

Delft University of Technology
Master's thesis

Discrete frequency seismology

Adriaan VISSER 4202724

Start date: 1st January 2018
End date: 20th December 2018

Supervisors:

Prof Dr Ir R F Hanssen
Dr Ir D J Verschuur
P Dheenathayalan MSc

Additional committee members:

Dr Ir G Blacquière
Dr J Kalkman

This thesis is written as partial fulfilment for obtaining the degree of Master of Science
in Applied Earth Sciences and the degree of Master of Science in Applied Physics
Courses: AESM2640, AP3902
Credits: 60 ECTS



Abstract

This work determines whether the amount of frequency components present in the data can be reduced, whilst still retaining image quality, whereas most efforts in seismological research are done in reducing spatial sampling. It is shown using a PCA on the frequency spectra of several data sets that indeed a large redundancy in frequency content is present in onshore seismic data, and an attempt is made to generate a distribution of frequencies in order of importance.

Given this redundancy in the frequency spectrum of onshore seismic data, it has been attempted to reconstruct the missing frequencies by applying the Fourier transformation iteratively to the data. However, this transform does not take spatial sampling into account, which is aimed at to compensate for the missing frequencies. Therefore it has been elected to use a linear Radon transformation instead, which keeps components which are connected in space-time connected in the transform domain. A CGNE scheme has been set up to reconstruct the data, which performs very well along the almost linear asymptots in the shot records, up to a reduction of 70% of frequency components. This scheme iteratively applies the linear Radon transform to a shot record, weighing the data in the transform domain with an amplitude based norm. The energy that was spread out due to aliasing because of the missing frequencies is refocused to the main reflectors, especially along the asymptots of the reflection hyperbola. Missing frequencies are reconstructed, up to a scaling factor, and band gaps of up to 6Hz get filled in very well.

Next, it is attempted in this work to give quantitative quality metrics, to make comparison between seismic images easier and based on data, rather than subjective visual inspection. Treating migration as a black box, several quality metrics have been devised for the migrated sections: correlation to the ground truth, contrast within an image, average length of found lines, and local SNR. Contrast is not a very good metric to compare between images as its average across an image is almost constant with reduction percentage. The other parameters are good metrics and show a clear trend that the fewer frequency components present in the shot records, the worse the quality of the final image. An increase in deterioration of image quality is observed around 70% reduction, which is in correspondence with the earlier found value for the shot records.

Contents

1	Introduction	1
1.1	Research question	3
1.2	Thesis outline	4
2	Determination of redundancy in the frequency spectrum	5
2.1	Data description	5
2.2	Data preparation	8
2.3	Results	8
2.4	Discussion	9
3	Fourier reconstruction of missing spatial samples	13
3.1	Reconstruction in the Fourier domain	14
3.1.1	One dimension, regular grid	14
3.1.2	Conjugate gradient scheme	15
3.1.3	Irregular grid	15
3.1.4	2D Fourier transform	16
3.2	Results	17
3.3	Discussion	18
4	Radon reconstruction of missing frequencies	21
4.1	Radon Transform	21
4.2	Mathematical problem statement	23
4.3	Iterative solvers	24
4.4	Model	27
4.5	Results	28
4.6	Discussion	34
5	Quality description and results	37
5.1	Quality metrics	37
5.2	Results	40
6	Conclusion, discussion and recommendations	45
6.1	Answers to research questions	45
6.2	Recommendations	46
	Appendices	53
A	Additional tables and plots	54
A.1	Yilmaz data	54
A.2	Additional plots	56

Chapter 1

Introduction

The current standard of onshore seismic acquisitions for characterisation of oil and gas fields, as well as geothermal reservoirs, is to plant tens of thousands of geophones in the ground, which are usually analog devices that measure ground motion (Tellier and Lainé [2017]). This motion is induced by a sound wave, generated by either exploding a piece of dynamite or vibrating a heavy truck, where such an experiment is called a ‘shot’. Reflections of these waves on geological layer boundaries in the Earth reflect in all directions, and the part reaching the surface is measured by the geophones. Usually multiple shooting crews are used, so that when one group shoots the other(s) can move into position for the next shot, minimising downtime (Vermeer [1998]). Shooting is done at many positions, in order to optimise the areal coverage of the data (Meunier [2011]). The average area which is being explored in a 3D onshore survey is 1137km² (Visser [2017]), see table 1.1 for source and receiver parameters. This acquisition method is very labour intensive and generates a very large amount of data. If the same knowledge can be inferred from less data, the cost of a survey could go down, as well as increasing safety by requiring less people in the field.

Whilst geophones themselves are analog devices, the limited time sampling rate and record length restrict the number of available frequencies. The maximum frequency that can be measured is $(2\Delta t)^{-1}$, where Δt is the time sampling of the signal and the factor of 2 arises from the Nyquist criterion. A 2ms sampling results in a maximum measurable frequency of 250Hz, which extends well beyond the seismic bandwidth, which is approximately 5 to 100Hz (Society of Exploration Geophysicists [2018]). Given a 6s signal length, this would contain $6/(0.002) = 3000$ unique frequencies. Half of these are positive and half are negative, and the value at a certain positive frequency is the complex conjugate of its negative counterpart, $\omega_k = \omega_{-k}^*$, such that the amount of independent frequencies is half the total.

Seismic data in a 3D world is measured in five dimensions, (x_s, y_s, x_r, y_r, t) , i.e. source/receiver location and time. Dense sampling in 3.5 of those dimensions is most commonly used (Dr Verschuur, pers. comm.). Usually these are time and either (x_r, y_s) , or (y_r, x_s) , as in the cross-spread format (fig. 1.1)(Vermeer [2002]), with the other source and receiver dimension being sparsely sampled, making up the last half dimension. This format has a high inline density, and a low crossline density, with the lines of sources and receivers perpendicular to one another. The three dense dimensions are thus

Table 1.1: Average acquisition sizes as found by Visser [2017]

	Average [m]
Inline receiver separation	39
Crossline receiver separation	223
Inline source separation	41
Crossline source separation	145

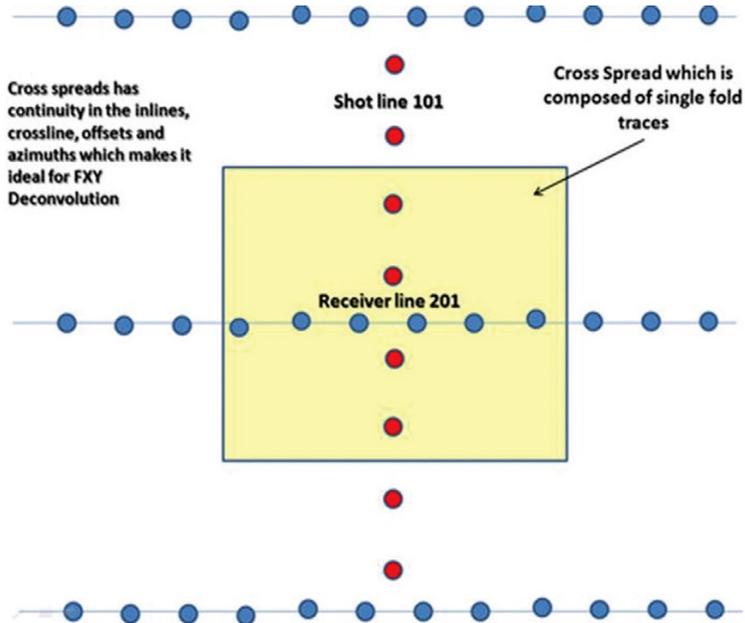


Figure 1.1: Red dots denote shot locations, blue ones show receiver position. From this figure it is clear that each shot has an (x, y) location, as well as each receiver. Measurement is amplitude during the time on each receiver location, thus the total amount of dimensions in an onshore acquisition is 5. Figure copied from [Schulte and Manthei \[2014\]](#).

inline source direction, inline receiver direction and time. The other two dimensions, crossline source and receiver direction, are sampled less dense (tab. 1.1). Efforts are made to reduce spatial sampling in a smart way, through (semi-) random positioning of sources and receivers. If it is possible to increase spatial coverage of receivers tremendously without increasing cost, the three densely sampled dimensions would be 2D receiver location and time, thus the amount of shots could possibly go down.

Attempts to reduce the amount of required data have been made using compressive sensing methods ([Romberg \[2009\]](#)), where the shots and receivers are randomly placed, in order to suppress aliasing of the signal with respect to using a regular grid. Processing data acquired in this manner uses a mathematical technique called ‘sparse inversion’. Data reduction is usually done for one of two reasons: faster processing due to redundancy in the data, or a trade-off with quality due to sparse sampling. In this thesis the latter will be looked at; it is attempted to retain image quality with less data.

Sparse inversion techniques have been around since at least the 1980’s (e.g. [Scales et al. \[1988\]](#)). These focus on sparsity in a transform domain to get high quality images with very few data available; this thesis will focus on sparseness in the frequency domain, i.e. reduction in Fourier transformed time. This sparseness can be enforced by changing the acquisition method, for example by emitting a single frequency at the source instead of a sweep over a range of frequencies, or in processing by removing frequencies to reduce calculation time. Sparse inversion can be employed in the data domain or on the final image. Reconstruction in the image domain is possibly faster, since only a single image needs to be processed. However, performing reconstruction already in the data domain will likely restore more of the missing information, resulting in a better final image.

Emitting single-frequency signals can possibly use lower instantaneous energy than in conventional seismology, since signals do not have to be pulsed and can thus be emitted at low energy, but over a long time, keeping total energy constant. This would make it safer to do in e.g. built-up areas, where high instantaneous energy can be detrimental to the structural integrity of buildings. The low-intensity

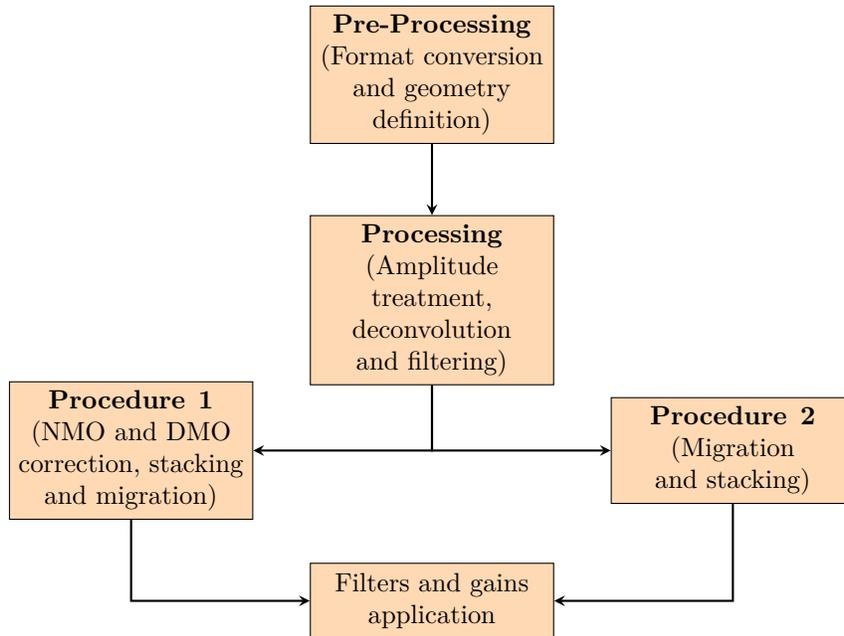


Figure 1.2: Flowchart of the seismic processing chain. Adapted from de Lucena Rêgo et al. [2017]. The discussed sparse inversion would enter the chain after the Processing step, before either Procedure.

signal would need to be emitted for a longer time than the usual 15s upsweep of a vibrating truck to reach the required signal-to-noise ratio (SNR). Instead of measuring the range of discrete frequencies as given by a geophone, it is proposed to reduce the number of frequency components present. This opens up possibilities for new acquisition methods that can handle only a few frequency components at the same time, but are cheap in deployment and usage, requiring less manual handling in the field, and have dense areal coverage.

The seismic processing chain starts with pre-processing of the raw data as obtained from the geophones, such that all meta-data like shot and receiver locations, trace numbers etc are correct (fig. 1.2). Then initial filtering is done to e.g. equalise amplitudes and remove ground roll and multiple reflections. Then the data can be filtered further using Normal Move Out (NMO) or Dip Move Out (DMO), and either stacked and then migrated, or migrated and subsequently stacked. Stacking is done by adding traces in a so-called common midpoint (CMP) gather, which is a data format where each trace has the same midpoint between source and receiver. Migration is an imaging technique which attempts to migrate all the reflections as obtained in a shot record to their correct reflection location in the subsurface image, i.e. this step is the one that builds the final image of the subsurface. Finally some image processing filters such as gain manipulation can be performed, but those are skipped in this thesis. The sparse inversion would enter this chain after the initial filtering, before stacking or migration.

Finally, quantification of image quality is required for a non-subjective comparison between methods. In practice a qualitative comparison is often used by placing two images side-by-side, one using the old processing methods and one using the discussed technique, or calculate the relative error between the obtained and desired result. Qualitative statements such as ‘the reflector at depth x has a clear higher contrast using the new method’ are of commonly used (e.g. (Kumar [2015]) paragraph 2.9). A more quantitative measure of image quality is thus needed.

1.1 Research question

The main question of this thesis is:

Is it possible to reduce the number of required frequencies in onshore seismology by increasing the spatial density of the measurements?

The reason for this question is the potential of new acquisition methods based on sparse frequency collection, which has not yet been researched in-depth for seismic purposes.

In order to know whether it is possible to reduce the amount of required frequencies, it is necessary to know whether there is redundancy in the frequency spectrum of onshore seismological data. If there is, knowing which frequencies carry the main load of information can help in selecting the optimal frequencies to retain.

If indeed redundancy can be found in the data, and frequencies are removed from a shot, the full frequency spectrum of the data needs to be reconstructed in order to use the present seismic processing chain, see figure 1.2, or [Verschuur \[2006\]](#) for a full treaty.

Reconstruction is usually done in a transform domain, therefore it has to be investigated which transformation will be best to use for the given case. Using the best transform we would like to know how well the data can be reconstructed.

To answer this the following subquestions are posed:

1. Is there redundancy in the frequency domain of the data? Thus, can less frequencies be used whilst still obtaining a ‘good’ result?
2. What frequency sampling is best if only a limited number of frequencies are to be measured?
3. Which transformation yields the best results for data with missing frequency components?
4. How well can we reconstruct the dense, complete data set from the sparse frequency data to allow usage of the currently available processing chain?
5. What metrics can be used to determine image quality? Thus what is a ‘good’ result in quantitative measures to be used under question 1.

1.2 Thesis outline

Before an attempt is made at reconstructing data with missing frequencies, it is necessary to determine whether there is redundancy in the data (chapter 2). After this has been established, frequencies are removed from shot records and an attempt is made to use the Fourier transformation to reconstruct the missing frequencies (chapter 3). The Fourier transform does not take spatial continuity of reflection events into account, thus instead the linear Radon transformation is used to reconstruct the aliased shot records (chapter 4). The reconstructed shot records are migrated to a depth section, for which several quality descriptors are presented (chapter 5). Finally the research subquestions as posed above are answered and recommendations for further research are made (chapter 6).

Chapter 2

Determination of redundancy in the frequency spectrum

Before starting to remove frequencies from existing data and attempting reconstructing those missing frequencies, it is useful to know whether there is any redundancy in the frequency spectrum of seismic data. There are various techniques to determine whether a data set has redundancy, such as singular value decomposition or wavelet analysis. The choice has been made to use Principle Component Analysis (PCA), since that is a well-known general mathematical tool for dimensionality reduction of a data set. It is implemented in most programming languages, and is it fast to use. In a set consisting of N dimensions it finds the direction of maximum variance, and assigns that to the first principal component (PC). The second largest variance, which has to be orthogonal to the first one, is assigned as the second PC and so on until as many PCs are found as there are dimensions, explaining the total variance in the data (Lay et al. [2016]).

If only a few PCs explain most of the total variance in the data, there is redundancy in the data, and performing sparse reconstruction can possibly help reduce the amount of required data to represent the same knowledge. If onshore seismic data has redundancy in its frequency domain it is likely that choosing frequencies based on those components will give a good representation of the data with only few frequencies.

In the book of Yilmaz et al. [1987], 40 shot records are presented, which have been collected by various companies all around the world, both on- and offshore (see table A.1). This is considered to be a standard, open data set in seismological research; hence it has been elected to use here as well.

First the data set will be discussed (sec. 2.1) and the methodology for data preparation is presented (sec. 2.2). Then the results of PCA will be presented (sec. 2.3), finally the results are compared to a real world data set, as well as a synthetic one (sec. 2.4).

2.1 Data description

All shot records from the Yilmaz data have been Fourier transformed to the (f, x) domain, and each panel has been averaged over all traces to obtain a single-sided Fourier spectrum. Since all shot records have a different number of time samples, rate of temporal sampling, amplitude scaling and amount of traces the vertical axis label, amplitude, has no relevant meaning other than showing the relative weight of frequencies with respect to one another per shot record (fig. 2.2).

In some panels, for instance 2 and 4 (fig. 2.1), it can be observed that above a certain frequency the data looks very much like that of a smooth filter (fig. 2.2). In others, such as 5, nodes reminiscent of a filter are present, but less obvious. Overall it can be said that the signals are band limited and usually have a high frequency cut-off somewhere between 60 and 100Hz. Shots 2 and 4 appear to be shot with a vibrating source, which is operated in a limited frequency range in the sweep, whereas shot 5 looks

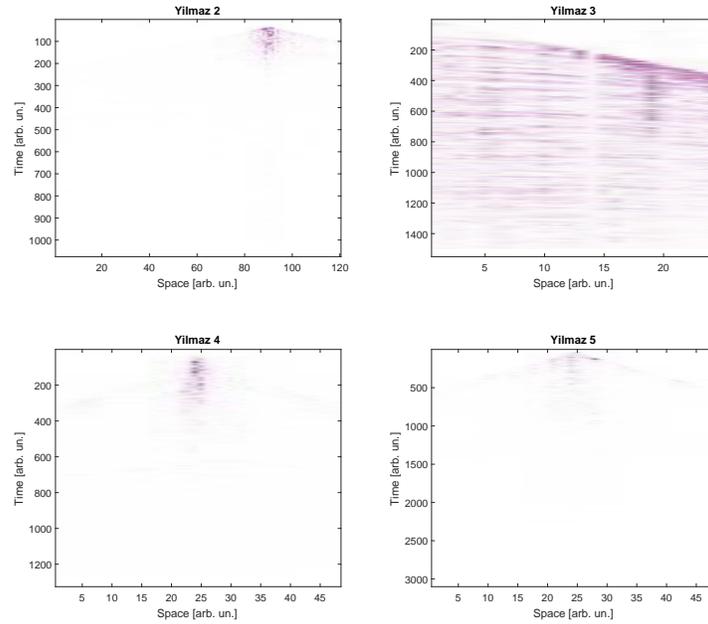


Figure 2.1: Unclipped shot records for the data from [Yilmaz et al. \[1987\]](#), the number on top of each panel corresponds to the one in Yilmaz' book. Since all shot records have a different number of time samples, rate of temporal sampling, amplitude scaling and amount of traces, the colour scale is not the same for all panels.

to have been shot with an airgun, which is characterised by a natural decline of energy in the high frequencies.

Some shots show a sharp notch at 60Hz, which is an indication that the data has already undergone some processing, as this is a tell-tale sign that resonance of the electrical grid at 60Hz has been removed. In others a large peak is visible around 180Hz, which can be a higher harmonic of the said electrical grid resonance, which possibly was not removed because it lies far outside the bandwidth of interest (c.f. [fig. 2.3](#)).

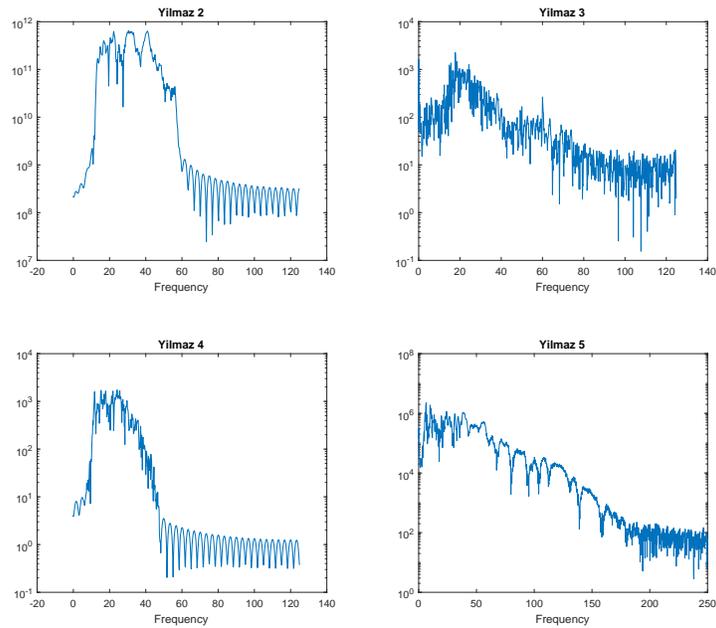


Figure 2.2: Logarithmic single-sided Fourier spectra of the shot records given in fig. 2.1. The spectra were obtained by horizontally averaging the Fourier transform of each panel. Note that the frequency spacing is different for each panel, and the maximum frequency in set 5 is 250Hz, as opposed to 125Hz found in sets 2, 3, and 4, since this shot has been sampled at 2ms instead of 4ms. One can clearly see the band limitation of onshore data shot using a vibroseis (2 and 4), and the slowly decaying amplitude of airgun data (3 and 5).

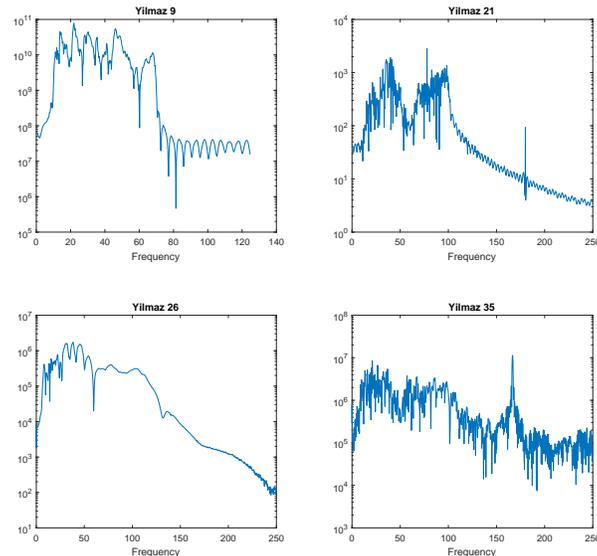


Figure 2.3: Logarithmic single-sided Fourier spectra of several shot records. A sharp dip in amplitude can be observed at 60Hz in sets 9 and 26, which is probably due to removal of the electrical grid resonance. In sets 21 and 35 a peak can be observed at 180Hz, which is likely a higher harmonic of the electrical grid resistance, which has not been removed since it is far outside the bandwidth of interest.

2.2 Data preparation

Since this thesis is concerned with onshore seismology, only the shots which have been made on land are included in this section. Each shot has a time sampling of either 2 or 4ms, and a different total measurement time, number of traces and amplitude scaling; so, first they are normalised. Each shot is bandpass filtered to a range of 0-90Hz, and subsampled to 4ms, meaning that in shots having 2ms sampling every second sample is thrown away, and subsequently zero-padded to the maximum time extent found in the data (about 8s). The shots then have their amplitudes normalised to $[-0.5, 0.5]$, centring the data around 0 so as to remove the DC component. Each shot record is Fourier transformed from (x, t) to (x, f) . Since each shot has been normalised, all shots will now contain samples on the same frequency components.

The full data contains 2175 frequency components, thus $(2175 - 1)/2 = 1089$ positive frequencies, and there are 1893 traces in total. The data matrix of 1893×1089 is ill-conditioned, therefore it has been decided to perform an interpolation in the frequency domain to half-Hz. Because 4ms sampling has been used in the data, thus recorded at 250Hz, it has a Nyquist frequency of 125Hz. This results in a 1893×251 matrix, since there are 251 frequency components in $[0, 0.5, 1, \dots, 124.5, 125]$ Hz. This matrix is not ill-conditioned and PCA can be performed on it.

Each principal component found in the PCA is an N dimensional vector, where N is the number of frequencies used, 251 in this case. It is thus difficult to tell which frequency component is actually the most important one. A method has been devised to decide which are the most important components:

- Each Principal Component (PC) has a given explained variance (see the first percentage on the right of e.g. fig. 2.4). This serves as input to the algorithm, together with N , the number of desired frequencies as output. Given these two inputs, each $\frac{1}{N}100\%$ of explained variance will determine one frequency to output.
- If $N = 100$, each 1% of explained variance will give 1 frequency to keep. Looking at the top 5 PCs of the Yilmaz data set, we see that the first PC will determine 56 frequencies to keep, the second 11, the third 7 etc., until a 100 frequencies will be determined.
- The first PC now gives the required amount of frequencies, based on absolute amplitude of each frequency component within the PC.
- Every next PC has the required amount of frequencies selected based on amplitude again, after which that list will be checked for duplicates with those already present in the output. E.g. if 5Hz is the top component of PC2, but is already given by PC1, it is discarded, and the next highest frequency in amplitude is selected. This continues until the required number of frequencies per PC has been selected.
- At the end, a list of N unique frequencies will be obtained; these are considered to be the most important ones.

2.3 Results

Figure 2.4 shows the first five principal components, which together explain almost 80% of the total variance in the data. The 5-15Hz component can be explained by ground roll. Some 90% of the shot energy goes into ground roll, which propagates at a low velocity and frequency ([Schlumberger Oilfield Glossary \[2018\]](#)). Furthermore, this shows that at least 80% of the energy is contained in frequencies below 60Hz, which is just halfway the seismic bandwidth. This is supported by looking at the second panel of figure 2.8, where it is clear that most energy is concentrated between 5 and 60Hz. Given that so much of the total variance is explained by so few components, redundancy can be assumed for this data set. Since the used set is widely considered as an open standard, this conclusion will be extended to onshore seismic data in general.

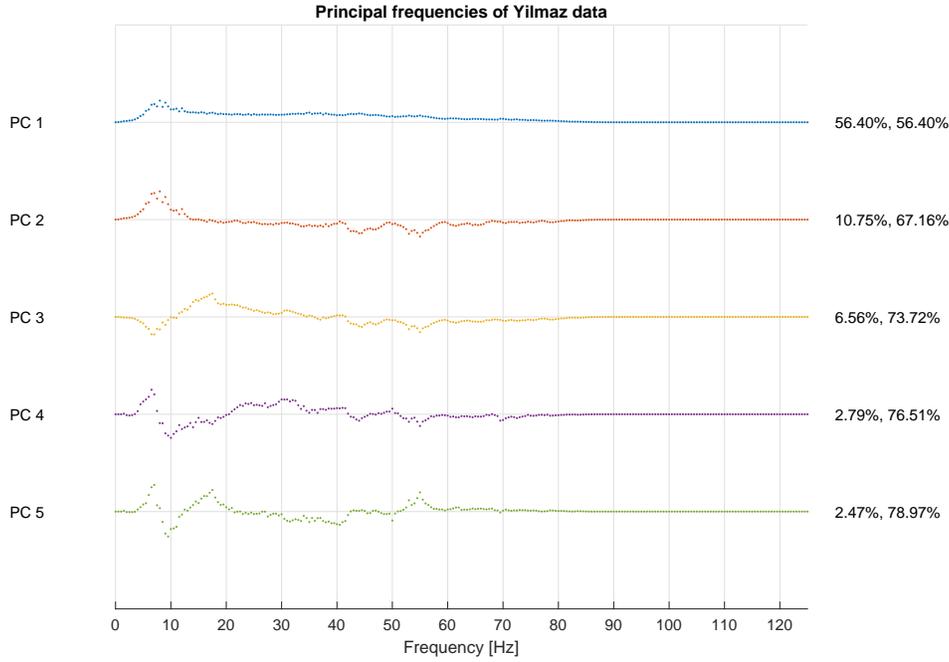


Figure 2.4: Top five principal components of the frequencies contained in the Yilmaz dataset. Each line is normalised to $[-1, 1]$. The components are plotted in decreasing order of importance, the first percentage mentioned is the percentage of data explained by that component, the second percentage is the cumulative sum of explained percentages by that component and all above.

2.4 Discussion

The results given in figure 2.4 are compared to an onshore data set from Poland shot with a vibrating source ([Geofizyka Torun S.A, Poland \[2005\]](#)), containing 251 shots with 284 traces each (fig. 2.5). Receiver separation is 25m, time sampling is 500Hz, with a total record length of 3s. The differences between the results of PCA on the Yilmaz data and this set show immediately the characteristics of dynamite and vibroseis as a source (fig. 2.6). Some of the shots in the Yilmaz data have been made with a dynamite source, localising a lot of energy at low frequencies, whereas the data from Poland have been shot with a vibroseis, equally spreading the energy over the full bandwidth.

The same has been done for the synthetic data, which will be introduced in section 4.4. It can be seen that the PCA of this synthetic data, fig. 2.7 is very clean and are very smooth. The data from Poland also has this same structure, albeit less smooth and with more noise, but this is as expected since the noise component of the data has not yet been removed. Even the principal components of the Yilmaz data are somewhat reminiscent of having the same shape, but even worse than those of the Poland set. Presumably this is because some shots have used dynamite instead of a vibrating source, as well as each shot being done by a different company in a different location. Since the shots are old, equipment was probably less sensitive than nowadays and more prone to noise.

For the synthetic model data have been reduced to contain only the 20% most important frequencies as given by its corresponding PCA (fig. 2.7). As can be seen in figure 2.8 more than half of the frequency bandwidth of the model, 60 to 125Hz, already contains little to no data, so it is easily understood that the central part of the 5-60Hz data lobe is retained, 14 to 37Hz with half-integer sampling in this case. The shot record after removing frequencies is slightly aliased, as can be seen around zero-offset in the rightmost panel near the apices of the hyperbola.

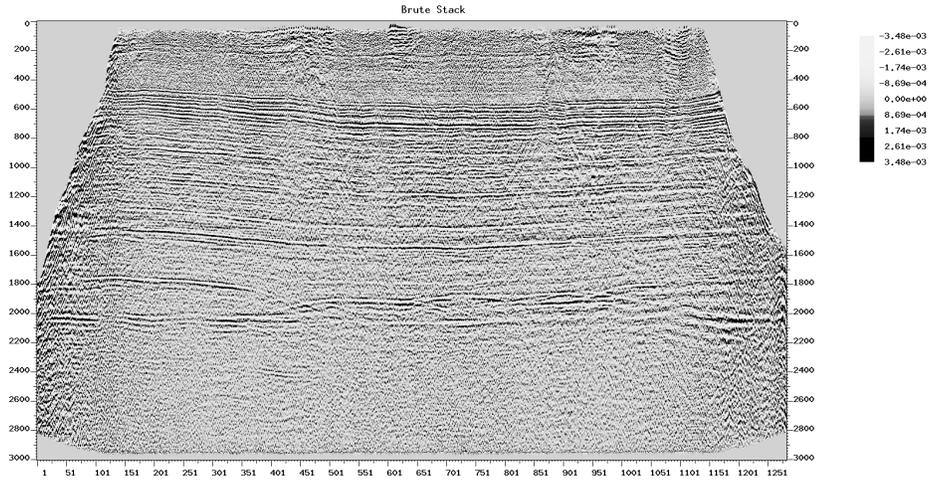


Figure 2.5: Brute stack of line 001 by [Geofizyka Torun S.A, Poland \[2005\]](#). Image obtained from the SEG Wiki.

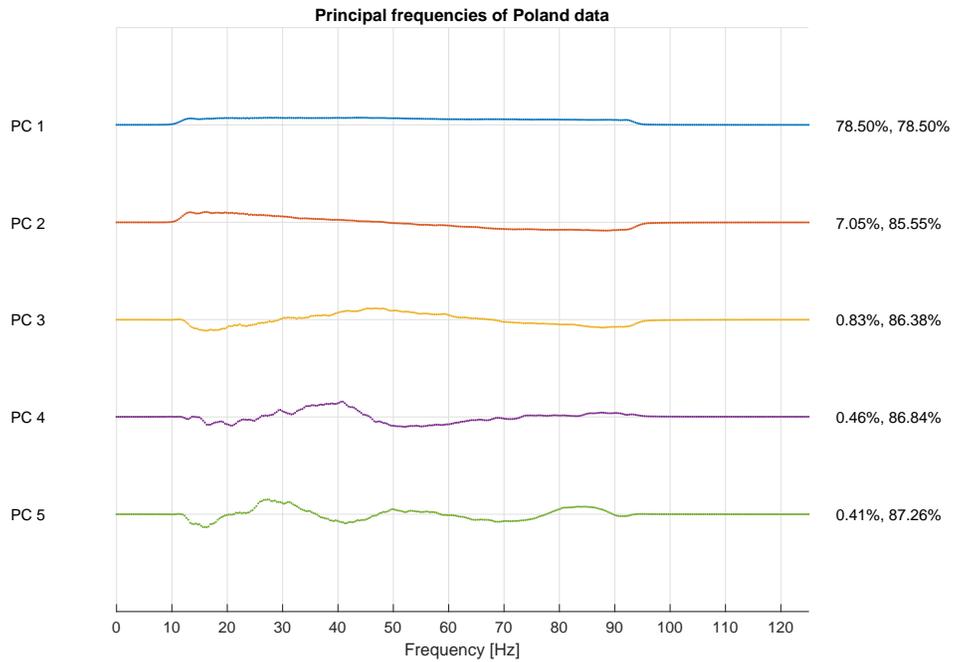


Figure 2.6: Top five principal components of the frequencies contained in a collection of 251 onshore shots in Poland. Each line is normalised to $[-1, 1]$. The components are plotted in decreasing order of importance, the first percentage mentioned is the percentage of data explained by that component, the second percentage is the cumulative sum of explained percentages by that component and all above.

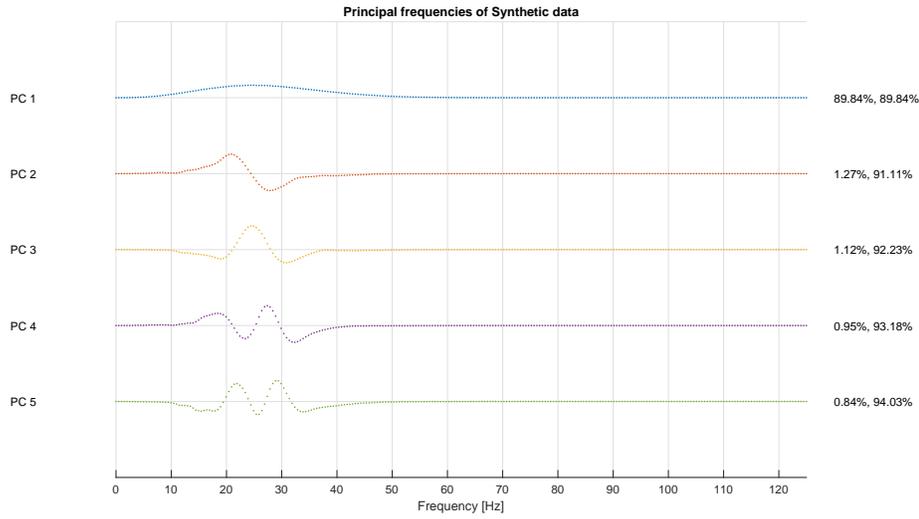


Figure 2.7: Top five principal components of the frequencies contained in a collection of 361 synthetic shots of the model presented in the next section. Each line is normalised to $[-1, 1]$. The components are plotted in decreasing order of importance, the first percentage mentioned is the percentage of data explained by that component, the second percentage is the cumulative sum of explained percentages by that component and all above. An example shot record of this set is presented in figure 4.2, left panel.

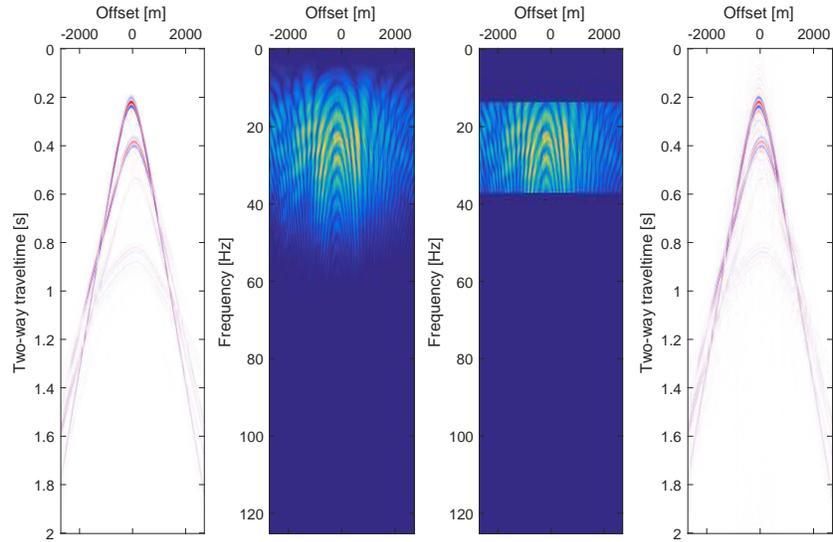


Figure 2.8: Shot record (left panel) with its associated Fourier transform (second panel). Only 20% of the frequencies have been kept, with the chosen frequencies given by the PCA (third panel), which shows a slightly aliased shot record (right panel).

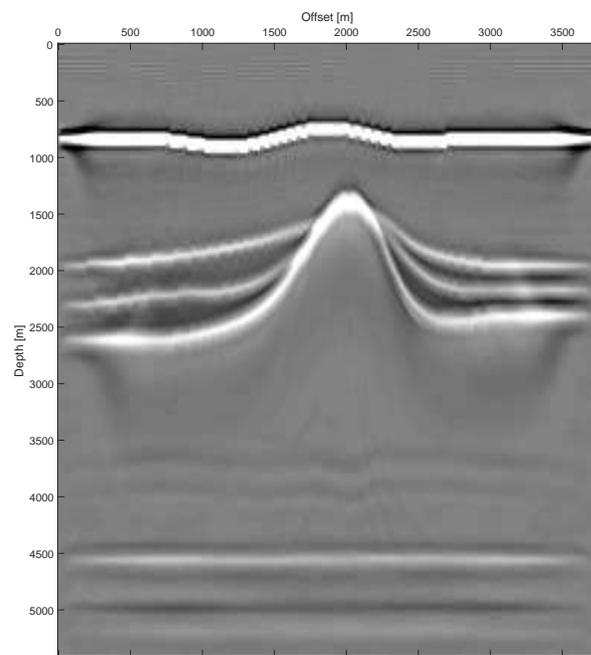


Figure 2.9: Migrated model when only the top 20% of frequencies as given by fig. 2.8 are retained. The synthetic model will be introduced in section 4.4; for now it suffices to say that all important features are still visible.

Chapter 3

Fourier reconstruction of missing spatial samples

Given the results in the previous chapter, it has been established that there is redundancy in the frequency spectrum of seismic data. Sparse reconstruction may thus help to retrieve information on the missing frequencies. Real-world data will always have defects, such as a broken receiver, missing data due to the need to go around buildings or natural obstacles etc. (fig. 3.1). Reconstruction can help to fill in these gaps or mistakes in the data.

There are three main groups of reconstruction techniques: based on transformation, filtering or wavefield operators (Zwartjes [2005]).

- Transformation based reconstruction uses the sparseness in the transform domain and finds the parameters which best fit the data under given constraints. Usually these constraints are concerned with noise reduction, energy focussing and making the output grid uniformly sampled. The parameters in the transform domain are iteratively obtained via least-squares inversion or a solver like Conjugate Gradients.
- Filter based methods try to interpolate over the missing or wrong data using an interpolating filter. Prediction-error filters can fill gaps in uniformly sampled data sets (Fomel [2000]).
- Wavefield operator based reconstruction use the Kirchhoff integral operator. In numerical implementation the integrals are replaced by summation, which relies on constructive and destructive interference. These integrals are flexible and allow for any sampling geometry, but artefacts are introduced when sampling is coarse or non-uniform, due to incomplete destructive interference.

This thesis will be restricted to transformation based methods. Transformations are easily implemented numerically, and can carry a wider range of constraints than a filter interpolation could encompass. Due to the redundancy being in the frequency domain, the first transformation which will be investigated is the Fourier transform. Zwartjes [2005] wrote a dissertation on this subject, which will be followed here. After introducing the Fourier transform and conjugate gradient scheme an attempt will be made to recover the results of Zwartjes.

Fourier reconstruction is geared towards reconstructing non-uniform spatial sampling, whereas the present case is non-uniform sampling in the frequency domain. After reproducing results from Zwartjes [2005], it will be attempted to alter the reconstruction to work in the frequency domain instead.

In this chapter the following mathematical notation will be used: non-bold letters denote scalars, e.g. α, β , bold lowercase letters denote vectors, e.g. \mathbf{p}, \mathbf{d} , and bold uppercase letters denote matrices, e.g. $\mathbf{A}, \mathbf{S}_{\odot}$.

First the mathematics of the Fourier transform will be discussed, as well as the conjugate gradient method (sec. 3.1), followed by results (sec. 3.2) and a discussion (sec. 3.3).

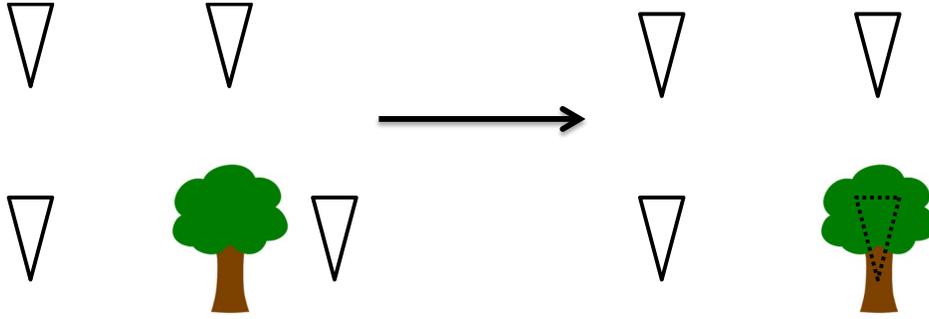


Figure 3.1: Illustration of the spatial sampling problem. The tree on the left blocks a receiver (triangles) from being properly aligned on the grid. Fourier reconstruction virtually moves the receiver in processing to the correctly aligned position, see right panel.

3.1 Reconstruction in the Fourier domain

Missing and irregular data can be reconstructed using an interpolation method in the Fourier domain. In this section first the general outline of the algorithm is given for the simple case: one dimension on a regular grid, which is subsequently extended to irregular grids and finally to two dimensions.

3.1.1 One dimension, regular grid

A discrete one dimensional signal is represented by $\mathbf{p}[x_n]$, which may be an amplitude, given in Volt, Pascal or any other unit, measured at locations x_n . First a regular grid is considered, i.e. the locations $n \in [0 \dots N - 1]$ are located on $n\Delta x$, where Δx is a constant.

We define a number of desired Fourier components, M , where $M = N$ for a regularly sampled case without missing samples. The wave number is given by k , thus the Fourier transformed signal is given by $\tilde{\mathbf{p}}[k_m]$, where in this case we use again a regular grid, thus $\Delta k = 2\pi/(N\Delta x)$ and $m \in [-M/2 \dots M/2 - 1]$ for M even and $m \in [-(M-1)/2 \dots (M-1)/2]$ for M odd.

We define two inverse Fourier transform matrices, \mathbf{F} and \mathbf{A} . \mathbf{A} transforms the data from Fourier domain to the measurement grid x_n ; \mathbf{F} transforms the found frequency components back to a regular grid. In the case where $N = M$ and both the input and output grid are equal, these two are equal matrices are equal.

$$F_{nm} = e^{-im\Delta k n\Delta x} \quad (3.1)$$

$$A_{nm} = \frac{\Delta k}{2\pi} e^{-im\Delta k x_n}. \quad (3.2)$$

For minimisation purposes we then define a weighting matrix \mathbf{W} , whose use will become apparent when the measurement grid is non-regular and/or samples are missing:

$$W_{nn} = N\Delta x_n, \quad (3.3)$$

where Δx_n in this case is constant.

When solving ill-posed problems, or to prevent overfitting of a solution, a mathematical problem is often constrained with a damping parameter. This is called regularisation. To regularise the CG scheme we introduce a regularisation parameter $\lambda = c^2/\sigma^2(\tilde{\mathbf{p}}_0)$, where c is a constant which can be set manually to control the amount of regularisation. $\sigma^2(\tilde{\mathbf{p}}_0)$ is the variance of the Fourier transform of the initial signal:

$$\tilde{\mathbf{p}}_0 = \mathbf{A}^H \mathbf{p}, \quad (3.4)$$

where H denotes the Hermitian conjugate.

3.1.2 Conjugate gradient scheme

We now present a CG scheme based on [Zwartjes \[2005\]](#), which is modified according to [Shewchuk \[1994\]](#). CG schemes are used to iteratively approximate an inverse solution, in case the inverse model itself is too large to be handled easily ([Shewchuk \[1994\]](#)). Another benefit of CG over direct inversion is that that would need to account for every little detail in the data, including all the noise, whereas CG minimises an objective function under some constraint. This thus allows for a ‘good enough’ solution, as opposed to the exact one. Noise can be discarded under the constraint, leaving a much cleaner model as opposed to what would be retrieved using direct inversion. Non-linear constraints cannot be inverted exactly, since their matrix representation is based on the data, thus changing the model each calculation step. The objective function we strive to solve is:

$$J = \frac{1}{c^2} \|\mathbf{W}^{1/2}(\mathbf{p} - \mathbf{A}\tilde{\mathbf{p}})\|_2^2, \quad (3.5)$$

which is equation 2.3.14 from [Zwartjes \[2005\]](#). c is some scalar which damps the function.

We initialise the scheme with the initial residual, and set the initial search direction equal to this:

$$\begin{aligned} \mathbf{r}_0 &= \mathbf{A}^H \mathbf{p} - (\mathbf{A}^H \mathbf{W} \mathbf{A} - \lambda \mathbf{I}) \tilde{\mathbf{p}}_0 \\ \mathbf{d}_0 &= \mathbf{r}_0; \\ i &= 1; \end{aligned}$$

where \mathbf{I} is the identity matrix, \mathbf{r}_i the residual at iteration i and \mathbf{d}_i the search direction at iteration i . The

Finally we need a maximum number of iterations and a threshold for the residual below which the algorithm stops.

The CG scheme iteratively steps towards solution with a step size α , then the new, reduced, residual is calculated. The next search direction must be orthogonal to the previous ones by definition; the new change of direction is given by β .

$$\begin{aligned} \text{while } i < i_{\max} \quad \& \quad \frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} > r_{\text{threshold}} \\ \alpha &= \frac{\mathbf{r}_k^H \mathbf{r}_k}{\mathbf{d}_k^H (\mathbf{A}^H \mathbf{W} \mathbf{A} + \lambda \mathbf{I}) \mathbf{d}_k} \\ \tilde{\mathbf{p}}_{k+1} &= \tilde{\mathbf{p}}_k + \alpha \mathbf{d}_k \\ \mathbf{r}_{k+1} &= \mathbf{r}_k - \alpha (\mathbf{A}^H \mathbf{W} \mathbf{A} + \lambda \mathbf{I}) \mathbf{d}_k \\ \beta &= \frac{\mathbf{r}_{k+1}^H \mathbf{r}_{k+1}}{\mathbf{r}_k^H \mathbf{r}_k} \\ \mathbf{d}_{k+1} &= \mathbf{r}_{k+1} + \beta \mathbf{d}_k \\ i &= i + 1. \end{aligned}$$

This scheme is iterated over until either the residual, $\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|}$, is below our pre-set threshold. This means the solution is good enough for our purposes. Or the scheme stops because the maximum number of allowed iterations has been reached.

3.1.3 Irregular grid

When the sampling grid is non-uniform, for example because of missing samples or varying positions, we can use different grids in the measurement and output domains with proper windowing of the Fourier coefficients. The final solution should be obtained on a regular grid, thus that is different from

the initial input grid. The same CG scheme as outlined in section 3.1.2 can then be used then to reconstruct the signal.

Let $\mathbf{p}[x_n]$ be our sampling vector, this time with $x_n - x_{n+1}$ not being constant. We also define a regular grid with constant Δx_r , with positions $x_{n,r}$, on which we want to have our interpolated result. Furthermore the number of Fourier coefficients, M , will not be equal to the number of measurement locations, N , but rather equal to fN_r , with N_r the number of points on the regular grid and f an oversampling factor. Δk is as given for the regular case, provided we use Δx_r ; m is equal to the definition above. \mathbf{A} is equal to the earlier use, \mathbf{F} uses the regular grid Δx_r . This makes sense as \mathbf{A}^H transforms our irregular grid measurement to the regular Fourier domain grid, and \mathbf{F} transforms from the Fourier domain to the regular output grid $x_{n,r}$. This means as well that \mathbf{A} and \mathbf{F} will no longer be square matrices. The weight matrix \mathbf{W} changes:

$$W_{nn} = N * \frac{N_r \Delta x_r}{\sum_n \Delta x_n} \Delta x_n \quad (3.6)$$

where Δx_n is given by $(x_{n+1} - x_{n-1})/2$. The division is to normalise the weights.

We use an oversampling factor $f = 2$, which is optimal according to [Duijndam and Schonewille \[1999\]](#). We might end up aliasing our Fourier transformed vector $\tilde{\mathbf{p}}[k_m]$; hence, we use a window on that to suppress this. The current implementation allows for the following window functions: Hamming, Blackman, Tukey, Exponential, Gaussian with $\sigma = 0.3$, and rectangular with edges $N_r/2$. The CG algorithm can then be used in the same way to solve for the Fourier transformed solution vector $\tilde{\mathbf{p}}[k_m]$, which can be transformed to the regular output domain via \mathbf{F} :

$$\mathbf{p}[x_{n,r}] = \mathbf{F}\tilde{\mathbf{p}}[k_m] \quad (3.7)$$

3.1.4 2D Fourier transform

The above algorithm can be extended from one dimension to multiple dimensions; the 2D case will be discussed here.

Consider a two dimensional signal $\mathbf{p}[x_n, y_n]$ sampled at n locations on an $[x, y]$ grid. The locations are not necessarily equally spaced in both directions, i.e. $\Delta x_n \neq \Delta y_n \neq \text{constant}$. The signal \mathbf{p} is stored in a 2D manner, whereas an array is required for further processing. This can be realised by storing the matrix in lexicographic, i.e. column major, order. This means a matrix \mathbf{A} will be stored as follows:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \rightarrow \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

This enables us to use an array for \mathbf{p} , which defines a 2D function given accompanying arrays for the x, y coordinates. Additionally we create a regular grid in the Fourier space with components $[m_x, m_y]$, which is also stored in lexicographic order, labelled with index m for the both of them, where the Fourier domain again is oversampled by a factor of f : $M_x = fN_x$, and $M_y = fN_y$.

The equations for \mathbf{F} and \mathbf{A} (eqs. 3.1, 3.2) can then be written for the two dimensional case:

$$F_{nm} = e^{-i[m_x \Delta k_x n \Delta x_r + m_y \Delta k_y n \Delta y_r]} \quad (3.8)$$

$$A_{nm} = \frac{\Delta k_x \Delta k_y}{4\pi^2} e^{-i[m_x \Delta k_x x_n + m_y \Delta k_y y_n]} \quad (3.9)$$

where each unique x, y pair is denoted by a single n and each unique m_x, m_y pair by a single m . The CG algorithm as described above can now be used with the multi-dimensional definitions.

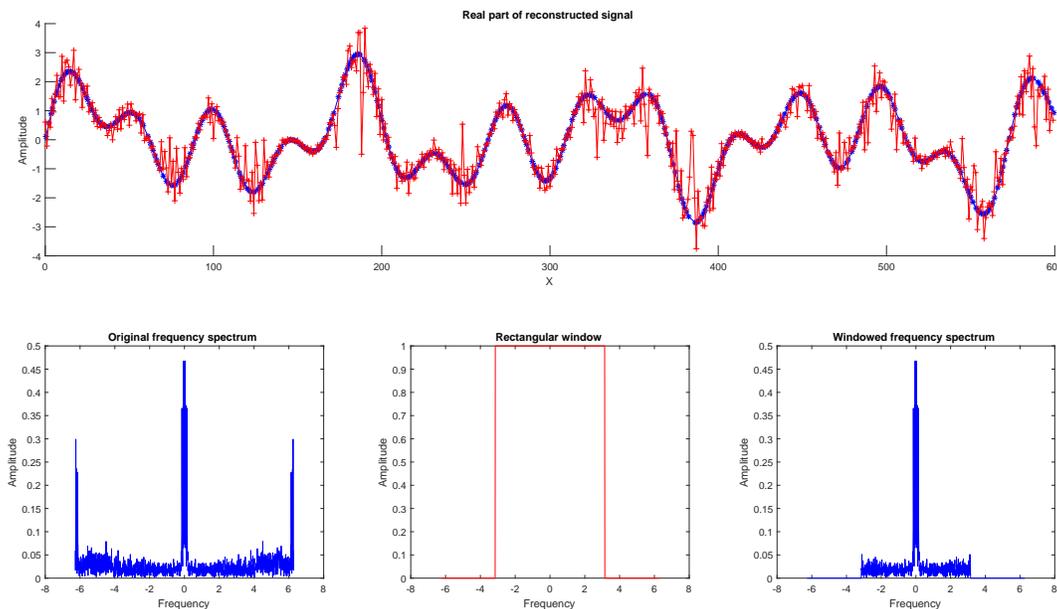


Figure 3.2: Top figure: original signal (blue), reconstructed signal (red) using the Fourier-reconstruction method with a rectangular window in the frequency domain. Bottom row left: unwindowed Fourier spectrum of the original signal, middle: rectangular window, right: the windowed Fourier spectrum.

3.2 Results

The algorithm presented in section 3.1 is tested by using a signal defined by:

$$\mathbf{p}[x_n] = \sin \frac{x_n}{7} + \sin \frac{x_n}{13} + \sin \frac{x_n}{23},$$

where x_n has been created as $\Delta x * n + \text{rand}$, where Δx is 1 initially, n runs from 0 to $N - 1$ where $N = 600$ and rand is a random number between 0 and 0.5, so as to make the location random. Next, 10 percent of the points have been randomly removed from the signal to create gaps. This means that N will be changed to 540 after removing the points. The output grid is desired to be regular, so $N_r = 600$.

The rectangular window as proposed by Zwartjes [2005] in paragraph 2.2.3, see figure 3.2, converges in 26 iterations and does a good job when there are no missing samples. When gaps in the initial data occur, however, the overshoot is quite large, and this cannot be suppressed by increasing the regularisation parameter λ . When looking at the Fourier spectrum it can be seen that there are still quite a few high frequency components in the data which constitute noise and aliasing, even after windowing.

Using an exponential window, see figure 3.3, the CG scheme also converges in 26 iterations, and does a better job than the rectangular window in the gaps, at the cost of under-estimating the local minima and maxima of the signal. Again the same Fourier spectrum can be observed on the lower left of the figure, with its central peak consisting of the three frequencies of the input signal, in the middle and aliasing at the far ends of the spectrum. Using a sharply peaked exponential window, high frequencies are strongly suppressed, generating a much smoother signal than when using the rectangular window. This performs well for the signal at hand, since only very low frequencies are present.

Other windows, such as the Tukey, Blackman, Hamming and Gaussian windows have also been implemented and show behaviour somewhere between these two extreme cases. The optimal window to be used for a given data set will depend on the frequency content of the original signal.

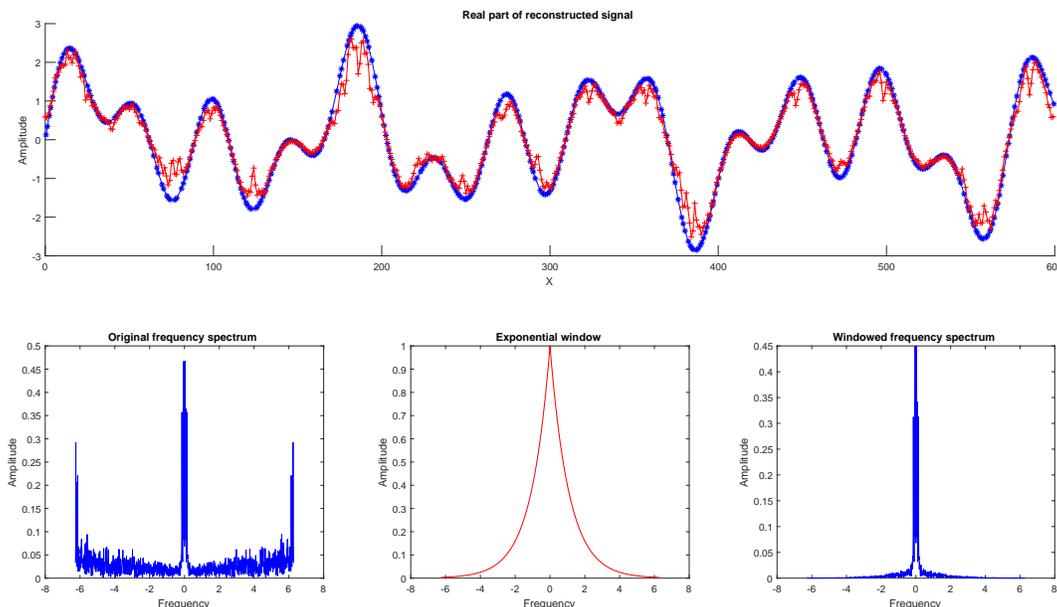


Figure 3.3: Top figure: original signal (blue), reconstructed signal (red) using the Fourier-reconstruction method with an exponential window in the frequency domain. Bottom row left: unwindowed Fourier spectrum of the original signal, middle: exponential window, right: the windowed Fourier spectrum.

3.3 Discussion

The changes between Zwartjes’ original algorithm and the one presented here are the definitions of $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha(\mathbf{A}^H \mathbf{W} \mathbf{A} + \lambda \mathbf{I}) \tilde{\mathbf{p}}_k$ and $\mathbf{d}_{k+1} = -\mathbf{r}_{k+1} + \beta \mathbf{d}_k$. Zwartjes’ algorithm returns a fit of comparable quality to the one used here, but for $N = 450$ and $M = 1199$ the algorithm presented here terminates after 7 iterations, and does not converge when using Zwartjes’ algorithm, all other parameters being equal. A variability of $\frac{1}{2} \Delta x$ has been used for all points, an amplitude variation of 30% of the maximum amplitude in the data and 25% of the points have been removed for both cases.

Besides these changes, Zwartjes [2005] appears to be wrong in their definition of α , since they divide an l_2 -norm by a *squared* \mathbf{H} norm. This means that α now has a dimension of $1/[\mathbf{d}]$, where $[\mathbf{d}]$ is the dimension of \mathbf{d} , which is $1/\text{Fourier}$. Subsequently the calculations of \mathbf{z}_k and \mathbf{r}_k are physically incorrect, since they both have dimensions of the Fourier domain, thus adding a dimensionless quantity to one with a dimension.

Whilst Zwartjes’ algorithm does give a reasonable result, see figure 3.5, it ran for the maximally allotted 1,000 iterations, and tests have been conducted with a larger number of iterations and the algorithm keeps on iterating. This means that even though the CG scheme produces correct output, the iterations themselves do not converge. It has been noted that the result is comparable with the implementation according to Shewchuck after the number of iterations Shewchuck’s version needs. The stopping criterion seems to go wrong due to the minus sign in the definition of \mathbf{d}_{k+1} , since this reverses the search direction each time. Subsequently updating \mathbf{r}_{k+1} using \mathbf{p}_k instead of \mathbf{d}_k makes the residue flip between positive and negative.

Matrices \mathbf{A} and \mathbf{F} grow very large, very fast in a multidimensional case. For instance, consider an x, y grid containing 100 points in both directions. This results in $N = 10^2 * 10^2 = 10^4$, $M = 2 * 10^2 * 2 * 10^2 = 4 * 10^4$ and thus \mathbf{A} and \mathbf{F} having $10^4 * 4 * 10^4 = 4 * 10^8$ elements already. A single shot record can consist of several hundred to thousand traces, each with 2,000 time samples. This would result in the matrices having up to $4 * 10^{12}$ elements. Storing these as complex double precision values would require 58TB of RAM. Therefore they should be implemented as functions as opposed to matrices. Having them as evaluable functions, such as a non-uniform FFT, can solve for the Fourier

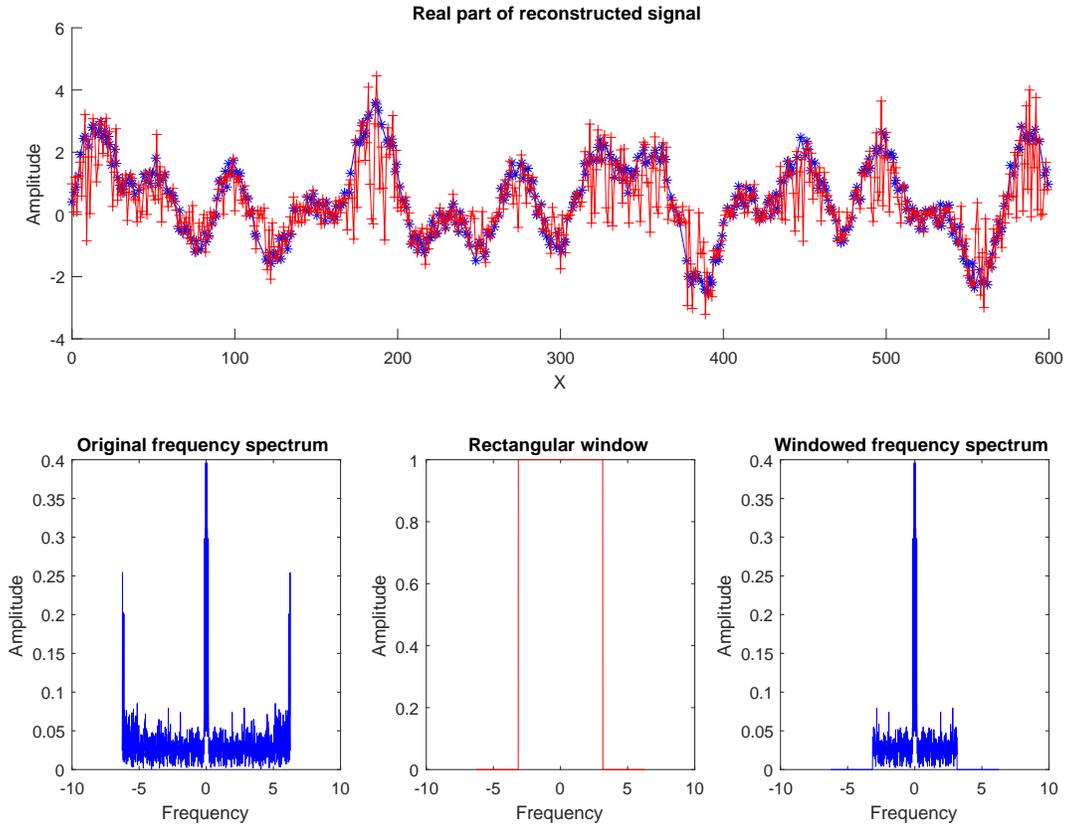


Figure 3.4: Top figure: original signal (blue), reconstructed signal (red) using the algorithm presented in section 3.1.2. Bottom row left: unwindowed Fourier spectrum of the original signal, middle: rectangular window, right: the windowed Fourier spectrum.

transform from a non-uniformly sampled input \mathbf{p} to a regularly sampled frequency vector $\tilde{\mathbf{p}}$ directly. This can improve the calculation efficiency from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$, due to better implementation of FFT compared to the naive implementation, which is used here for \mathbf{A} and \mathbf{F} . As well as a function requiring almost no memory to store, thus RAM requirements go down considerably.

This reconstruction allows for 2D non-uniform grids to be aligned (fig. 3.1), thus $(\mathbf{x}, \mathbf{y}, t)$, per scalar time and thus it operates independently on each time step. It is imperative that continuity across all three dimensions is taken into account. Also it does not keep reflection events which are connected in the space-time domain connected in the Fourier domain. This connection is of paramount importance to this work, since the reduced number of frequencies present in the data should be compensated by a much denser sampling in the spatial domain. Measuring the wavefield using a different acquisition method than the current ones, as explained in chapter 1, would increase spatial coverage tremendously. It is hoped that by increasing sampling density in both receiver coordinates, sampling density in the frequency domain can go down. Another transformation, which does take spatial continuity into account, is thus required. One of the transformations which does that is the Radon transform (Radon [1917]).

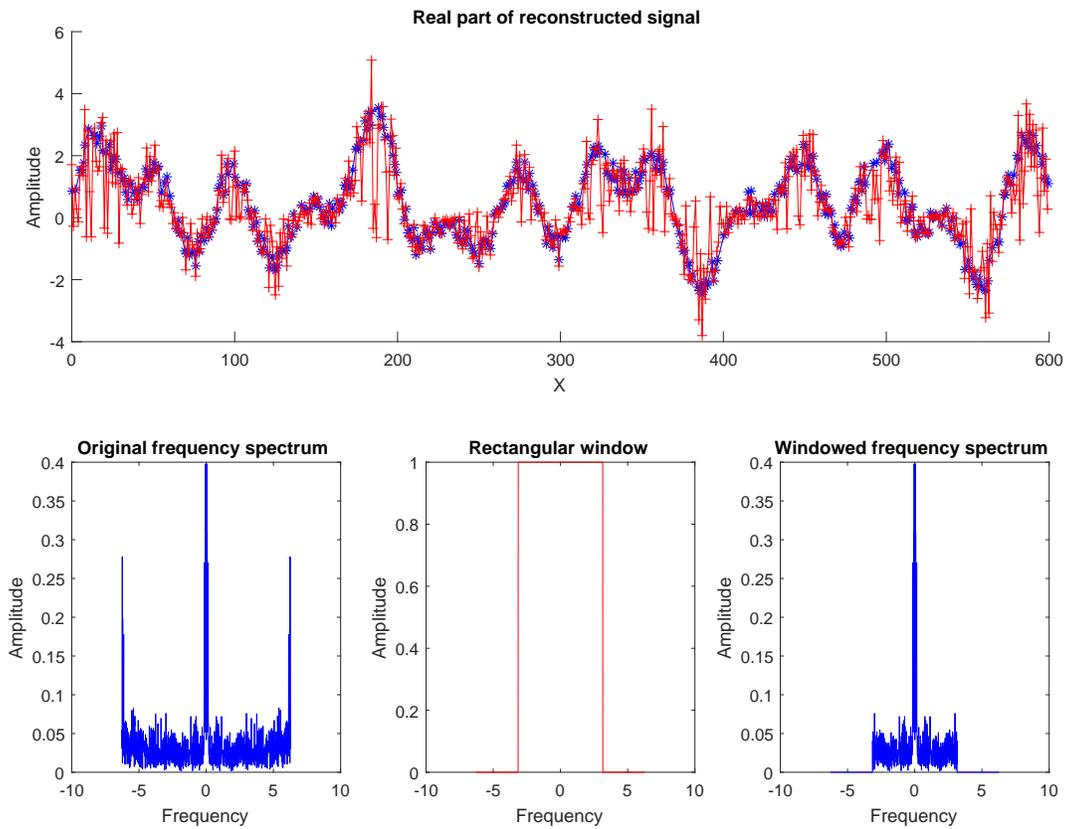


Figure 3.5: Top figure: original signal (blue), reconstructed signal (red) using the algorithm by Zwartjes (2005). Bottom row left: unwinded Fourier spectrum of the original signal, middle: exponential window, right: the windowed Fourier spectrum.

Chapter 4

Radon reconstruction of missing frequencies

In order to reconstruct data with missing frequency components, the Fourier transformation was found to be insufficient in taking spatial connectivity into account (ch. 3). A linear Radon transform will thus be used (Radon [1917]). The Radon transform keeps connected points in time-space connected in the transform domain, i.e. a continuous reflection event in a seismic record will show up as a continuous event when transformed to the Radon domain as well. This means that spatial continuity factors into the transform, providing information on the image beyond what a Fourier transform would do.

The linear Radon transform will be iteratively applied to the data, in a Conjugate Gradient (CG) scheme. CG schemes are used to approximate an inverse solution iteratively, in case the model itself is too large to be inverted easily (Shewchuk [1994]). Another benefit of CG over direct inversion is that direct inversion would need to account for every little detail in the data, including all the noise, whereas CG minimises an objective function under some constraint. This thus allows for a ‘good enough’ solution, as opposed to the exact one. Noise can be discarded under the constraint, leaving a much cleaner model as opposed to what would be retrieved using direct inversion. Non-linear constraints cannot be inverted exactly, since their matrix representation is based on the data, thus changing the model in each calculation step.

First the Radon transform is presented (sec. 4.1), then the problem is posed in mathematical form (sec. 4.2), followed by introduction of the iterative solver techniques (sec. 4.3). The model used for testing is introduced (sec. 4.4) and finally results of the procedure are presented and discussed (sec. 4.5 and 4.6). The same notation convention as in chapter 3 will be used.

4.1 Radon Transform

The Radon transform is a transformation of a data panel where one sums the energy along a specified trajectory through the panel. The generic Radon transform $\mathbf{U}(q, \tau)$ is given by (Cao et al. [2003]):

$$\mathbf{U}(q, \tau) = \int_{-\infty}^{\infty} \mathbf{D}(x, t = \tau + q\phi(x)) dx, \quad (4.1)$$

where $\mathbf{D}(x, t)$ is the initial shot record in space-time, τ the intercept time in seconds, q the slope or curvature and $\phi(x)$ the path to integrate the energy over. In seismology most often the slant-stack, or τ, p transform is used on CMP gathers (Verschuur [2006], chapter 2). This uses a linear path, i.e. $\phi(x) = x$, thus equation 4.1 results in:

$$\mathbf{U}(p, \tau) = \int_{-\infty}^{\infty} \mathbf{D}(x, t = \tau + px) dx, \quad (4.2)$$

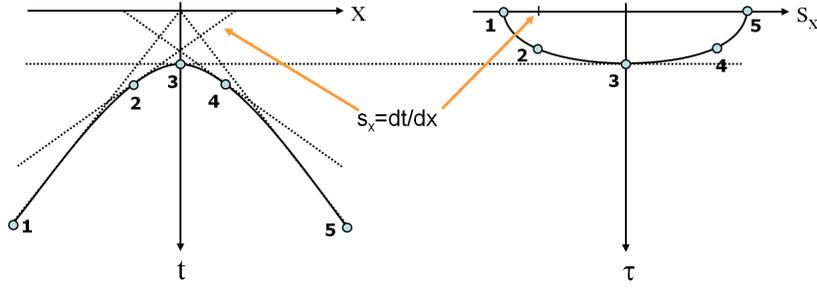


Figure 4.1: Mock-up of the linear Radon transformation. The hyperbola in (t, x) spanned by points 1 to 5 is mapped along tangent lines to an ellipse in (τ, S_x) , with τ the intercept time of the line and S_x the slowness. Image courtesy of Dr Verschuur (Verschuur [2006]).

where p is the slowness in s/m, and can be seen as the angle along which the summation takes place, see fig. 4.1 for an illustration.

When calculating this transform directly from the (x, t) domain to the (τ, p) domain, calculations are slow and imprecise, since interpolation is necessary. Both problems can be solved by going through the space-frequency, (x, ω) , domain by using the Fourier transformation \mathcal{F} :

$$\mathcal{F}\{\mathbf{U}(\tau, p)\} = \mathbf{U}(p, \omega) = \iint_{-\infty}^{\infty} \mathbf{D}(x, \tau + px) e^{-i\omega\tau} dx d\tau, \quad (4.3)$$

or

$$\iint_{-\infty}^{\infty} \mathbf{D}(x, \tau') e^{-i\omega(\tau' - px)} dx d\tau' = \int_{-\infty}^{\infty} \mathbf{D}(x, \omega) e^{i\omega px} dx, \quad (4.4)$$

where the substitution $\tau' = \tau + px$ has been made from the first to the second line.

Seismic data is discretised by nature, even though geophones are analog devices, due to the storage on disks. This means that eq. (4.4) will have to be discretised as well; resulting in the following discrete transform relations:

$$\mathbf{U}(p_m, \omega) \approx \sum_{n=1}^N \mathbf{D}(x_n, \omega) e^{i\omega x_n p_m} \quad (4.5)$$

$$\mathbf{D}(x_n, \omega) \approx \sum_{m=1}^M \mathbf{U}(p_m, \omega) e^{-i\omega x_n p_m}. \quad (4.6)$$

Equations (4.5) and (4.6) can be evaluated very quickly through usage of FFT packages, and a simple loop over all desired values of p and summing over x gives the Fourier transformed linear Radon Transform, which can be inverse-FFTed to obtain the τ, p transform itself (fig. 4.2).

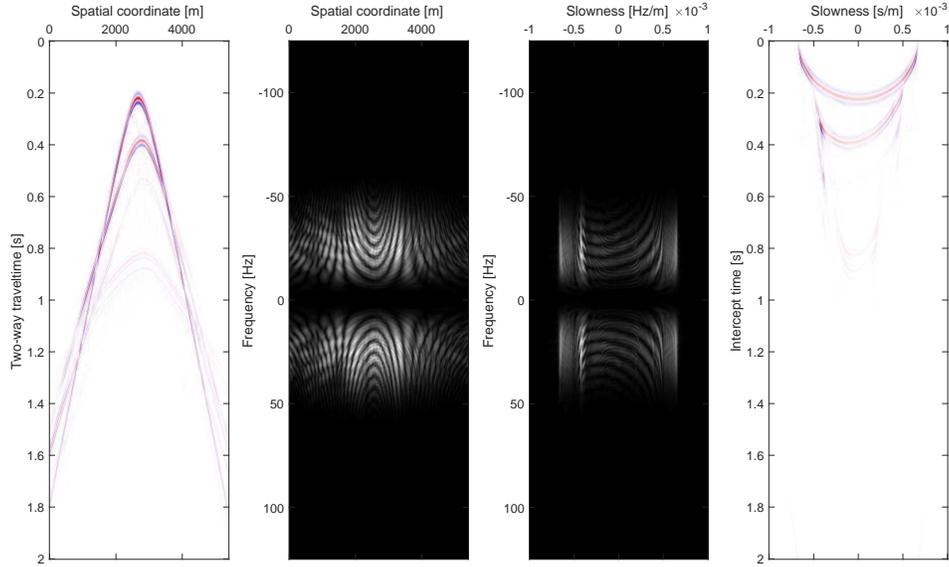


Figure 4.2: The linear Radon transform of a shot record (left panel). The second panel shows the temporal Fourier transform of the initial data, the third panel shows the Fourier transformed Radon transformation; the fourth panel shows the Radon transform in the τ, p domain. Amplitudes of the various plots are irrelevant, the shape and axes labels are what matters for this showcase example, hence it has been elected not to show a colour bar.

4.2 Mathematical problem statement

A two dimensional shot record $\mathbf{P}(x, t)$ is usually presented in a 2D manner, such as an image, whereas an array is required for further processing. This can be realised by storing the data in lexicographic, i.e. column major, order. Thus the data in matrix \mathbf{P} will be stored as follows:

$$\mathbf{P} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \rightarrow \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} = \mathbf{p}.$$

This enables us to use an array for the vector \mathbf{p} . The same column major storage is used for all vectors in this section.

The problem we strive to solve can now be posed as:

$$\mathbf{A}\tilde{\mathbf{p}} = \mathbf{p}, \quad (4.7)$$

where \mathbf{p} is our measured data, i.e. the shot record in the (x, t) domain, \mathbf{A} is the conjugate Linear Radon transform (eq. 4.6), i.e. it transforms the shot from the (τ, p) domain back to (x, t) , and finally $\tilde{\mathbf{p}}$ contains the desired Radon coefficients of the shot under certain conditions, such as minimal energy spread, reflector continuity etc. The usual way of solving this would be to simply invert \mathbf{A} :

$$\tilde{\mathbf{p}} = \mathbf{A}^{-1}\mathbf{p}, \quad (4.8)$$

but this is seldom possible since \mathbf{A} is rarely a square matrix, and usually poorly conditioned. The textbook way of solving this, is to left-multiply eq. 4.7 with the Hermitian conjugate of \mathbf{A} . Since $\mathbf{A}^H\mathbf{A}$ is square and invertible, we have:

$$\mathbf{A}^H\mathbf{A}\tilde{\mathbf{p}} = \mathbf{A}^H\mathbf{p}, \quad (4.9)$$

Finally a preconditioner, \mathbf{M} , is applied to make computation of the final solution easier and more robust, as well as to contain more stringent conditions on the model space:

$$\mathbf{M}^{-1}\mathbf{A}^H\mathbf{A}\tilde{\mathbf{p}} = \mathbf{M}^{-1}\mathbf{A}^H\mathbf{p}. \quad (4.10)$$

\mathbf{M} must be self-adjoint and the closer \mathbf{M}^2 is to $\mathbf{A}^H\mathbf{A}$, the faster convergence will be. To make calculation easier one additional step is taken:

$$\mathbf{M}^{-1}\mathbf{A}^H\mathbf{A}\mathbf{M}^{-1}\mathbf{z} = \mathbf{M}^{-1}\mathbf{A}^H\mathbf{p} \quad (4.11)$$

$$\mathbf{z} = \mathbf{M}\tilde{\mathbf{p}}, \quad (4.12)$$

since the $\mathbf{M}^{-1}\mathbf{A}^H\mathbf{A}\mathbf{M}^{-1}$ is sparse and symmetric and can thus be evaluated very fast, at the cost of having to perform $\tilde{\mathbf{p}} = \mathbf{M}^{-1}\mathbf{z}$ once at the end. Various preconditioners based on norms are given in table 4.1, where $\mathbf{C}_x^{-1/2} = \mathbf{M}^{-1}$.

Table 4.1: Several penalty functions, $\rho(x)$, with corresponding influence function $\varphi(x) = \frac{\partial\rho(x)}{\partial x}$ and model weight $\mathbf{C}_x^{-1} = \frac{\varphi(x)}{x}$. We see that a determines the type of norm used. Based on [Zwartjes \[2005\]](#).

Norm type	$\rho(x)$	$\varphi(x) = \frac{\partial\rho(x)}{\partial x}$	$\mathbf{C}_x^{-1} = \frac{\varphi(x)}{x}$
$l_{1,2}$	$\sqrt{1+x^2}$	$x(1+x^2)^{-1/2}$	$(1+x^2)^{-1/2}$
Cauchy	$\frac{1}{2}\ln(1+x^2)$	$x(1+x^2)^{-1}$	$(1+x^2)^{-1}$
Geman-McClure	$\frac{q}{2}\frac{x^2}{1+x^2}$	$x(1+x^2)^{-2}$	$(1+x^2)^{-2}$
Generalised function	$\frac{1}{2(1-a)}(x^2+1)^{1-a}$	$x(1+x^2)^{-a}$	$(1+x^2)^{-a}$

4.3 Iterative solvers

Inversion of $\mathbf{M}^{-1}\mathbf{A}^H\mathbf{A}\mathbf{M}^{-1}$ is, even though possible, difficult since it is usually a large matrix, and often poorly conditioned. This is circumvented by using a CG scheme. Since CG requires a positive-definite \mathbf{A} , the scheme has to be changed to CGNE: Conjugate Gradient for the Normal Equations, also called CGLS: Conjugate Gradient for Least Squares ([Kaipio and Somersalo \[2005\]](#)). We strive to minimise the following objective function:

$$J = \|\mathbf{p} - \mathbf{A}\tilde{\mathbf{p}}\|_2^2 + \|\mathbf{C}_p^{-1/2}\tilde{\mathbf{p}}\|_2^2 \quad (4.13)$$

$$\mathbf{C}_{\mathbf{p}_{jj}} = \left((\mathbf{A}^H\mathbf{p}_j)^* \mathbf{A}^H\mathbf{p}_j + \sigma_{\mathbf{A}^H\mathbf{p}}^2 \right)^a, \quad (4.14)$$

where the preconditioner is given as $\mathbf{M}^{-1} = \mathbf{C}_{\tilde{\mathbf{p}}^{i,jj}}^{-1/2}$; the subscript jj denotes that this is a diagonal matrix which is calculated for each entry j of the vector $\tilde{\mathbf{p}}$ ([Zwartjes \[2005\]](#)). $\sigma_{\mathbf{A}^H\mathbf{p}}^2$ is the variance of the Radon transformed input, and a controls the type of norm as presented in table 4.1. The CGNE scheme to solve equation (4.13) can be implemented as follows:

$$\begin{aligned}
k &= 0 \\
\mathbf{z}_0 &= \mathbf{0} \\
\mathbf{r}_0 &= \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{A}^H \mathbf{p} \\
\mathbf{d}_0 &= \mathbf{r}_0 \\
\mathbf{t}_0 &= \mathbf{p} \\
\text{while } \frac{\mathbf{r}_k^H \mathbf{r}_k}{\mathbf{r}_0^H \mathbf{r}_0} &> \varepsilon, \quad k \leq k_{\max} \\
\alpha &= \frac{\mathbf{r}_k^H \mathbf{r}_k}{\mathbf{d}_k^H \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{A}^H \mathcal{F}^{-1} \mathbf{S}_{\odot} \mathcal{F} \mathbf{A} \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{d}_k + \mathbf{d}_k^H \lambda \mathbf{I} \mathbf{d}_k} \\
\mathbf{z}_{k+1} &= \mathbf{z}_k + \alpha \mathbf{d}_k \\
\mathbf{t}_{k+1} &= \mathbf{t}_k - \alpha \mathcal{F}^{-1} \mathbf{S}_{\odot} \mathcal{F} \mathbf{A} \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{d}_k \\
\mathbf{r}_{k+1} &= \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{A}^H \mathbf{t}_{k+1} + \alpha \lambda \mathbf{I} \mathbf{d}_k \\
\beta &= \frac{\mathbf{r}_{k+1}^H \mathbf{r}_{k+1}}{\mathbf{r}_k^H \mathbf{r}_k} \\
\mathbf{d}_{k+1} &= \mathbf{r}_{k+1} + \beta \mathbf{d}_k \\
k &= k + 1 \\
\tilde{\mathbf{p}} &= \mathbf{C}_{\mathbf{p}_{jj}}^{-1/2} \mathbf{z}_k
\end{aligned}$$

where α is the step size to take in the direction prescribed by \mathbf{d} and β gives the change of direction. $\mathcal{F}^{(-1)}$ is the forward (inverse) Fourier transform, which lets the sampling operator \mathbf{S}_{\odot} select the frequencies that are present in the data. λ , is the regularisation parameter, \mathbf{I} the identity matrix. The scheme should end based on the squared norm, $\mathbf{r}_k^H \mathbf{r}_k$ being sufficiently small, i.e. the residual will be small and thus \mathbf{z} will be very close to the true solution.

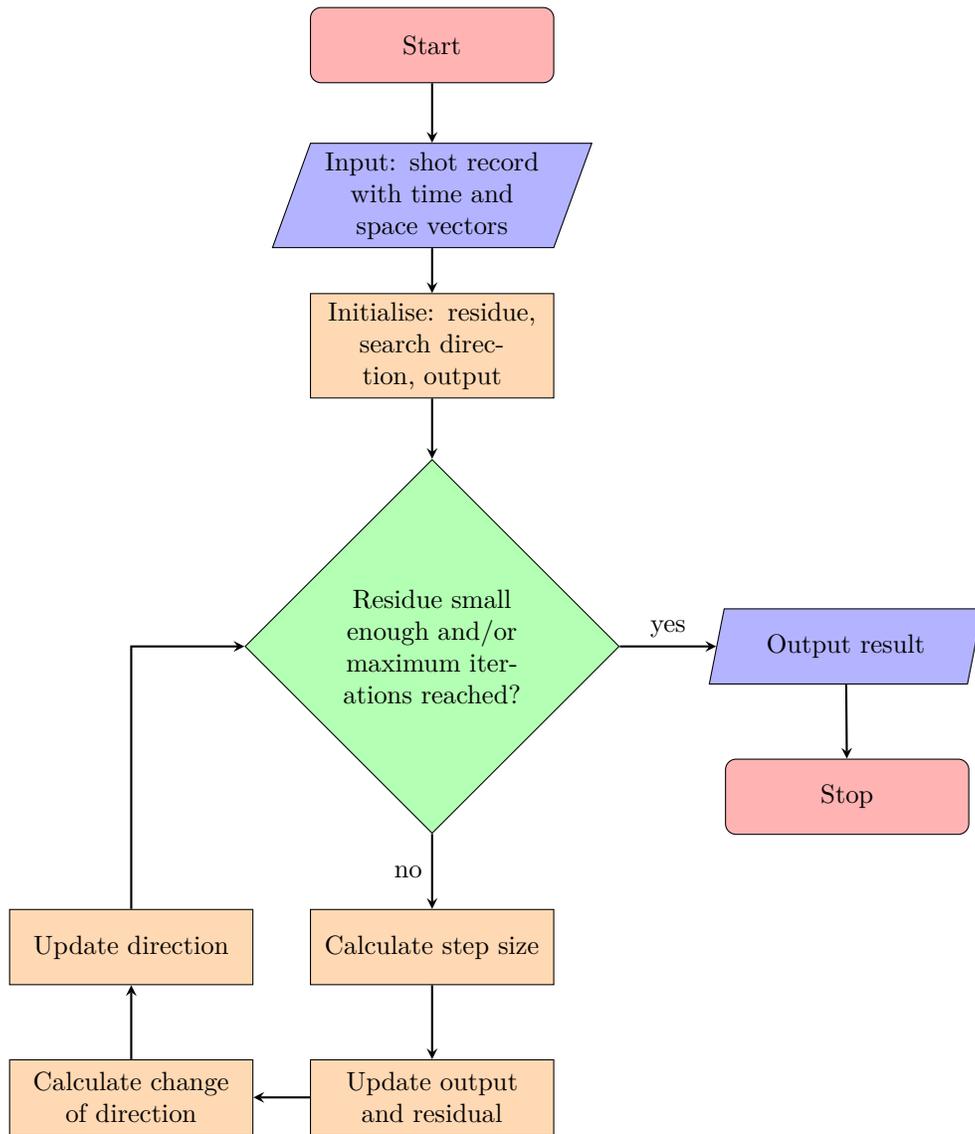


Figure 4.3: Flowchart of the CGNE algorithm as presented in sec. 4.3.

4.4 Model

In this section a synthetic data example is discussed. The model was kindly provided by Dr Ir Verschuur, and is presented in figure 4.4.

The model is a grid of 5400 by 1250m, (x, z) , using a grid size of 5×5 m, containing velocity and density at each grid point. Using an acoustic finite difference modelling, second order in time, fourth order in space, staggered grid, 361 shots into 361 receivers have been modelled along the profile, with source and receiver separation both at 15m. Shot numbers run from 1 to 361, are shot from left to right through the model. Shot 1 thus has offsets 0 to 5400m, shot 181 -2700 to 2700m, and shot 361 -5400 to 0m. The model contains eleven layers, their P-wave velocities and densities are given in table 4.2. The layer with 3500m/s is considered a “salt layer”, which is the thick, black layer in the middle of the model. The layer underneath that is the reservoir proper.

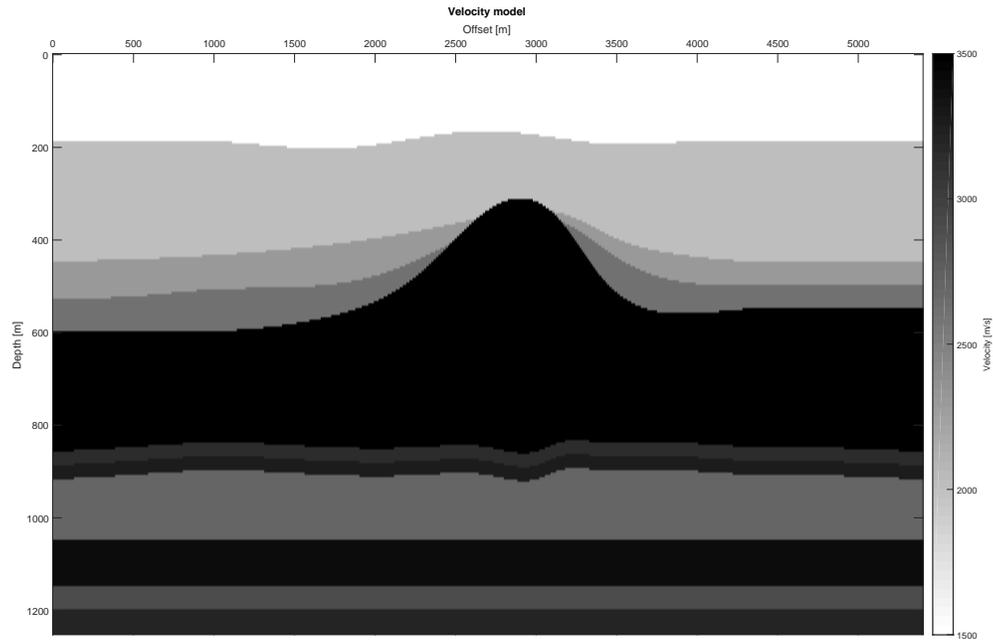


Figure 4.4: Velocity model used in the rest of the thesis. Colour denotes the velocity; white is 1500m/s up to black, 3500m/s.

Table 4.2: P-wave velocities and densities of the synthetic model. Layers are from top to bottom in the model

P-wave velocity [m/s]	Density [kg/m ³]
1500	1000
2000	1800
2300	2000
2600	2300
3500	2500
3150	2600
3250	2420
2700	2800
3400	3200
2900	2900
3200	3000

4.5 Results

As can be seen in figure 4.6 the reconstruction removes most of the aliasing. Figure 4.5 clearly shows that reconstruction is rather good along the asymptots of the main reflectors, but the region around zero offset gets reconstructed quite badly, which is not surprising given that a linear Radon transformation has been used. The problem with a linear transformation on a hyperbolic shape is that the apex of the hyperbola is mapped to a large ellipsoid in the Radon domain (fig. 4.1). The asymptots, on the other hand, are almost straight lines, so they map to nearly a point in the Radon domain, making reconstruction easier. This is reinforced by looking at traces along the asymptot (fig. 4.8a) and at zero offset (fig. 4.8b). The trace along the asymptot shows lower amplitudes of the aliasing, especially along 0.6-1.2s TWT, whereas for the zero offset the trace is almost untouched by the reconstruction.

Reconstruction fills in the missing frequencies, albeit with a lower amplitude (fig. 4.9), and increases amplitudes at high frequencies, which is undesired. A bandpass filter to cut out frequencies above some 70Hz might be of help to reduce this effect.

For each shot record a Q -factor has been calculated as the sum of the squared difference between the original and reconstructed shot, normalised with the squared sum of the original shot record:

$$Q = \frac{\sum (\mathbf{p} - \mathbf{p}_{\text{rec}})^2}{\sum \mathbf{p}^2}. \quad (4.15)$$

For several reconstruction parameters this has been calculated. The mean and variance of the Q -factors of each data set for all shots are presented in table 4.3, the accompanying plot can be found in figure 4.7. It can be observed that up to 30% reduction the Q -factors are almost equal, with a clear tendency to be lower for the central shot records. This means that the reconstruction of far offsets is worse than that of near ones. As the reduction percentage goes up, so does the Q -factor, with increasingly bigger steps after 70% reduction. Variance shows the same trend, see table 4.3, which suggests that at high reduction percentages performance of the reconstruction is highly dependent on the frequencies chosen. This confirms the findings of chapter 2 that indeed care has to be taken when choosing frequencies.

Finally, a test has been done with a single band gap. Band gaps of 3, 6, 9, 12, 15, 18, and 21Hz width centred around 30Hz have been created in the frequency domain, after which 25 iterations using the Cauchy norm have been performed (fig. 4.10). It can be seen that small gaps, especially the one of 3Hz, show very good reconstruction in terms of shape, even at zero offset where the problems of the linear Radon transform are largest due to the highly non-linear behaviour of hyperbolas near their apex. The larger the gap, the worse the reconstruction performs the infill. Deterioration starts with the zero offset, and then progresses from the far offset towards the near offset. The overall shape is reconstructed rather well, but the amplitudes are decreased. The 3Hz gap already shows decreased

Table 4.3: Q -factor mean and variance per data set as calculated from (4.15)

Reduction %	Norm	Iterations	Mean	Variance
10	Cauchy	10	0.1158	0.0009
20	Cauchy	10	0.1233	0.0009
30	Cauchy	10	0.1383	0.0012
40	Cauchy	10	0.1726	0.0019
50	Cauchy	10	0.2205	0.0034
60	Cauchy	10	0.2981	0.0059
70	Cauchy	10	0.4029	0.0074
80	Cauchy	10	0.5590	0.0109
90	Cauchy	10	0.7887	0.0097
50	Cauchy	25	0.2169	0.0040
50	$l_{1,2}$	25	0.3624	0.0040

amplitude in the gap, and the deterioration is mainly an even greater lack of amplitude.

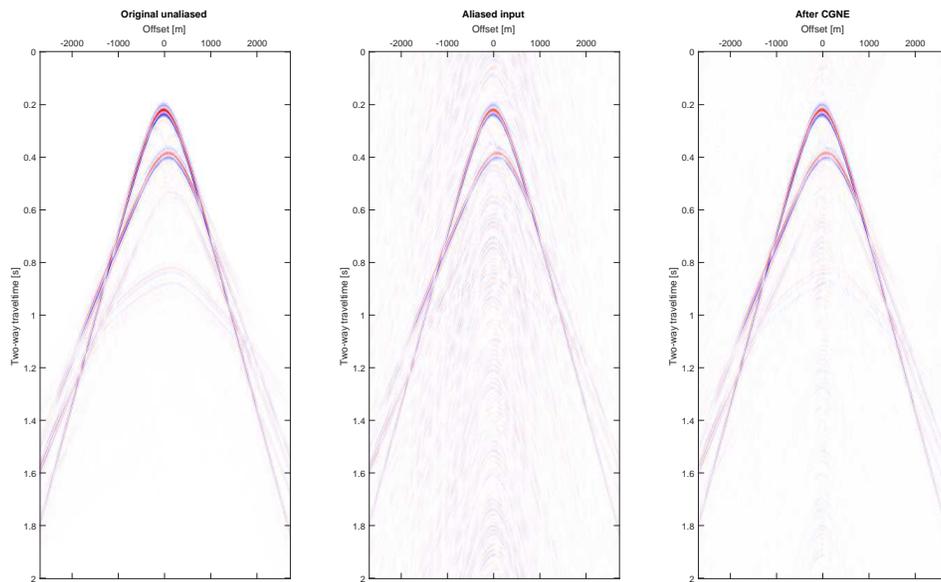


Figure 4.5: The left panel shows the original, unalised shot record. The second panel shows the shot record after removing 30% of its frequencies at random. The third panel shows the result after reconstruction with 15 steps. Colour runs in $[-12; 12]$ in all panels. $a = 1$, $k_{\max} = 15$, $\lambda = 0$. This clearly shows that the energy spread due to spectral leakage as present when selecting a subset of frequencies is moved back to the main reflecting peaks as much as possible, although not all aliasing is suppressed, especially around $x = 0$.

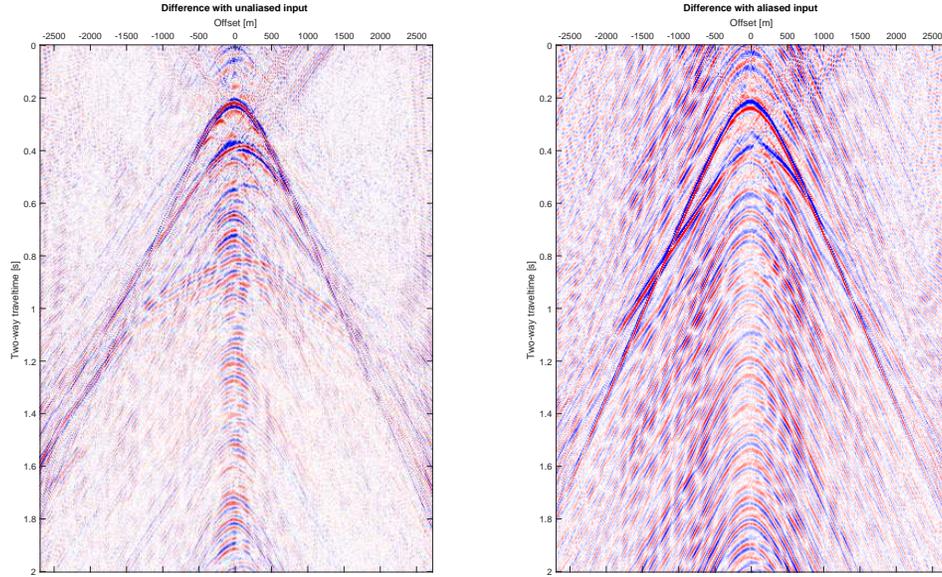


Figure 4.6: The left panel shows the difference between the reconstruction and unaliased original, i.e. panels 1 and 3 of figure 4.5. The right panel shows the difference between the reconstruction and aliased original, i.e. panels 2 and 3 of figure 4.5. Colour scale is decreased tenfold to $[-1.2; 1.2]$. Reconstruction does remove most of the aliasing, as evidenced by the right panel, but does leave a lot of spread out energy around the zero offset.

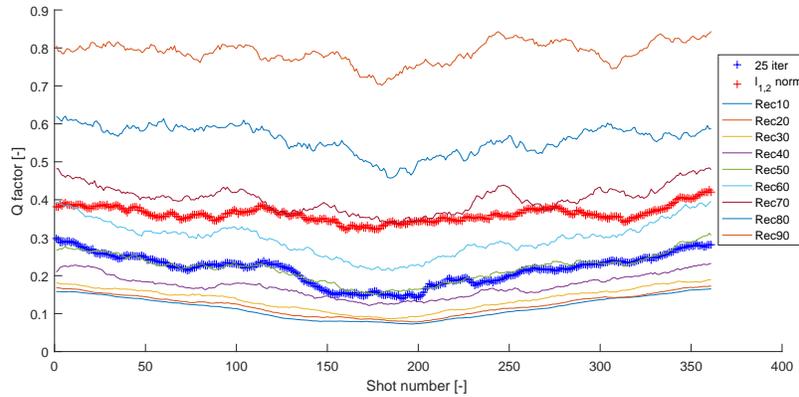


Figure 4.7: Each line contains the 25-point moving average of the Q factors per shot record for the given reduction parameters of table 4.3. For low reduction percentages the Q factors are almost equal, and lower in the centre, getting progressively higher with increasing reduction. The moving average has been used to reduce visual variance of the lines due to each shot having different frequency components removed, whilst preserving the general trend of each line.

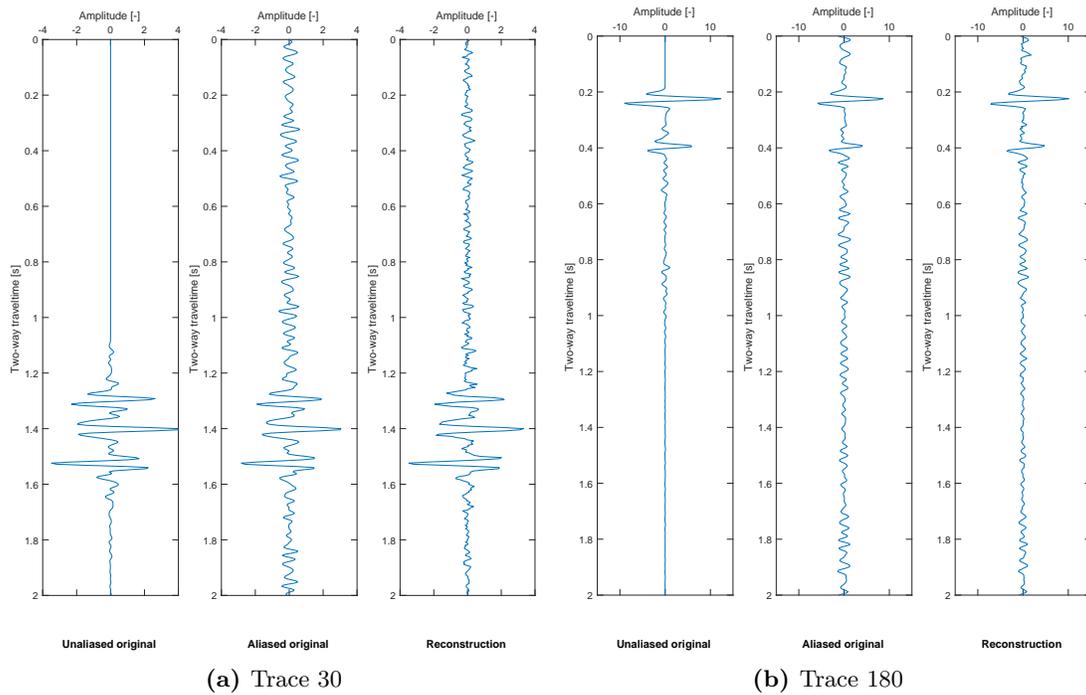


Figure 4.8: Cut-out of trace 30 (offset -2250m) and 180 (zero offset) out of figure 4.5. The left three panels are trace 30 out of the unaliased input, aliased input and reconstructed result; the right three are the same for trace 180. Between 0 and 1.2s TWT the amplitude on trace 30 goes down, which it should because this is energy before the first arrival; amplitude on trace 180 is almost unchanged after reconstruction, confirming poor reconstruction.

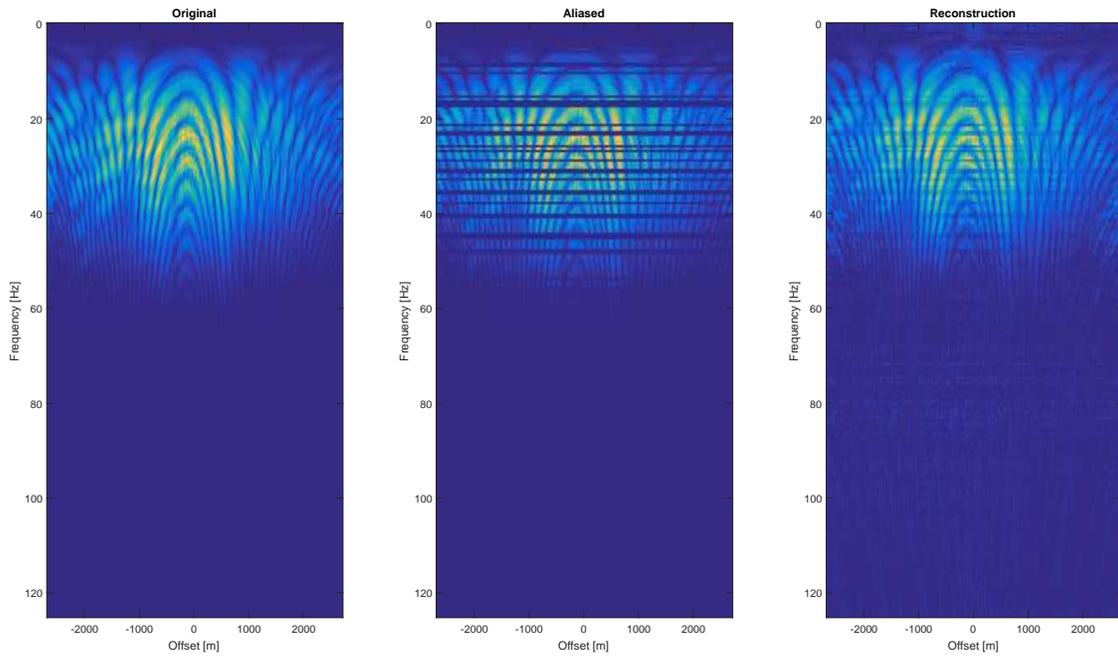


Figure 4.9: The left panel shows the frequency spectrum of the original of figure 4.5. The middle panel shows the frequency spectrum after removing frequencies, where frequencies have been removed. Horizontal stripes are visible on the missing frequencies since they have been set to zero amplitude. The reconstruction, given in the right panel, of missing frequencies is quite good, although some energy smearing is visible. Colour scale is the same for all panels.

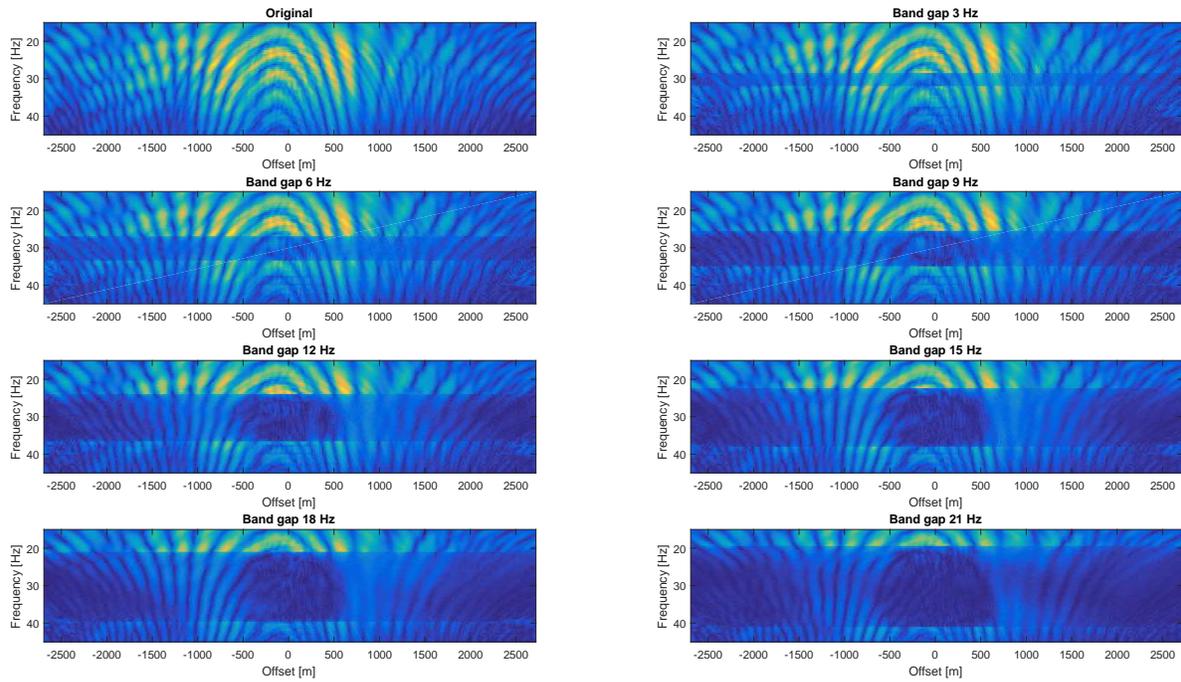


Figure 4.10: Plots of the frequency spectra of the central shot reconstructed with the Cauchy norm and 25 iterations. The title of each subplot mentions the band gap size, which is symmetric around 30Hz; the top left plot shows the original. Colour scale has been kept constant for each plot. Reconstruction gets progressively worse the larger the band gap, starting at zero offset and then at large offset, growing inwards.

4.6 Discussion

Aliases of the main reflectors are clearly visible in the central panel of fig. 4.11, which are removed nicely in the reconstruction, albeit at a cost of removing the weak reflectors with apex around 0.8s TWT. This shows that weighting based on signal amplitude can be problematic if spectral noise of large peaks is stronger than the amplitude of weak reflectors, such as the reservoir below the salt dome.

Reconstruction using the $l_{1,2}$ norm is rather slow, in the sense that a lot of iterations are necessary to remove the aliases, especially around zero offset. A benefit of using this norm would be that almost no artefacts get introduced in the data (fig. 4.12). The Cauchy norm is a lot faster compared to the $l_{1,2}$, but at the price of small artefacts around zero offset and small time. Geman-McClure, finally, is far too aggressive. It removes the aliasing almost instantly, but also introduces strong artefacts at the same time. Therefore all results given in this work have been made with the Cauchy norm.

Reconstruction in the presence of ground roll is very difficult, see the reconstruction of a shot from the same data set from Poland as in section 2.4 (fig. 4.13). Ground roll propagates at low velocity, and thus is centred around the zero-offset, with very high amplitude, up to 90% of shot energy (Schlumberger Oilfield Glossary [2018]). This means that any type of SNR-based weighting, such as in equation (4.14) will consider all the data noise. Ground roll filtering will thus have to be performed before reconstruction, or incorporated in it.

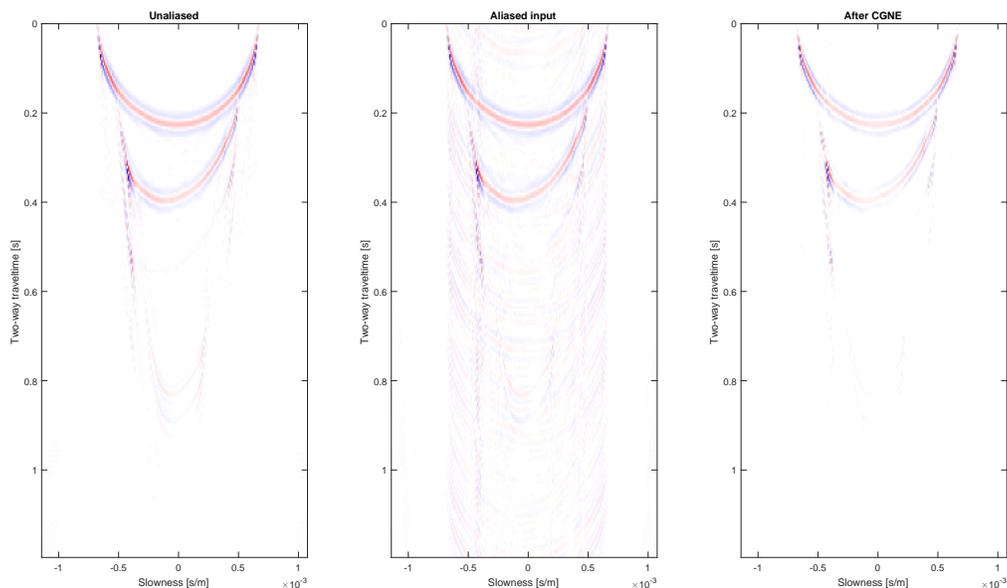


Figure 4.11: The left panel shows the linear Radon transform of the original unalised shot record as presented in figure 4.5, the middle panel shows the linear Radon transform of the aliased shot record, the right panel the linear Radon transform of the reconstructed shot. Aliasing is reduced, especially the aliased ellipses of the strongest top one, between -0.4ms/m and 0.4ms/m. The ellipses going down to 0.8s TWT in the left panel are almost invisible in the right panel though, they got lost in the reconstruction.

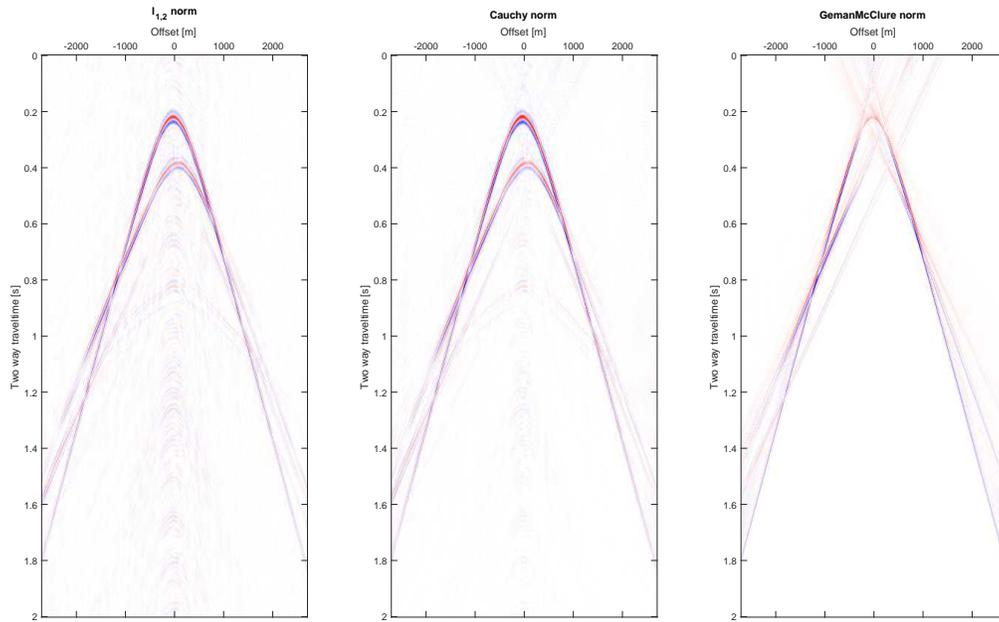


Figure 4.12: Reconstruction results of the same shot record, with the same 25% missing frequencies, using 20 iterations each. The used norm differs per panel: left is the $l_{1,2}$ norm, middle uses Cauchy and finally right the Geman-McClure norm. Note the presence of strong aliasing around zero offset in the left panel, and strong artefacts in the right.

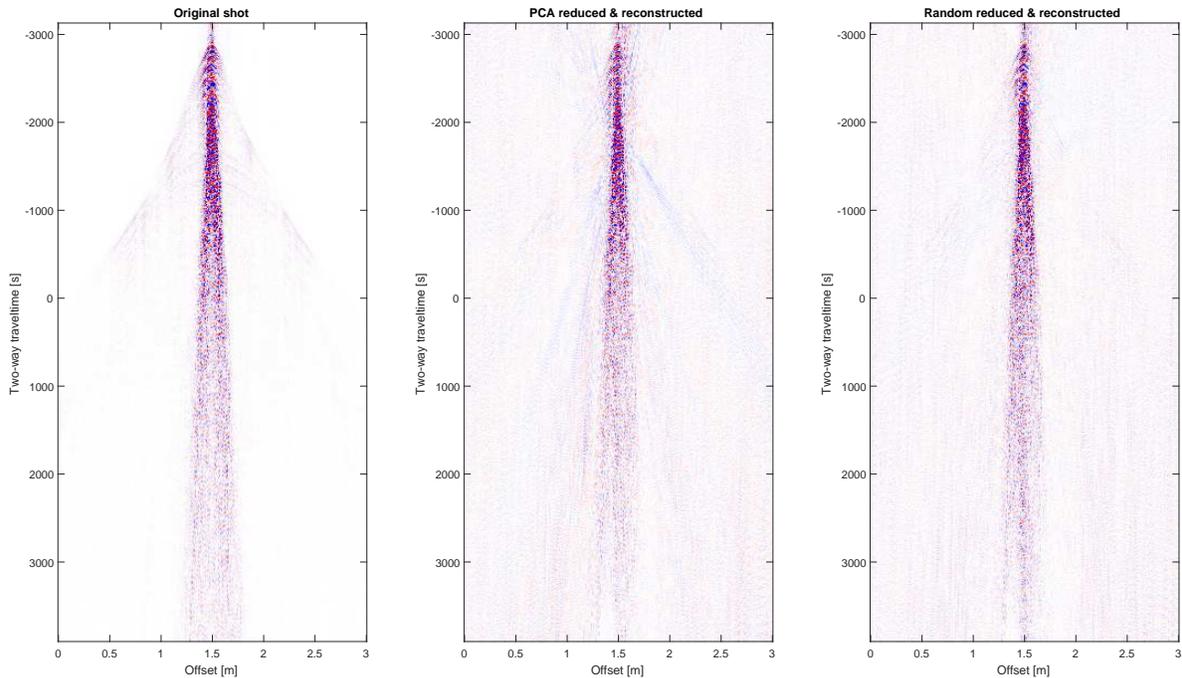


Figure 4.13: The left panel shows the original, unaliased shot record. The second panel shows the reconstructed shot record after removing 50% of its frequencies, as given by the PCA on its dataset, on the right is the reconstructed shot with 50% random frequencies removed. Both reconstructions have been performed with 10 iterations and the Cauchy norm. A clip of 90% of the amplitude has been used to allow visualisation of the areas outside the ground roll.

Chapter 5

Quality description and results

In the previous chapter the reconstruction has been shown to reduce aliasing per shot record. The quality of the migrated section will be discussed in this chapter. Migration will be considered as a black-box algorithm here and will be applied the same for all different parameter selections.

The quality of a seismic image, usually presented in the form of a migrated depth section, after some processing chain is often expressed qualitatively, but not quantitatively. Usually results are presented by putting the image obtained by the previously used processing chain next to that of the proposed new one, and explaining in words things as ‘the reflector at depth x has a better continuity using the new method’, or ‘the new AVP method clearly covers more angles’ (e.g. [Kumar \[2015\]](#), chapter 2.9). Therefore a more quantitative method is desired to make an objective comparison between images. [Alcalde et al. \[2017\]](#) show that how an image is presented is very important for the interpretation. They use two quality indicators: contrast in the image and continuity of reflectors. Line continuity was also used as a quality indicator by [Visser \[2017\]](#), who additionally used the correlation with a base image, as they had a synthetic set and thus a model without noise and maximum shot/receiver coverage to compare other images against.

First the quality metrics used to describe the migrated sections will be introduced (sec. 5.1), then the results will be presented (sec. 5.2).

5.1 Quality metrics

[Alcalde et al. \[2017\]](#) describe two main quality indicators on migrated seismic sections: contrast and average line length. We define contrast as the interquartile range of pixel values:

$$C = Q_{75} - Q_{25}, \quad (5.1)$$

where Q_{75} and Q_{25} are the third and first quantiles of pixel values for the area under consideration respectively. The reasoning here is that a broader interquartile range means more variation in light and dark colours, and thus a higher contrast. In the presence of noise this can be misleading, for instance the top middle panel of figure 5.5 shows a much broader distribution than the reconstructed one, which is again broader than the original image, but in this case the smaller the amplitude distribution is, the better.

Amplitude variations between neighbouring pixels can be found using Canny’s edge detection algorithm. This method finds edges by looking for local maxima in the gradient of the input image, using the derivative of a Gaussian filter. Using two thresholds, a strong and weak one, both strong and weak edges can be detected. For a full mathematical treatment the reader is referred to the original article ([Canny \[1986\]](#)). After edge detection a binary image will be obtained. Within that binary image connected areas of ones can be distinguished and labelled, giving the amount of pixels per connected area (fig. 5.1). This value can be interpreted as length for mainly linear features, such as reflectors in



Figure 5.1: Figure showing how components are connected in a binary image. Obtained from the official MATLAB documentation [The MathWorks \[2018\]](#)

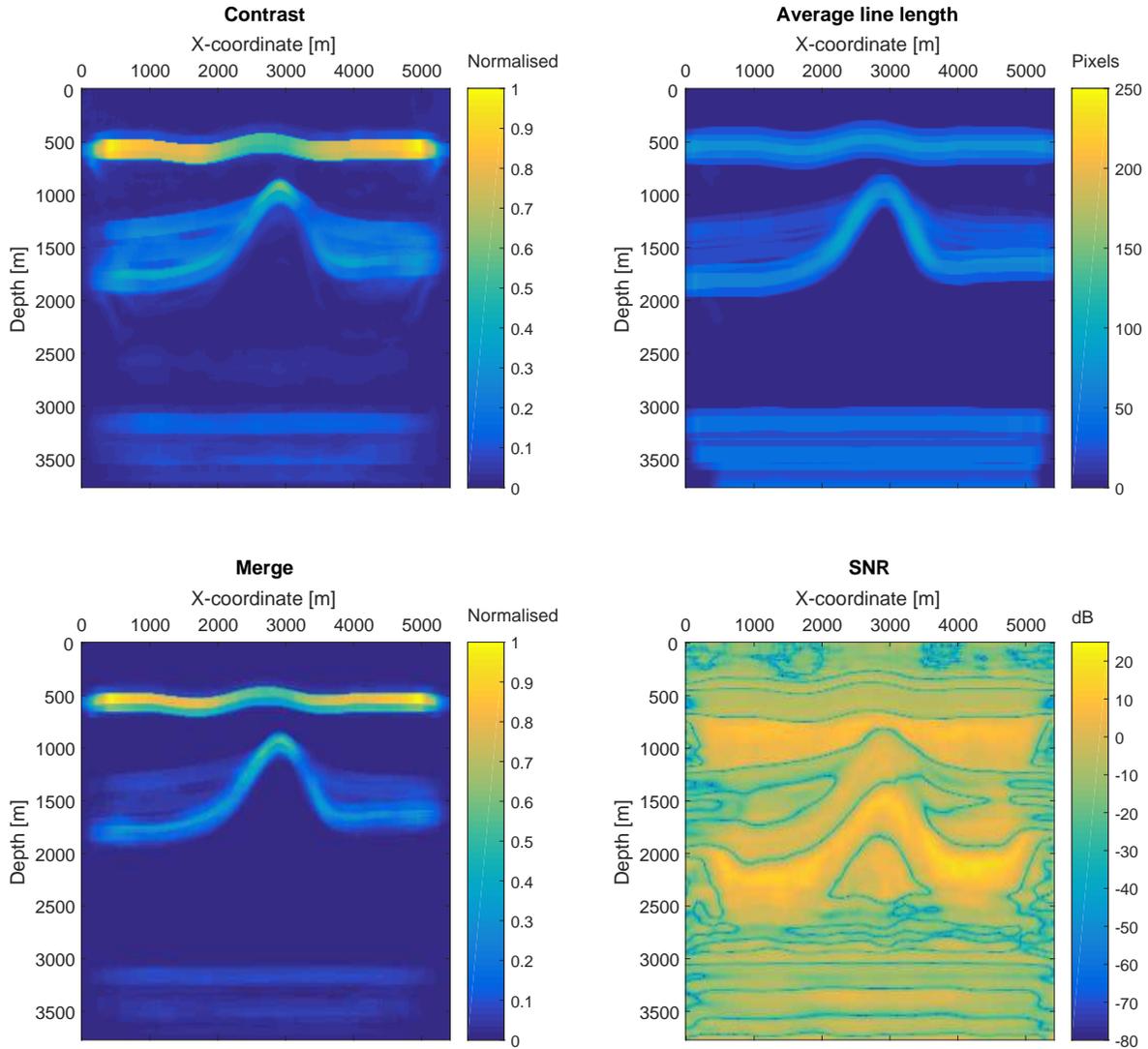
seismology. With a lot of noise, components can easily get connected though, such that the average line length goes up with more noise, instead of going down due to broken lines, see the top right panel in figure 5.7 at 20% reduction.

Using a sliding window the contrast and average line length are calculated for each pixel (fig. 5.2). Additionally, the Signal-to-Noise ratio (SNR), given by μ^2/σ^2 , where μ is the mean and σ^2 the variance of an area, is calculated. This is done for pixel values, i.e. amplitudes, and presented in dB. In areas having a mean of zero amplitude this will result in a very low SNR, even though it might be true that there is no amplitude or variation there, such as the top left panel of figure 5.4.

The 2D correlation coefficient between images **A** and **B** is calculated as:

$$r = \frac{\sum_{m,n} (A_{mn} - \bar{\mathbf{A}}) (B_{mn} - \bar{\mathbf{B}})}{\sqrt{\left(\sum_{m,n} (A_{mn} - \bar{\mathbf{A}})^2\right) \left(\sum_{m,n} (B_{mn} - \bar{\mathbf{B}})^2\right)}}, \quad (5.2)$$

where $\bar{\mathbf{A}}$ is the mean of all pixel values of image **A**, and the same for $\bar{\mathbf{B}}$.



Reconstructed model; window size 15 x 15

Figure 5.2: Figure showing the results of the quality metrics for a migrated section, all calculated with a square sliding window of 15×15 pixels. The upper left plot shows the local contrast, the upper right the average line length, the lower left the merger of the former two, i.e. $\text{contrast} \times \text{line length}$, and the lower right finally shows the local SNR in dB. The data had 50% missing frequencies, and 25 reconstruction steps using the Cauchy norm.

5.2 Results

Looking at the final, migrated section (fig. 5.3) we see that choosing random frequencies for each shot already gives a result in which all the reflectors are visible. This happens because the migration algorithm, treated as a black box for this research, takes neighbouring shots into account when imaging, thus reducing aliasing noise when the frequency components are randomly chosen per shot. Note that no noise has been added to the model, all noise present is due to aliasing, which manifest as migration smiles, best seen between reflectors. For instance above the first reflector, see the red box, the migration using all frequencies shows a clean zero amplitude, whereas the non-reconstructed migrated section shows a lot of aliasing noise (top row, figure 5.4). The reconstruction does reduce the amplitude of noise in this area considerably. As can be seen in figure 5.5, the noise amplitude has a Gaussian profile over this area, and the reconstruction narrows this distribution considerably.

Weaker reflectors are, as noted above, toned down during the reconstruction. Looking at a close-up of the reservoir, the blue box in figure 5.3, we see that in the full model the reservoir has smooth edges and a strong amplitude, albeit with some internal multiples of the salt flank superimposed (fig. 5.4). The same narrowing of the amplitude spread as in the red box can be seen in the lower row of figure 5.5. This illustrates that the weak reflector is indeed weakened even further, however its bimodal structure is kept intact.

The metrics discussed previously have been calculated for reduction percentages of 10 to 90, in steps of 10, reconstructed using the Cauchy norm and 10 iterations (fig.5.6). The quality control metrics are plotted in figure 5.7. The left and middle columns are structured the same as figure 5.2, as averaged per migrated section. Top left thus shows the average contrast, bottom left the average merge-value, top middle shows the average line length, bottom middle the average SNR. Additionally the 2D correlation coefficient, top right, and total number of lines found by the edge detection, lower right, are plotted.

Contrast and the merger of contrast and line length are not very clear indicators, as the average contrast is almost constant between 10 and 80% reduction with a value of 0.085; the merger is normalised within each panel separately, making it difficult to use for comparison between images. Average line length goes down steadily, which is as expected, as the number of edges found as well as the number of lines found increase. This means that a lot of small, connected features are found in the data, which are small migration artefacts due to aliasing noise. SNR of the original data is rather low, since the areas between reflectors are almost perfectly smooth around 0, resulting in a very low mean, see the top left of fig. 5.4. Correlation shows that the reconstruction results in almost the same image, until 40% or more of the frequencies are removed, then the correlation drops drastically. Finally, the number of lines found shows a behaviour opposite to the average line length, which is as expected since more and more aliasing noise components are found to be connected.

A single reconstruction with 25 iterations and the Cauchy norm has been performed, as well as one with the $l_{1,2}$ norm and 25 iterations, see the red and black stars in figure 5.7. The correlation to the original becomes better, suggesting improved reconstruction, but more edges are found, and the SNR drops considerably below the corresponding value for 10 iterations.

Overall it can be observed that the trend across four of the six metrics shows a kink at 70 to 80% reduction, leading to the conclusion that removing more than 70% of the frequencies speeds up image deterioration considerably. This is in accordance with the number found earlier for the Q -factors (fig. 4.7). It is therefor recommended to keep at least 30% of the frequencies. Contrast and the merger of this with average line length show little trend over the range of reduction. Those metrics have therefore to be used on an image to identify difference in quality within that image, but are not suitable to use as a comparison metric between images.

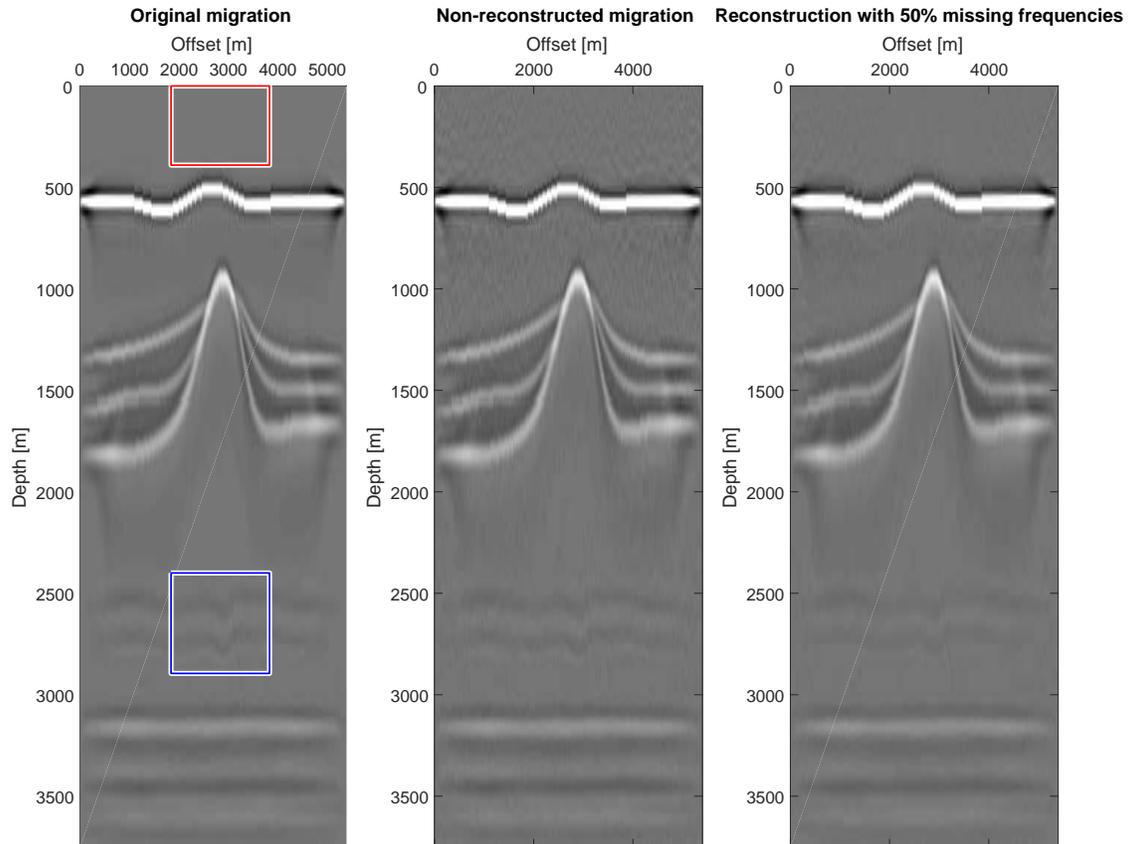


Figure 5.3: Results of the reconstruction after migration. A random 50% of frequencies has been removed for each shot, which subsequently was reconstructed using the Cauchy norm, 25 iterations and no regularisation. The left panel shows the migrated result of the full model, in the middle the migrated result after removing 50% of the frequencies without reconstructing, and on the right the migrated result after reconstruction. The red, top, and blue, bottom, boxes are areas enlarged in fig. 5.4.

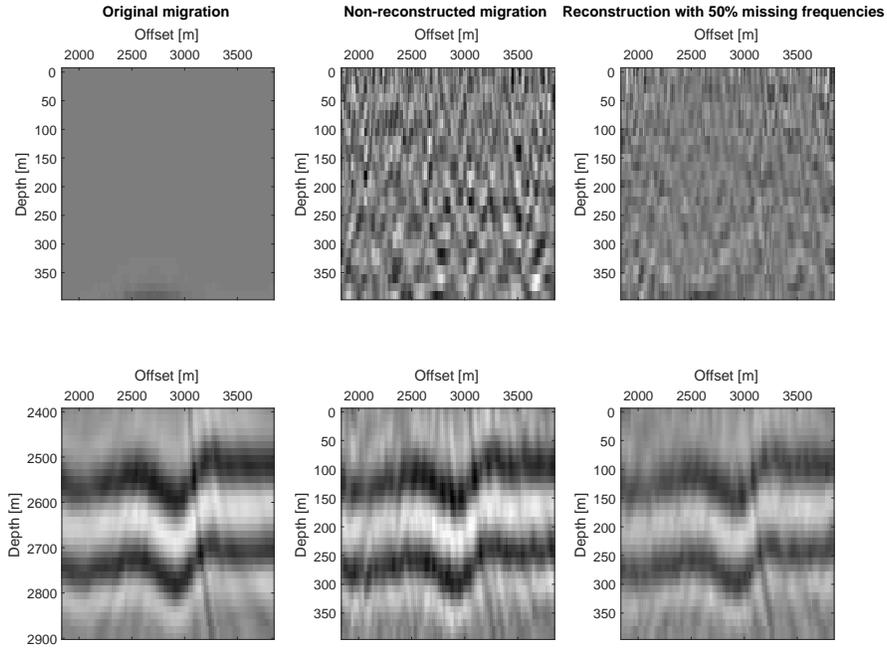


Figure 5.4: Close up of the three panels shown in figure 5.3. The top row corresponds to the red square, the bottom row to the blue square. Noise due to aliased shots in top-middle panel is clearly visible, but is reduced after the reconstruction, top-right. The reservoir has very ragged edges in the non-reconstructed migration, bottom-middle, which is corrected during the reconstruction, albeit with a loss of amplitude.

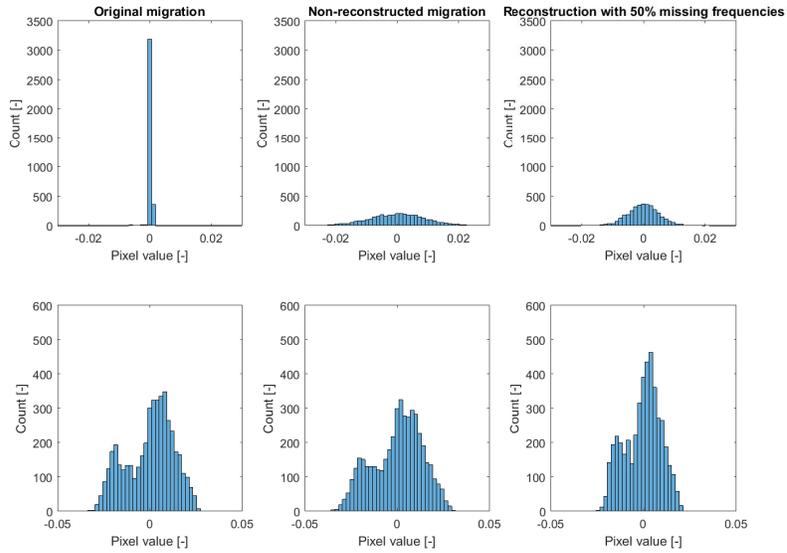


Figure 5.5: Same structure as figure 5.4, where a histogram of pixel values, thus amplitude, has been made. The top row clearly shows that the model with all frequencies has almost no noise in the layer above the first reflector, since most pixels have a value of close to 0. The aliased result shows a very broad distribution, which becomes smaller after reconstruction. The bottom row shows the same broadening of amplitude spectrum after aliasing the input, and likewise a narrowing of the distribution after reconstruction. However, as opposed to the lower-right panel of figure A.3, the bimodal nature is more pronounced here, suggesting that doing more iterations will restore the weak reflectors.

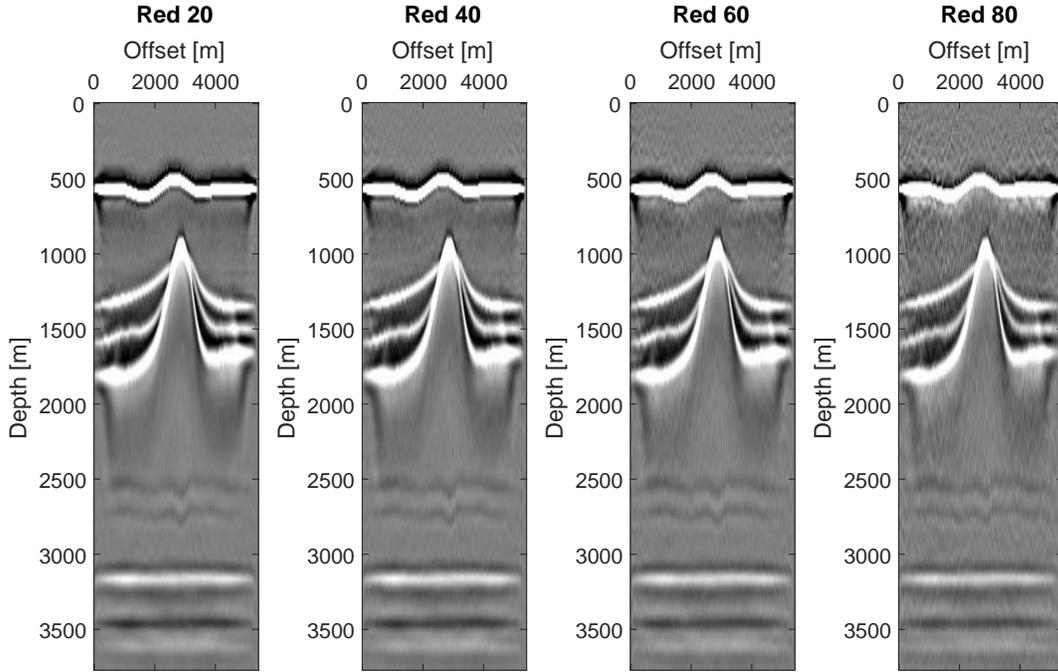


Figure 5.6: Migrated result of the reconstructed shots; the title of each subplot shows the percentage of missing frequencies. All features remain visible up to 80% reduction, but noise increases towards higher reduction.

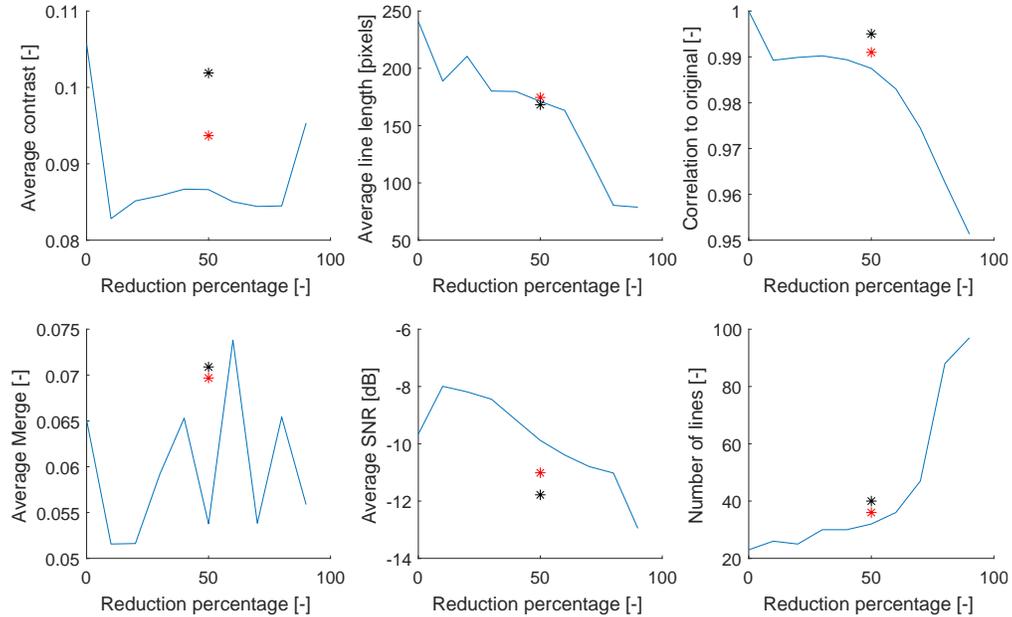


Figure 5.7: Plot showing several quality indicators as a function of frequency reduction percentage. The left and middle columns are the same as fig. 5.2, the right most column shows the correlation to the original on top, and the number of lines as found in the edge detection below. All reconstructions have been made with the Cauchy norm and 10 iterations, except for the red star, that has 25 iterations, and the black star which uses the $l_{1,2}$ norm and 25 iterations.

Chapter 6

Conclusion, discussion and recommendations

In order to reduce cost a lot of research in onshore seismology is conducted to reduce the amount of data required to obtain good images. This work attempts to reduce the number of frequency components present in the data, thereby reducing the total amount of data. First a PCA has been performed on a general, open source data set as presented by [Yilmaz et al. \[1987\]](#), which showed that indeed a large redundancy is present in seismic data, as confirmed by a different onshore data set from Poland, as well as a synthetic example (chapter 2).

An attempt has been made at reconstruction of the data using the Fourier transform (chapter 3), but the Fourier transform lacks connectivity in the spatial direction, so it has been disregarded. Instead, the linear Radon transformation has been chosen (chapter 4). A conjugate gradient scheme has been set up (sec. 4.3) to reconstruct iteratively the shots in an attempt to reduce aliasing. Various percentages of reduction in the frequency domain, as well as a varying number of iterations and different norms have been tested. As expected, a higher percentage of reduced frequencies results in a worse reconstruction (tab. 4.3, fig. 4.7) of the shot records. Gaps of a band of frequencies can be well reconstructed up to some 6Hz (fig. 4.10), and start deteriorating at larger gaps.

Contrast, line length, correlation to the original and SNR have been proposed as quality metrics (chapter 5). It can be seen that the reconstruction reduces aliasing noise in the final migrated section, although it also lowers the amplitude of weak reflectors (fig. 5.3). More research will have to be done in parameter selection, so that noise is attenuated but weak reflectors are retained. When more than 70% of the frequencies are removed, image deterioration accelerates (fig. 5.7). This is in accordance with the result of the Q -factors (fig. 4.7), thus that is interpreted as a threshold for the current procedure.

Answers to the subquestion as posed in section 1.1 are given (sec. 6.1), after which the results of the thesis will be discussed and recommendation for further research will be made (sec. 6.2).

6.1 Answers to research questions

1. *Is there redundancy in the frequency domain of the data? Thus can less frequencies be used whilst still obtaining a 'good' result?*

Yes. Chapter 2 shows that 80 to 90% of the data is carried by only five principal components (fig.2.4). Furthermore, results of reconstruction using fewer frequencies show a good result up to 75% reduction in the frequency domain (tab. 4.3 and fig. 5.7).

2. *What frequency sampling is best if only few frequencies are to be measured?*

No method has been implemented to randomly draw from a user-defined function based on the PCA results. Looking at reconstructed shots only, there is no clear preference in the PCA

frequencies or the random frequencies, but using different frequencies per shot definitely helps, as the migration algorithm then further suppresses aliasing through connecting shots.

3. *Which transformation yields the best results for data with missing frequency components?*

The Fourier transform has shown to be lacking in spatial continuity. Since that is proposed to compensate for the missing frequency components, this transform has been discarded. Instead, the linear Radon transformation has been used, which shows a good spatial connectivity.

4. *How well can we reconstruct the dense, complete data set from the sparse frequency data to allow usage of the currently available processing chain?*

- Energy gets refocused, especially along the asymptots of hyperbolas (fig. 4.5).
- Aliasing noise in the final image is not too broad, with weak reflectors being recovered at 80% reduction in frequencies, provided no noise is present in the shot records (fig. 5.6).
- Gaps in frequency get filled in completely up to a gap of 6Hz, for larger gaps the high-offsets start deteriorating (fig. 4.10).

5. *What metrics can be used to determine image quality? Thus what is a ‘good’ result in quantitative measures to be used under question 1.*

Currently the correlation to a ground truth, percentage of edges as found in the data, the average number of pixels per connected component of edges, the total number of connected components found and the average SNR are used to get quantitative information on the data (fig. 5.2 and 5.7). This list is certainly not exhaustive, and may benefit from image processing research in general.

The main question of this thesis is:

Is it possible to reduce the number of required frequencies in onshore seismology by increasing the spatial density of measurement?

Yes, this is possible, using the method proposed in this thesis about 70% of the frequencies can be removed from a seismic data set whilst still obtaining a good image.

6.2 Recommendations

Instead of the linear Radon transform, using a hyperbolic Radon transform would likely increase refocussing along the apices of hyperbolas. Nikitin et al. [2017] present a fast implementation of the hyperbolic transform and its Hermitian conjugate, with the CUDA source publicly available. Restructuring the algorithm given in section 4.3 to use this transform will likely improve reconstruction of the shots themselves.

The migration algorithm, treated as a black box for this work, takes into account neighbouring shots for each reconstruction point, thus reducing aliasing in the migrated section, provided the aliasing is not the same per shot. Thus extending the algorithm from reconstructing shots to fully migrating the data, taking into account neighbouring shots, for example based on a distance weighted correlation function, will likely result in less aliasing noise leaking into the final migrated section.

CG schemes are well known and widely used in tomography, since they do not require an explicit matrix form. For example, as shown in section 4.2, \mathbf{A} denotes a conjugated Radon transform, which in the numerical implementation is not expressed as a matrix, but as a function. It has been suggested that CG optimisation of tomography problems might be problematic for certain cases since large systems require a good preconditioner (Zang et al. [2018]). But precisely because in that case \mathbf{A} becomes so large, it is implemented as a function, limiting the choice of possible preconditioners. Another problem is that \mathbf{A}^H is not the exact Hermitian conjugate of \mathbf{A} , rather a function which does the conjugated of what \mathbf{A} does. Due to numerical noise and/or discrepancies in implementation between the forward

and inverse operator they are not each others exact Hermitian conjugate, thus introducing errors in each solver iteration. [Zang et al. \[2018\]](#) propose to use the Simultaneous Algebraic Reconstruction Technique (SART) to circumvent these issues.

IRLS or Iteratively Reweighted Least Squares, can be used to stabilise the CG scheme when it is prone to end in local minima instead of the global minimum, which is often the case when searching for sparse solutions. This entails a reset of the \mathbf{r} and \mathbf{d} vectors every now and then to reset the search for a minimum. It can also be used to reset the weighting function every few iterations, which in the present case might result in a better recovery of weak reflectors by reducing aliasing using CG, then recalculating the norm with the lowered noise level and running the CG again. An outline of a possible implementation:

$$\begin{aligned}
& i = 0 \\
& \tilde{\mathbf{p}}_0 = \mathbf{0} \\
& \mathbf{C}_{\mathbf{p}_{0,jj}} = \left((\mathbf{A}^H \mathbf{p}_j)^* \mathbf{A}^H \mathbf{p}_j + \sigma_{\mathbf{A}^H \mathbf{p}}^2 \right)^a \\
& \text{while } i \leq i_{\max} \\
& \quad k = 0 \\
& \quad \mathbf{z}_k = \tilde{\mathbf{p}}_i \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{1/2} \\
& \quad \mathbf{r}_k = \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{-1/2} \mathbf{A}^H \mathbf{p} \\
& \quad \mathbf{d}_k = \mathbf{r}_k \\
& \quad \mathbf{t}_k = \mathbf{p} \\
& \quad \text{while } \frac{\mathbf{r}_k^H \mathbf{r}_k}{\mathbf{r}_0^H \mathbf{r}_0} > \varepsilon, k \leq k_{\max} \\
& \quad \quad \alpha = \frac{\mathbf{r}_k^H \mathbf{r}_k}{\mathbf{d}_k^H \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{-1/2} \mathbf{A}^H \mathcal{F}^{-1} \mathbf{S}_{\odot} \mathcal{F} \mathbf{A} \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{-1/2} \mathbf{d}_k + \mathbf{d}_k^H \lambda \mathbf{Id}_k} \\
& \quad \quad \mathbf{z}_{k+1} = \mathbf{z}_k + \alpha \mathbf{d}_k \\
& \quad \quad \mathbf{t}_{k+1} = \mathbf{t}_k - \alpha \mathcal{F}^{-1} \mathbf{S}_{\odot} \mathcal{F} \mathbf{A} \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{-1/2} \mathbf{d}_k \\
& \quad \quad \mathbf{r}_{k+1} = \mathbf{C}_{\tilde{\mathbf{p}}_{i,jj}}^{-1/2} \mathbf{A}^H \mathbf{t}_{k+1} + \alpha \lambda \mathbf{Id}_k \\
& \quad \quad \beta = \frac{\mathbf{r}_{k+1}^H \mathbf{r}_{k+1}}{\mathbf{r}_k^H \mathbf{r}_k} \\
& \quad \quad \mathbf{d}_{k+1} = \mathbf{r}_{k+1} + \beta \mathbf{d}_k \\
& \quad \quad k = k + 1 \\
& \quad \tilde{\mathbf{p}}_{i+1} = \mathbf{C}_{\tilde{\mathbf{p}}_{jj}}^{-1/2} \mathbf{z}_k \\
& \quad \mathbf{C}_{\tilde{\mathbf{p}}_{i+1,jj}} = (\tilde{\mathbf{p}}_{i+1,j}^* \tilde{\mathbf{p}}_{i+1,j} + \sigma_{\tilde{\mathbf{p}}}^2)^a \\
& \quad i = i + 1
\end{aligned}$$

where the parameters are the same as in the CGNE algorithm presented in section 4.3.

Figures 6.1, 6.2, and 6.3 illustrate well the trade-off between the $l_{1,2}$ norm on one hand and the Cauchy norm on the other. Looking at the top rows of figures 6.2, and 6.3 one sees that noise is more attenuated when using the more aggressive Cauchy norm, but this also reduces the amplitude for the true reflecting layers, see the bottom row. A study varying the number of iterations and preferably using IRLS could determine which norm is preferred for usage. Result may well be that using the Cauchy norm for a quick result is optimal, and the $l_{1,2}$ norm being preferred when accuracy is the main focus, or even a couple of iterations with a more aggressive norm to quickly reduce a lot of noise and then switching to a less aggressive norm to fine-tune the weak events.

Quantifying the drop in amplitude in the band gap, see fig. 4.10, could help in the reconstruction, if the amplitudes are multiplied with a pre-calculated constant such that it will be the same as outside the band gap. It has to be investigated what this amplitude depends on; size of the band gap seems irrelevant judging figure 4.10, but number of iterations and norm used might still influence it.

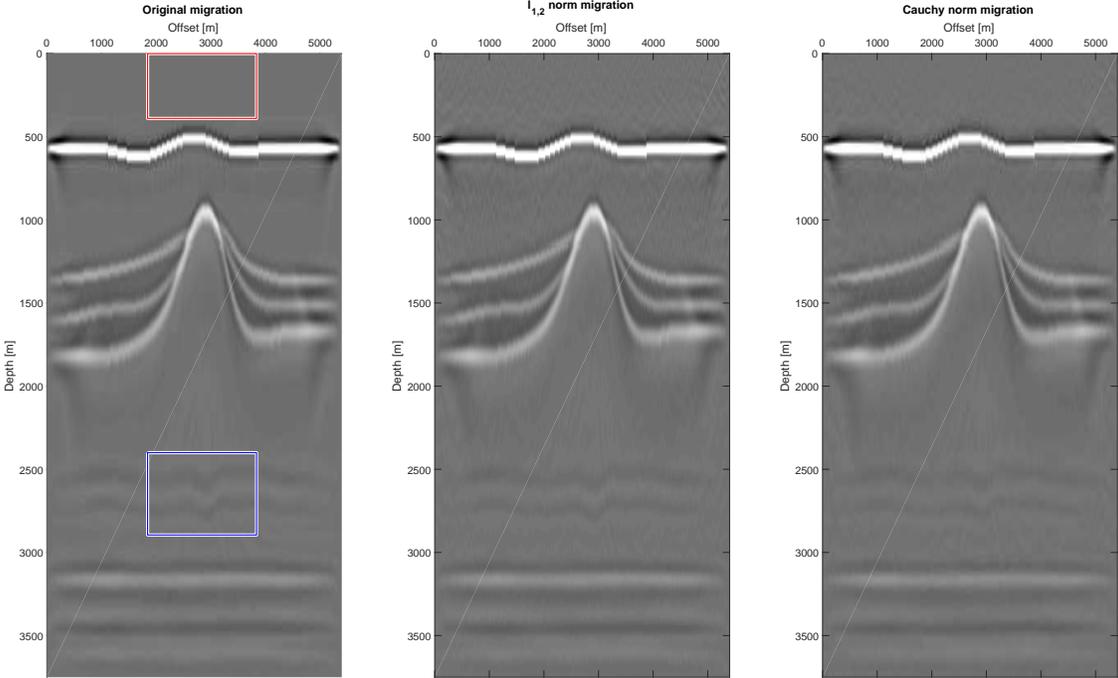


Figure 6.1: Results of the reconstruction after migration. A random 50% of frequencies has been removed for each shot, which subsequently was reconstructed using either the $l_{1,2}$ or Cauchy norm, using 25 iterations and no regularisation. The left panel shows the migrated result of the full model, in the middle with $l_{1,2}$ norm, and the right with Cauchy. The red, top, and blue, bottom, boxes are areas enlarged in fig. 6.2.

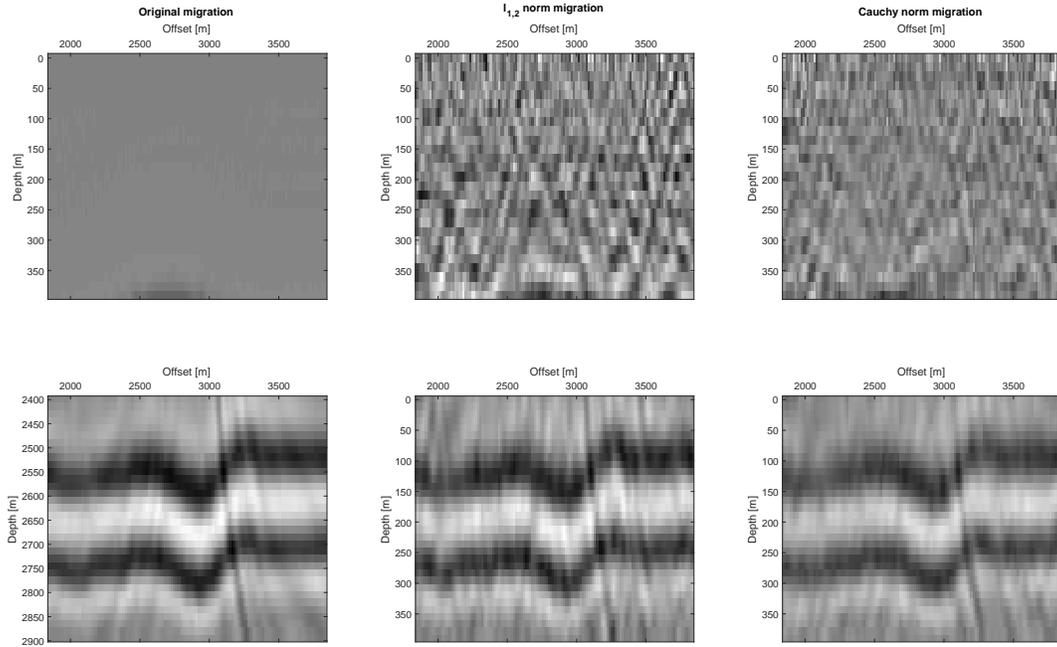


Figure 6.2: Close up of the three panels shown in figure 6.1. The top row corresponds to the red square, the bottom row to the blue square. Noise due to aliased shots in top-middle panel is clearly visible, and is attenuated using the Cauchy norm, top-right. The reservoir is slightly more noisy in the $l_{1,2}$ norm, bottom middle, than in the original, bottom left, but at least stronger in amplitude than when using the Cauchy norm, bottom right.

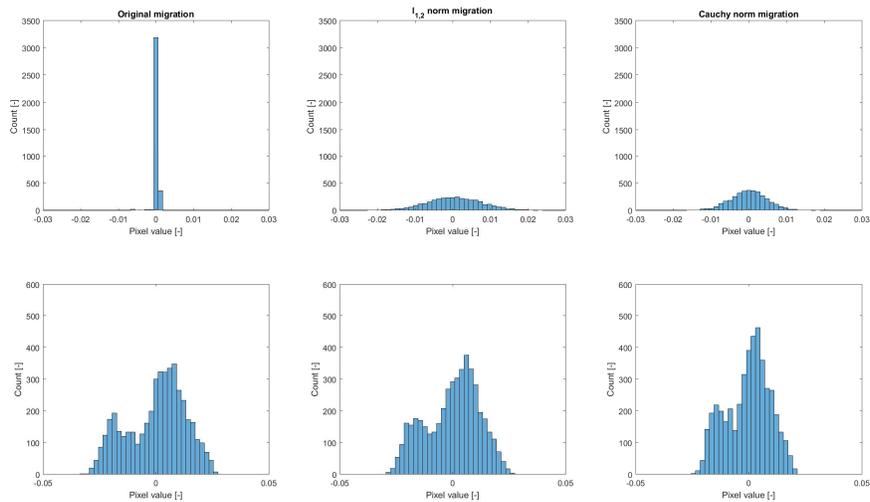


Figure 6.3: Same structure as figure 6.2, where a histogram of pixel values, thus amplitude, has been made. The top row clearly shows that the model with all frequencies has almost no noise in the layer above the first reflector, since most pixels have a value of close to 0. Both norms show a broad, Gaussian distribution of amplitudes, albeit the Cauchy norm has the smaller variance. The bottom row shows that the Cauchy norm also limits the spread in amplitude values, whereas the $l_{1,2}$ norm is a better match in terms of width and height of the two bimodal peaks.

Bibliography

- J. Alcalde, C. E. Bond, G. Johnson, J. F. Ellis, and R. W. Butler. Impact of seismic image quality on fault interpretation uncertainty. *GSA Today*, 2017.
- J. Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.
- Z. Cao, J. Bancroft, R. Brown, and C. Xaio. Radon transform and multiple attenuation. Technical report, CREWES Research Report, 2003.
- R. C. de Lucena Rêgo, O. A. de Lima Neto, and C. C. N. da Silva. Conventional seismic processing flow analysis: 2d line of the rio grande chain region southeastern brazil. In *15th International Congress of the Brazilian Geophysical Society & EXPOGEF, Rio de Janeiro, Brazil, 31 July-3 August 2017*, pages 1507–1512. Brazilian Geophysical Society, 2017.
- A. Duijndam and M. Schonewille. Nonuniform fast fourier transform. *Geophysics*, 64(2):539–551, 1999.
- S. Fomel. *Three-dimensional seismic data regularization*. PhD thesis, Stanford University, 2000.
- Geofizyka Torun S.A, Poland. 2d vibroseis line 001. Wiki entry, open data, 2005. URL https://wiki.seg.org/wiki/2D_Vibroseis_Line_001.
- J. Kaipio and E. Somersalo. Statistical and computational inverse problems. *Springer Verlag*, 2005.
- A. Kumar. *3-D seismic acquisition geometry design and analysis: Investigation of the requirements to include illumination from all multiples*. PhD thesis, TU Delft, Delft University of Technology, 2015.
- D. C. Lay, J. J. McDonald, and S. R. Lay. *Linear Algebra and Its Applications*. Pearson, 2016.
- J. Meunier. *Seismic acquisition from yesterday to tomorrow*. Society of Exploration Geophysicists, 2011.
- V. V. Nikitin, F. Andersson, M. Carlsson, and A. A. Duchkov. Fast hyperbolic radon transform represented as convolutions in log-polar coordinates. *Computers & Geosciences*, 105:21–33, 2017.
- J. Radon. über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, 69:262–277, 1917.
- J. Romberg. Compressive sensing by random convolution. *SIAM Journal on Imaging Sciences*, 2(4):1098–1128, 2009.
- J. A. Scales, A. Gersztenkorn, and S. Treitel. Fast l_p solution of large, sparse, linear systems: Application to seismic travel time tomography. *Journal of Computational Physics*, 75(2):314–333, 1988.
- Schlumberger Oilfield Glossary. Ground roll. Wiki entry, 2018. URL https://www.glossary.oilfield.slb.com/Terms/g/ground_roll.aspx.

- B. W. Schulte and D. Manthei. Chasing density – an introduction to seismic acquisition, processing, and interpretation methods leading to quantitative interpretation. *Recorder*, 39(08), 2014.
- J. R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain, 1994.
- Society of Exploration Geophysicists. Bandwidth and vertical resolution. Wiki entry, 2018. URL https://www.wiki.seg.org/wiki/Bandwidth_and_vertical_resolution.
- N. Tellier and J. Lainé. Understanding mems-based digital seismic sensors. *First Break*, 35(1):93–100, 2017.
- The MathWorks. Label and measure objects in a binary image. Documentation, 2018. URL <https://mathworks.com/help/images/label-and-measure-objects-in-a-binary-image.html>.
- G. J. Vermeer. Creating image gathers in the absence of proper common-offset gathers. *Exploration Geophysics*, 29(3/4):636–642, 1998.
- G. J. Vermeer. *3-D seismic survey design*. Society of Exploration Geophysicists, 2002.
- D. J. Verschuur. *Seismic multiple removal techniques: past, present and future*. EAGE publications, 2006.
- A. H. Visser. The impact of spatial sampling and signal-to-noise ratio on image quality in onshore seismics, 2017.
- O. Yilmaz, R. Chambers, D. Nichols, and R. Abma. Fundamentals of 3-d migration, part 1. *The Leading Edge*, 6(11):22–30, 1987.
- G. Zang, M. Aly, R. Idoughi, P. Wonka, and W. Heidrich. Super-resolution and sparse view ct reconstruction. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 137–153, 2018.
- P. M. Zwartjes. *Fourier reconstruction with sparse inversion*. PhD thesis, TU Delft, Delft University of Technology, 2005.

Acknowledgements

I would like to thank Prabu, Eric and Ramon for their great support. Not only during the project, but also their enthusiasm in setting up this two-faculty project, so that I would not have to write two theses. Thank you for all the long discussions and input, I hugely appreciated your door being always open.

Thanks to Jim, Joost, Junhai, Léon, Ronald, Edo and Henry for help in programming and technical issues. It's a shame you graduated six months earlier than I did Joost, I did miss the long ramblings about Fourier spaces. Thank you Robert for doing the artistic work with the cover, as well as printing the thesis.

Thanks to Andrés, Ander, Cris, and the others over in the *CHATLAB and TALKTAVE* room of Stack Overflow. I learned more about MATLAB there than in the rest of my studies, and I also enjoyed the discussions on all subjects of life, and physics and mathematics in particular.

Pa, ma, bedankt voor de ondersteuning gedurende deze 8,5 jaar durende studie. Ik kon altijd bellen of langskomen als er iets was. Zonder jullie steun, niet in het minst financieel, had ik dit niet kunnen doen. Hartelijk dank voor de controle van mijn scriptie pa. Als laatste wil ik Anne bedanken. Anne, jij was er voor me. Als ik weer eens vast zat in een algoritme, beurde jij me op en spoorde je me aan om door te gaan, maar in de weekeinden mocht ik onder geen beding aan de studie, want ik moest ook ontspannen. Bedankt voor alle fijne avonden, fiets- en wandeltochten en nog veel meer.

The cover depicts the correlation matrix of the frequency data which served as input for the PCA in figure 2.4. The structure of the matrix was so appealing, that I decided to copy it four times, so all quadrants would be symmetric, and attached a nonsensical colour map which at least looks pretty. If science does not pan out, I can always go into carpet-designs.

Appendices

Appendix A

Additional tables and plots

A.1 Yilmaz data

Table A.1: Shot records as given by [Yilmaz et al. \[1987\]](#). Other information such as the survey location and equipment numbers have been omitted as they are not used in this research. The only four shot records with trace separation mentioned are: 1 (330ft), 10 (25m), 11 (220ft), and 12 (330ft).

Number of shot	Traces	Sampling [ms]	Land/Marine
1	48	4	Land
2	120	4	Land
3	24	4	Marine
4	48	4	Land
5	48	2	Marine
6	48	4	Marine
7	48	2	Land
8	96	4	Marine
9	96	4	Land
10	120	4	Land
11	96	4	Land
12	49	4	Land
13	48	4	Marine
14	48	4	Land
15	48	2	Marine
16	48	4	Marine
17	48	4	Land
18	120	4	Marine
19	216	4	Land
20	120	4	Land
21	48	2	Land
22	48	2	Land
23	48	2	Marine
24	48	2	Land
25	96	2	Marine
26	48	2	Marine
27	92	4	Land
28	48	2	Marine
29	48	4	Land
30	96	4	Land
31	96	4	Land
32	96	4	Land
33	96	4	Marine
34	60	4	Land
35	52	2	Marine
36	48	4	Marine
37	48	4	Land
38	60	2	Marine
39	60	2	Land
40	96	2	Land

A.2 Additional plots

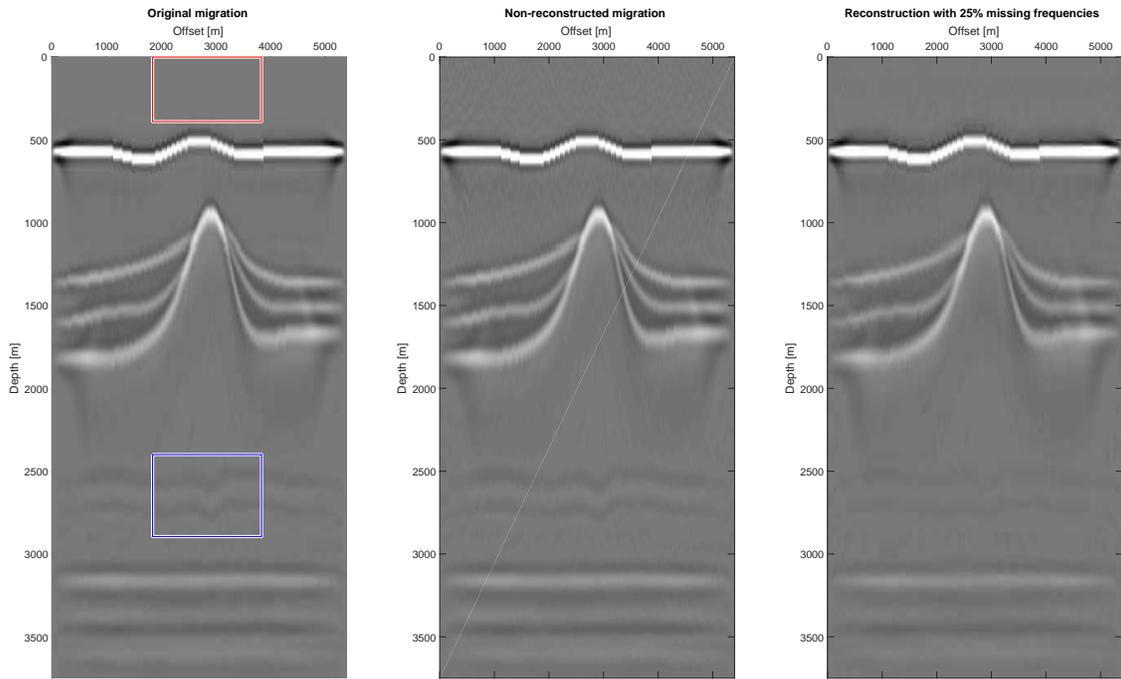


Figure A.1: Results of the reconstruction after migration. A random 25% of frequencies has been removed for each shot, which subsequently was reconstructed using the Cauchy norm, 10 iterations and no regularisation. The left panel shows the migrated result of the full model, in the middle the migrated result after removing 25% of the frequencies without reconstructing, and on the right the migrated result after reconstruction. The red, top, and blue, bottom, boxes are areas enlarged in fig. A.2.

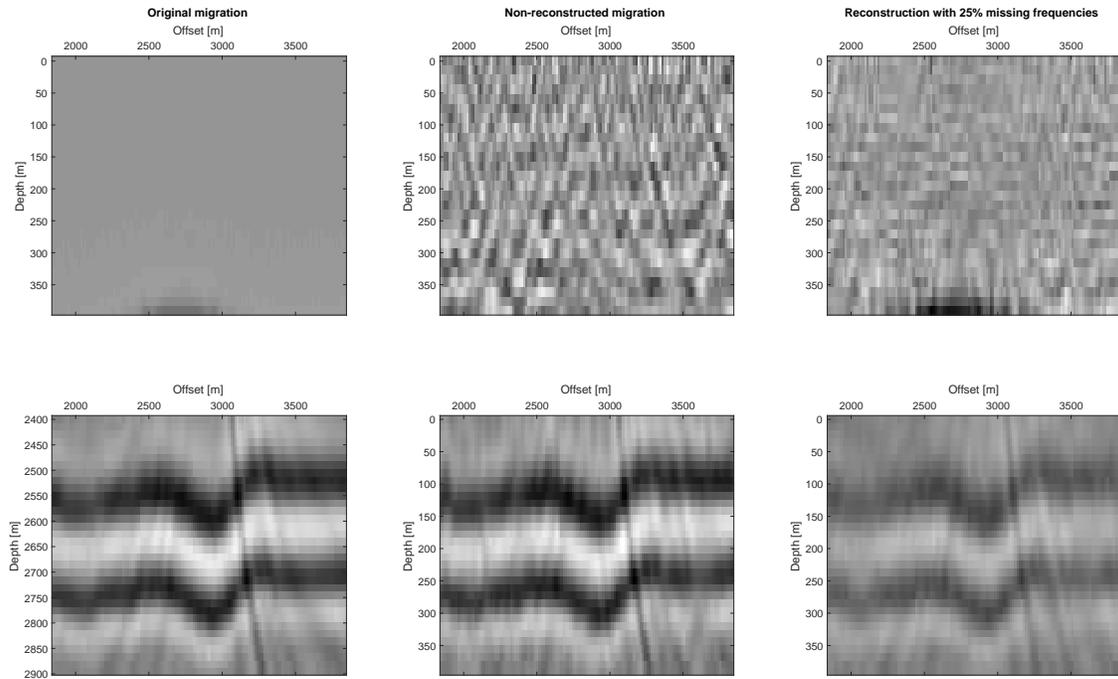


Figure A.2: Close up of the three panels shown in figure A.1. The top row corresponds to the red square, the bottom row to the blue square. Noise due to aliased shots in top-middle panel is clearly visible, but is reduced after the reconstruction, top-right. The reservoir has very ragged edges in the non-reconstructed migration, bottom-middle, which is corrected during the reconstruction, albeit with a loss of amplitude.

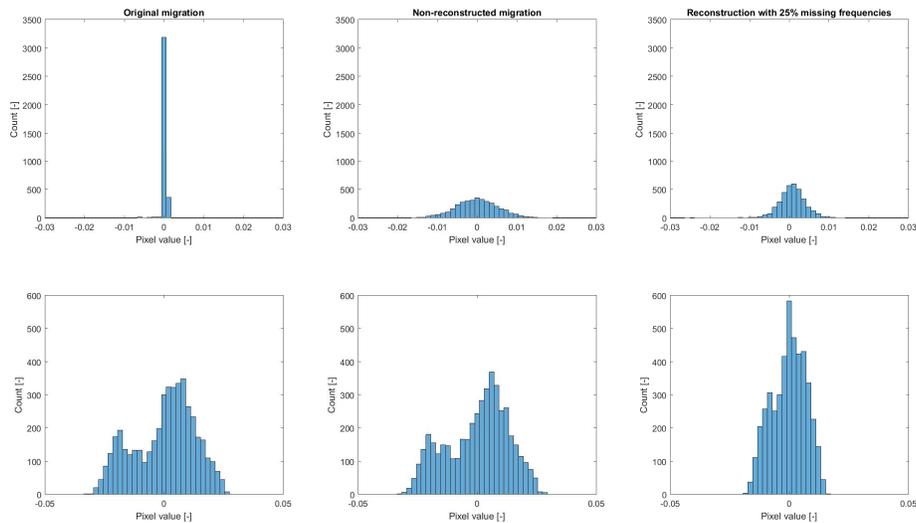


Figure A.3: Same structure as figure A.2, where a histogram of pixel values, thus amplitude, has been made. The top row clearly shows that the model with all frequencies has almost no noise in the layer above the first reflector, since most pixels have a value of close to 0. The aliased result shows a very broad distribution, which becomes smaller after reconstruction. The bottom row shows the same broadening of amplitude spectrum after aliasing the input, and likewise a narrowing of the distribution after reconstruction. This means that weak reflectors are suppressed using the given parameters of reconstruction.

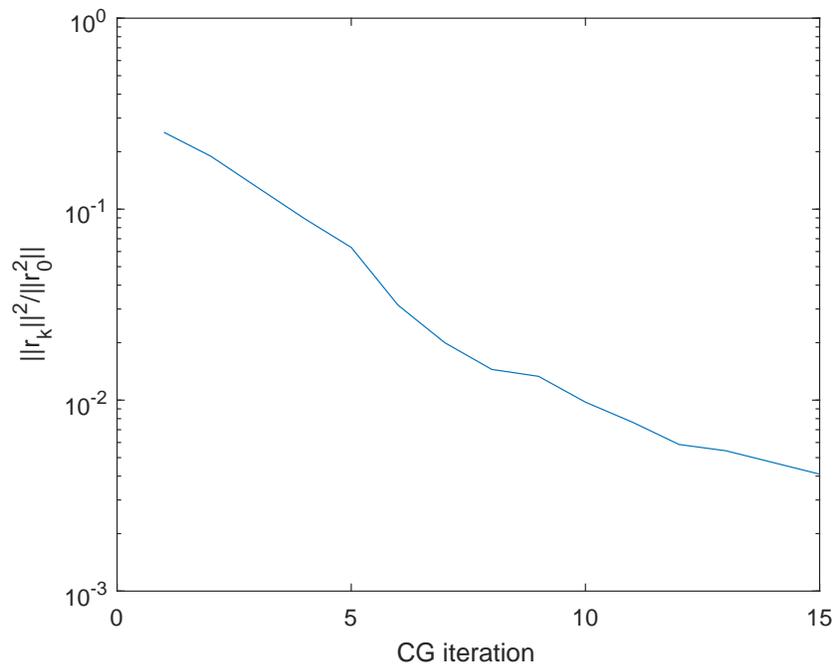


Figure A.4: The squared norm of each CGNE iteration for the shot in figure 4.5. It is monotonically decreasing.