $N(e_1^T e_3 - e_3^T e_1)b$ and $N(e_3^T e_2 - e_2^T e_3)b$
2. Form $(e_2^T e_1 - e_1^T e_2)N[e_3^T e_3 N(e_2^T e_1 - e_1^T e_2) + e_2^T e_3 N(e_1^T e_3 - e_3^T e_1) + e_1^T e_3 N(e_3^T e_2 - e_2^T e_3)]b$
3. Form $-(e_3^T e_1)N[e_3^T e_2 N(e_2^T e_1 - e_1^T e_2) + e_2^T e_2 N(e_1^T e_3 - e_3^T e_1) + e_1^T e_2 N(e_3^T e_2 - e_2^T e_3)]b$
4. Sum the last two images to obtain $Adj(N) * b \approx det(N) * x$

The above procedure amounts to 5 applications of the normal operator N, followed by one extra application to approximate $det(N)$. Bringing the total cost to 6 applications of N for $p = 3$.

I will implement the three parameter case to form the factors enumerated above. Also, I will investigate the possibility of reducing the number of required applications

of the normal operator, possibly in the setting of a particular problem like linear elasticity.

Bibliography

Aki, K. and Richards, P. (1980). *Quantitative Seismology: Theory and Methods*. Freeman, San Francisco.

Bao, G. and Symes, W. (1996). Computation of pseudo-differential operators. *SIAM J. Sci. Comput.*, 17(2):416–429.

Beylkin, G. (1985). Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *Journal of Mathematical Physics*, 26:99–108.

Beylkin, G. and Burridge, R. (1990). Linearized inverse scattering problem of acoustics and elasticity. *Wave Motion*, 12:15–22.

Bourgeois, A., Jiang, B., and Lailly, P. (1989). Linearized inversion: A significant step beyond pre-stack migration. *Geophysics J. Int.*, 99:435–445.

Charara, M., Barnes, C., and Tarantola, A. (1996). Constrained waveform inversion of seismic well data. In Jacobsen, B., Mosegaard, K., and Sibani, P., editors, *Inverse*

Methods, volume 63 of *Lecture Notes in Earth Sciences*, pages 98–112. Springer Berlin / Heidelberg. 10.1007/BFb0011767.

Claerbout, J. and Nichols, D. (1994). Spectral preconditioning. Technical Report 82, Stanford Exploration Project, Stanford University, Stanford, California, USA.

Foss, S.-K., de Hoop, M. V., and Ursin, B. (June 2005). Linearized 2.5-dimensional parameter imaging inversion in anisotropic elastic media. *Geophysical Journal International*, 161:722–738(17).

Guitton, A. (2004). Amplitude and kinematic corrections of migrated images for nonunitary imaging operators. *Geophysics*, 69:1017–1024.

Herrmann, F., Brown, C., Erlangga, Y., and Moghaddam, P. (2008a). Curvelet-based migration preconditioning. Technical Report 7, The University of British Columbia.

Herrmann, F., Moghaddam, P., and Stolk, C. (2008b). Sparsity- and continuity-promoting seismic image recovery with curvelet frames. *Applied and Computational Harmonic Analysis*, 24:150–173.

Jang, U., Min, D., and Shin, C. (January 2009). Comparison of scaling methods for waveform inversion. *Geophysical Prospecting*, 57:49–59(11).

Lörtzer, G. J. M. and Berkhout, A. J. (1989). Linear avo inversion of multicomponent seismic data. *SEG Technical Program Expanded Abstracts*, 8(1):967–972.

Minkoff, S. E. (1995). *Multiparameter Inversion and Energy Source Estimation for a Reflection Seismic Experiment*. PhD thesis, Rice University, Houston, Texas, USA.

Rakesh (1988). A linearized inverse problem for the wave equation. *Communications on Partial Differential Equations*, 13(5):573–601.

Rickett, J. E. (2003). Illumination-based normalization for wave-equation depth migration. *Geophysics*, 68:1371–1379.

Rutherford, S. R. and Williams, R. H. (1989). Amplitude-versus-offset variations in gas sands. *Geophysics*, 54:680–688.

Santosa, F. and Symes, W. W. (1988). High-frequency perturbational analysis of the surface point-source response of a layered fluid. *Journal of Computational Physics*, 74:318–381.

Shuey, R. T. (1985). A simplification of zoeppritz equations. *Geophysics*, 50:609–614.

Stolk, C. (2000). *On the modeling and inversion of seismic data*. PhD thesis, Universiteit Utrecht.

Symes, W. (1998). Mathematics of reflection seismology.

Symes, W. W. (2008). Approximate linearized inversion by optimal scaling of prestack depth migration. *Geophysics*, 73:R23–R35.

Tarantola, A. (1987). *Inverse Problem Theory*. Elsevier.

Taylor, M. (1981). *Pseudodifferential Operators*. Princeton University Press, Princeton, New Jersey.

Versteeg, R. and Grau, G. (1991). Practical aspects of inversion: The Marmousi experience. Proceedings of the EAEG, The Hague.

Virieux, J., Jin, S., Madariaga, R., and Lambaré, G. (1992). Two dimensional asymptotic iterative elastic inversion. *Geophys. J. Int.*, 108:575–588.