IDEA League

Master of Science in Applied Geophysics

Research Thesis

Nonlinear beamforming seismic data reconstruction: A novel kinematic wavefront based method

Quentin Hirsch

August 5, 2022

Nonlinear beamforming seismic data reconstruction: A novel kinematic wavefront based method

Master of Science Thesis

for the degree of Master of Science in Applied Geophysics at
Delft University of Technology
ETH Zürich
RWTH Aachen University
by

Quentin Hirsch

August 5, 2022

Department of Geoscience & Engineering · Delft University of Technology
Department of Earth Sciences · ETH Zürich
Faculty of Georesources and Material Engineering · RWTH Aachen University

Delft University of Technology

Printed in The Netherlands

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *August 5, 2022*

Supervisor(s):

Dr. Ir. G.G.Drijkoningen

Dr. Yimin Sun

Committee Members:

Dr. Ir. G.G.Drijkoningen

Dr. Yimin Sun

Prof. Dr. Florian Wellman

Dr. Ir. D.S. Draganov

Abstract

In 3D seismic data acquisition, sufficiently dense spatial sampling is most often not possible, because of physical, financial, and temporal constraints. The resulting aliased data is a major obstacle for accurate subsurface images. Therefore, methods are needed to reconstruct these sparse data to an adequate sampling. We develop a novel kinematic wavefront-based seismic data reconstruction method that uses the existing nonlinear beamforming (NLBF) framework, which allows for building a more detailed sampling from a sparse input. We present the theory and methodology of our NLBF reconstruction algorithm and test it on a synthetic Society of Exploration Geophysicists Advanced Modeling (SEAM) Arid dataset and a field dataset. We attempt to answer the following question: can the NLBF framework be used for seismic data reconstruction, and how do these results compare to some conventional reconstruction methods? Control parameter tests are performed to find the optimal NLBF reconstruction, and results are compared to those from several control methods including the convergent alternating projection onto convex sets (POCS) and bootstrap POCS methods, which were also developed for this study. Our NLBF reconstruction method successfully reconstructs high-quality data on both datasets. We find that the optimal NLBF control parameters are ultimately dataset-dependent as the concrete data acquisition geometry plays a central role. Specifically for our datasets, optimal parameters include a time window of $12\Delta t_s$, an operator aperture of 600 m x 600 m, and a parameter trace interval of $\Delta x = \Delta y = 60$ m. The synthetic SEAM Arid NLBF reconstruction results show high trace fidelity to the ground truth. The field data NLBF results show better-reconstructed gaps, and wavefield variations closer to those of physical propagating waves in comparison to our control methods. These results show the potential of the NLBF reconstruction method to become a common data reconstruction tool used in the seismic industry.

Table of Contents

Acronyms

$\mathbf{A}_{est}$ Estimation Aperture

$\mathbf{A}_{op}$ Operator Aperture

$\mathbf{A}_{sum}$ Summation Aperture

APOCS Alternating Projections Onto Convex Sets

BP Bootstrap Alternating Projections Onto Convex Sets

CP Convergent Alternating Projections Onto Convex Sets

CRS Common Reflection Surface

DMO Dip-Moveout

eGA Efficiency-improved Genetic Algorithm

$f-k$ Frequency-Wavenumber

GA Genetic Algorithm

LSQM Least-Squares Migration

NLBF Nonlinear Beamforming

NMO Normal Moveout

OO Operator Oriented

PEF Prediction Error Filter

POCS Projections Onto Convex Sets

SEAM SEG Advanced Modeling

SP Single-Pass Alternating Projections Onto Convex Sets

SNR Signal-to-Noise Ratio

TO Target Oriented

$x - t$ Space - Time

Chapter 1

Introduction

In this chapter, we first present the necessary background to understand the purpose, the acquisition, and the challenges of seismic data for subsurface imaging. Then, we outline several mathematical and geophysical frameworks in existing literature created to deal with the data sparsity challenge. Finally, we reveal how we attempt to tackle it. Here, we introduce our research question and present our outline.

1-1 Seismic data acquisition

Imaging the earth's subsurface is necessary to understand underlying geological settings, especially for the exploration of sedimentary basins with the potential to host natural resources such as oil and gas. Seismic data acquisition is based on the physics of wave propagation in the earth [Yilmaz, 2001].

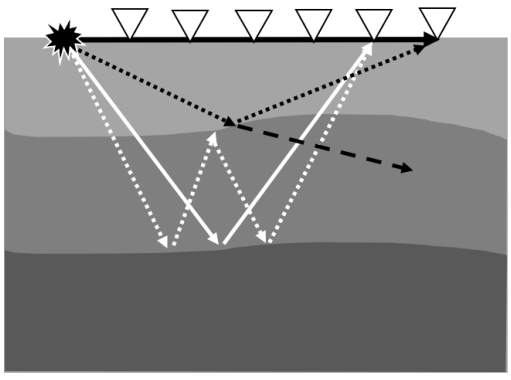

Figure 1-1: A 2D seismic data acquisition illustration, with a source (black star) and receivers (triangles). The black arrows are the direct (solid), the reflected (dotted), and the refracted (dashed) wavefronts; the white arrows are the primary (solid) and the multiple (dotted) reflections.

The acquisition of seismic data can be performed on land or in a marine setting with their respective instruments and setups. A source, such as an airgun for the marine setting or a seismic vibrator on land, is used to emit a downgoing wavefield into the subsurface. Figure 1-1 illustrates some of the energy's possible propagation directions. The wavefield that propagates directly from a source to a receiver, is the direct wavefield (the solid black line). The wavefield that propagates into the subsurface may be reflected upwards (the dotted black line), or refracted downwards (the dashed black line) at interfaces according to Snell's law [Wapenaar, 2014]. These subsurface interfaces are present between rocks with contrasts in physical properties that affect the wavefront's behavior. At the surface, the wavefields are measured by receivers and are stored in a seismic data format for processing. To image subsurface reflectors, surface-recorded wavefields are backward-propagated into the subsurface via the wave propagation theory, where they interact with forward-simulated wavefields to reveal buried reflectors. Seismic imaging works in practice but also presents inherent issues hindering the quality and accuracy of the target subsurface model [Kutscha, 2014]. These can be divided into subsurface-based problems and acquisition-based problems:

- The subsurface is a complex environment and the wavefield may behave differently than our simple model's predictions. Near-surface environments are known to be weathered which creates scattering noise and can distort the wavefield. Subsurface structures such as high impedance salt domes are major obstacles that limit wave propagation [Xie and Gajewski, 2017]. Furthermore, receivers capture many signals at the surface, including primary reflections, multiples, surface waves, noise, etc. This results in information that can not be dealt with by certain imaging methods. Pre-processing of seismic data is necessary to overcome these limitations.

- The acquisition setup is also important as it controls the amount of collected data and the survey's coverage [Qin et al., 2018; Zhang et al., 2021]. Adequate subsurface coverage must be obtained to create accurate images, which implies fine wavefield sampling in both time and space. Temporal sampling is rarely an issue with modern receivers. However, spatial sampling is usually too coarse to create well-resolved images, resulting in aliasing. Coarse spatial sampling is primarily financially driven, as additional equipment, manpower, and processing power are required for denser data acquisition.

To obtain accurate images, the missing traces in these coarse seismic gathers can be reconstructed during data preprocessing [Kutscha, 2014; Xie and Gajewski, 2017]. This thesis investigates the reconstruction of such sparse seismic data using the nonlinear beamforming (NLBF) framework, to be discussed in Chapter 3. We define data sparsity as the relative fraction of missing seismic data traces. In this chapter, we introduce the fundamental issue of seismic data sparsity and existing methods that attempt to reconstruct such sparse data.

1-2 Data sparsity challenge

The challenge faced in seismic data acquisition is ensuring that the survey area is sufficiently covered by data traces. The wavefield must be uniformly and densely sampled in space and time to fully characterize the subsurface. Sparse signal sampling can lead to aliasing

and erroneous signal representations [Kutscha, 2014]. Figure 1-2 illustrates the principle of aliasing in both time and space. In seismic imaging, this causes strong artifacts and unwanted noise [Yilmaz, 2001].

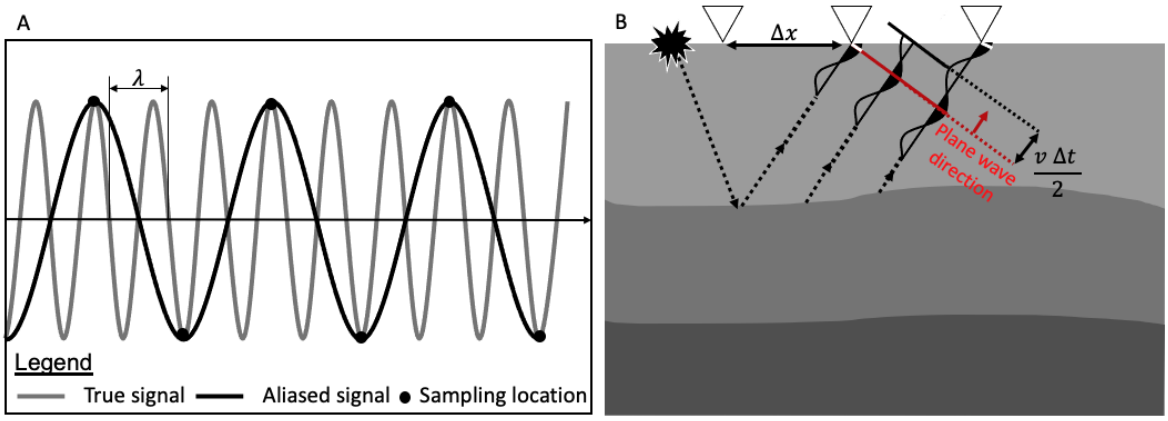

Figure 1-2: Illustration of temporal and spatial aliasing in seismic data acquisition. (A) Temporal aliasing [1]: coarse sampling (black dots) of the true signal (gray), reconstructs an aliased erroneous signal (black). (B) Spatial aliasing [2]: receiver spacing is too coarse to adequately sample the wavefield in space.

According to the Nyquist criterion, the signal must be sampled at least twice per wavelength to be fully characterized [Geldart and Sheriff, 2004]. Sampling the wavefield in time at the Nyquist[1] frequency is achievable with the use of modern sampling instruments. In fact, typical sources for hydrocarbon exploration emit up to 150 Hz which receivers can uniformly sample [Kutscha, 2014]. For spatial sampling, in 3D seismic surveys the wavefield is sampled along four spatial dimensions: the respective x-y coordinates of source and receiver pairs (Figure 1-3). Seismic data are rarely densely [2] sampled in all directions because of economic and time-driven constraints [Kutscha, 2014; Zwartjes and Gisolf, 2007]. In addition, spatial sampling issues can result from equipment failures, human errors, and geographic inaccessibility. Figure 1-4 illustrates coarse spatial sampling and missing traces on preprocessed seismic data, leading to discontinuous and poor quality reflections. Such gaps in coarse or non-uniformly sampled data cause erroneous subsurface images [Kutscha, 2014; Zwartjes and Gisolf, 2007].

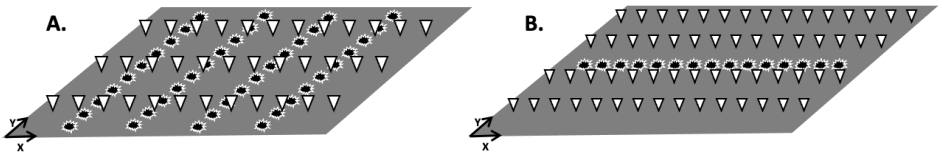

Figure 1-3: Example seismic data acquisition geometries: black stars represent the sources and white triangles receivers. (A) A typical land acquisition, where sources and receivers are densely sampled along different directions. (B) A typical marine acquisition, where sources and receivers are densely sampled in one direction.

[1] The sampling theorem: a continuous signal can be reconstructed if $f_s \geq \frac{1}{2f_{max}}$. [Geldart and Sheriff, 2004]
[2] Seismic spatial sampling criterion is $\Delta x = \frac{v}{2 \sin \Theta}$, in practice rarely achieved. [Geldart and Sheriff, 2004]

The acquisition geometry is a key factor in the quality of acquired seismic data. Dense and regularly positioned sources and receivers on a uniform grid offer the best resolution but are not economically viable. The common land setup places densely sampled sources in the inline dimension and densely sampled receivers in the crossline direction (Figure 1-3A). Dense and uniformly sampled data in a marine setting is inherently more difficult because of the terrain, so dense sampling is only performed in one direction and therefore systematic spatial aliasing occurs [Zwartjes and Gisolf, 2007] (Figure 1-3B). Methods are used to increase the spatial sampling through different acquisition geometries, for example, wide-azimuth surveys are common in the industry today [Xie and Gajewski, 2017; Zhang et al., 2021].

1-3 Seismic data reconstruction methods

The inherent irregularities and spatial sparsity of seismic data can be accounted for during data preprocessing. The goal of these methods is to reconstruct sparse data and dead traces to dealias coarsely sampled data (Figure 1-4). Five main categories of interpolation methods have been identified by Xie and Gajewski [2017]. These are sparse transform-based methods, prediction filtering methods, Cadzow rank-reduction methods, wave equation-based methods, and wavefront operator-based interpolation methods. They differ in their applications, assumptions, and results, but all attempt to reconstruct sparse data. We will briefly discuss each of these five methods next.

1-3-1 Prediction filtering methods

Prediction filtering is a powerful tool to reconstruct coarsely sampled and aliased seismic data. These methods are based on the predictability of linear events and are applied by convolution of interpolating filters on sparse data, equivalent to reshaping the data spectrum. The filtered data is then reinserted on a dense regular grid with enhanced spatial resolution.

Prediction error filters can reconstruct uniformly sampled aliased data [Claerbout and Abma, 1992; Spitz, 1991]. Irregularly sampled data must first be regularized - e.g., convolution with a Gaussian window - and subsequently placed on a regular grid [Knutsson and Westin, 1993]. Spitz [1991] presents a typical reconstruction workflow to insert unrecorded samples into spatially under-sampled multi-channel data with random noise. The basis of this method is the predictability of linear events in the frequency-space ($f-x$) domain [Spitz, 1991]. Prediction error filters (PEF) record the dips of non-aliased low-frequency energy, which are used as interpolation operators to reconstruct aliased high-frequency energy. The result greatly enhances the quality of the migrated stacked data [Spitz, 1991]. A similar method described by Claerbout and Abma [1992] estimates the PEF in the time-space ($t-x$) domain, and it registers the dips of the data. Stretching and inverting the PEF enhances the spatial sampling rate thereby dealiasing the filter [3]. These methods use prediction filters in different ways: Spitz [1991] scales the frequencies in the $f-x$ domain, whereas Claerbout and Abma [1992] scale the physical axes of the PEF.

The main limitation of filter-based approaches is that they assume stationarity and linearity of events, which are not the case for seismic data. Methods such as data windowing [Spitz, 1991],

[3] Using the scale-invariant property [Claerbout, 1985]

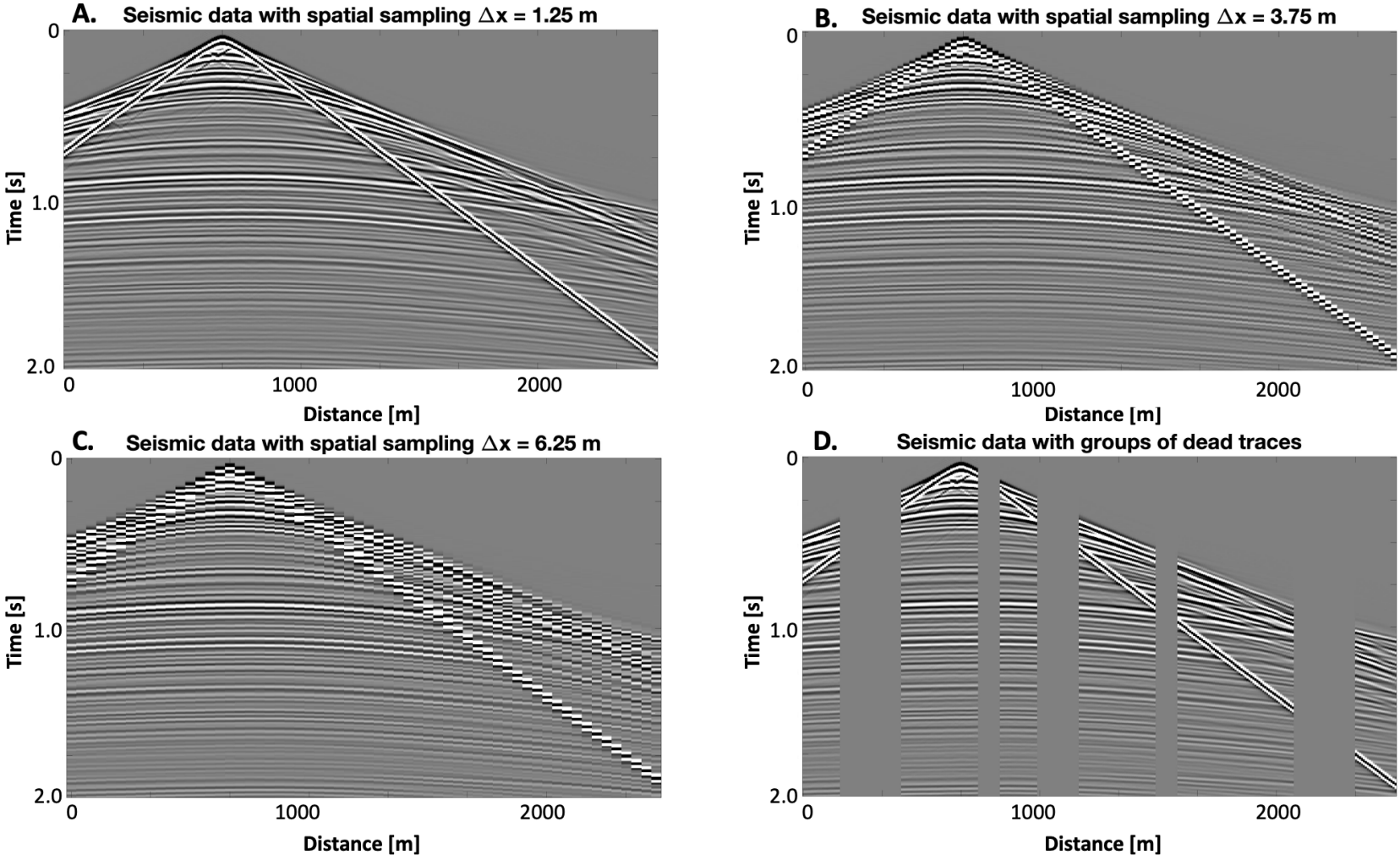

Figure 1-4: Raw seismic data with varying receiver spacing and gaps illustrating data sparsity. (A) $\Delta x = 1.25$ m for 300 traces; (B) $\Delta x = 3.75$ m for 100 traces; (C) $\Delta x = 6.25$ m for 60 traces; (D) Faulty geophones create large gaps in the data. The quality of the data visibly decreases with increasing spatial sampling distance.

adaptive prediction filters [Naghizadeh and Sacchi, 2009], or using smoothly nonstationary PEF coefficients [Liu and Fomel, 2011] have been developed to allow filter-based interpolations. Additionally, prediction filter methods require regularly sampled data [Kutscha, 2014; Spitz, 1991].

1-3-2 Sparse transform methods

Sparse transform methods interpolate and regularize sparse data using the signal's transform domain properties. Artifacts can also be attenuated through iterative thresholding in the transform domain [Kutscha, 2014]. Common transforms used to reconstruct data are the Fourier transform [Liu and Sacchi, 2004; Zwartjes and Sacchi, 2007], the Radon transform [Kabir and Verschuur, 1995; Zhang and Lu, 2014], and the curvelet transform [Naghizadeh and Sacchi, 2010]. Kabir and Verschuur [1995] demonstrate that the insertion of zero traces and subsequent forward and inverse band-limited transform to the parabolic Radon domain results in partial trace restoration, notably for missing near-offsets. Abma and Kabir [2006] interpolate missing data in the Fourier domain using a projection onto convex sets (POCS) algorithm. Iterative transformation and frequency-wavenumber ($f-k$) amplitude thresholding show successful interpolation of irregularly spaced and aliased seismic data. An in-depth discussion of the POCS algorithm will be presented in Chapter 2.

Transform-based reconstruction methods require optimal mapping into the transform domain and derivation of correct transform coefficients from the data. However, the forward transformation usually presents distorted aliased data indistinguishable from unaliased energy [Kutscha, 2014]. Other transform-based methods use constrained inversion, which allows optimal mapping to a transform space without aliasing through regularization [4] [Liu and Sacchi, 2004; Zwartjes and Sacchi, 2007]. While regularization is method-specific, the overall objective is the minimization of the following function:

$$J = ||\vec{p}_m - \mathbf{L}\vec{x}||_2 + \lambda F(\vec{x}) \quad , \tag{1-1}$$

where J is the objective function to be minimized, $\vec{x}$ contains the vectorized model domain parameters, $\vec{p}_m$ is the measured data (therefore is also a function of $\vec{x}$), L is the forward transform operator, $F(\vec{x})$ is the regularization constraint, and λ is the damping factor [Kutscha, 2014].

Liu and Sacchi [2004] propose the reconstruction of spatially undersampled data in the Fourier domain by solving an inverse problem including bandlimited wavenumber regularization. This method allows control over the bandwidth and spectral shape of the reconstructed data [Liu and Sacchi, 2004]. Zwartjes and Sacchi [2007] propose another Fourier-based method that can handle both non-uniformly sampled and aliased data. Fourier coefficients are estimated for low frequency, temporally unaliased, undersampled data, which are used as an a priori model to distinguish aliased energy at high frequencies. Results show suppression of transform domain artifacts from non-uniform sampling and aliasing, and successful sparse trace interpolation [Zwartjes and Sacchi, 2007].

[4] The addition of a priori information and constraints to approximate inverse problems [Engl et al., 1996]

1-3-3 Rank-Reduction methods

Cadzow Rank-Reduction filtering, also called truncated singular value decomposition, is a method that can be used to denoise seismic data and preserve primary components of interest. Results have shown improvements over filter-based signal enhancement methods such as deconvolution and projection filtering, particularly for very noisy datasets [Trickett, 2008]. Additionally, it can also be used for pre-stack trace interpolation to improve spatial sampling [Trickett et al., 2010].

Cazdow filtering of constant frequency slices can both reduce noise and interpolate data. For a given set of traces, frequencies are extracted and placed into a matrix. Research suggests different viable matrix forms such as the Hankel structure, eigenimage strategy, or eigenimage-Cadzow [Trickett, 2008; Trickett and Burroughs, 2009]. A singular value decomposition of this matrix allows its reconstruction as a sum of matrices of rank one, called weighted eigenimages. Reducing the rank of the matrices can reduce noise in traces for all frequencies, as coherent energy is usually concentrated in the first few eigenimages, and noise is normally distributed across all eigenimages [Trickett et al., 2010]. Therefore, the smaller the rank the greater the noise reduction, but the weaker the details preservation. Similarly to its use for denoising, it can perform multidimensional trace interpolation on constant frequency slices [Trickett et al., 2010]. Rank reduction is performed in singular value decomposed matrices with unknown values at missing trace positions. The use of matrix completion algorithms [Kurucz et al., 2007; Olsson and Oskarsson, 2009] results in the addition of traces at new spatial locations.

1-3-4 Wave equation methods

Wave equation-based seismic reconstruction methods are used to reconstruct the wavefield in the subsurface based on the physical properties of wave propagation. The basis of these methods is the Kirchhoff integral that represents the convolution of the data with the wavefield operator [Claerbout and Abma, 1992; Zwartjes and Gisolf, 2007]. Numerically, this equates to the discrete constructive and destructive summation of the data in the propagation direction. However, aliasing artifacts can occur when sampling is too coarse or non-uniform due to incomplete destructive interference [Kutscha, 2014]. Wave equation methods generally are computationally expensive and require a good understanding of subsurface velocity models. The absence of this knowledge can interfere with the model's accuracy [Zwartjes and Gisolf, 2007; Kutscha, 2014].

Several wave equation methods are used to reconstruct data and suppress aliasing artifacts. Artifacts can be reduced through inversion, as is done in imaging with Kirchhoff operators via least-squares migration (LSQM) [Nemeth et al., 1999; Kuehl and Sacchi, 2002]. Key components of the LSQM are migration and de-migration steps, in which a migrated image is calculated from data, and missing data are calculated from the previously computed image [Malcolm et al., 2014]. Other examples of wave equation data reconstruction methods are offset continuation [Bagaini and Spagnolini, 1996], shot continuation [Spagnolini and Opreni, 1996], and dip-moveout (DMO) that transforms prestack data into zero-offset data [Deregowski, 1986]. The combination of DMO followed by inverse DMO is often applied to reconstruct data to a uniform grid. Ronen [1987] uses the wave equation to interpolate traces on spatially aliased multi-channel seismic data. He proposes an inverse problem where the

desired output is well-sampled zero-offset data, and the available input data is normal move-out (NMO) corrected. These domains are linearly related by the wave equation and Fourier analysis [Ronen, 1987]. He finds that interpolating traces by inserting zero-traces, performing DMO, and stacking is equivalent to the application of the transpose operator in the absence of spatial aliasing.

1-3-5 Wavefront attribute methods

Wavefront attribute interpolation methods reconstruct missing traces using the physical attributes of the measured wavefield. These attributes are determined from prestack data and have physical meanings such as the angle of emergence, traveltime, or wavefront curvature. They are used in processing to predict and interpolate neighboring traces - e.g., through partial stacking [Hoecht et al., 2009]. Several methods exist to determine these wavefront attributes, such as non-hyperbolic common reflection surface (CRS) [Fomel and Kazinnik, 2013], i-CRS [Schwarz et al., 2014], multifocusing [Gelchinsky et al., 1999], and nonlinear beamforming [Sun et al., 2022].

For example, the CRS interpolation method uses the kinematic attribute of the measured wavefield for data reconstruction [Hoecht et al., 2009]. These kinematic traveltimes are estimated using a second-order wavefront surface approximation:

$$\Delta t = t(\Delta x, \Delta y) - \hat{t} = b_0 \Delta x + b_1 \Delta y + a_{00} \Delta x^2 + a_{01} \Delta x \Delta y + a_{11} \Delta y^2 \quad , \tag{1-2}$$

where $t(\Delta x, \Delta y)$ is a time sample of a trace located at (x, y), $\hat{t}$ is a time sample of a trace located at $(0, 0)$, and $(\Delta x, \Delta y)$ are the respective (x, y) distances between these two traces. Available data are used to solve for the five unknown parameters $b_0, b_1, a_{00}, a_{01}, a_{11}$ along all available dimensions to build wavefront traveltime operators around parameter traces. Target traces are reconstructed by weighted interpolation of surrounding parameter trace operators [Hoecht et al., 2004]. Further advances to the CRS method by Hoecht et al. [2009] suggest an alternative operator-oriented (OO) interpolation scheme. This method constructs multiple wavefields using the operators at several parameter trace locations (Equation 1-2), and their intersection with the target trace defines the value for that time sample. Here, several operators contribute to a target sample's position along the trace, which stabilizes the interpolation. OO scheme results show better continuity and reduced noise [Hoecht et al., 2009]. Further optimization of wavefront attributes CRS interpolation methods such as global optimization search strategies for the extension to 5D data have shown significant improvements in data resolution compared to the conventional 3D method [Xie and Gajewski, 2017].

1-4 Thesis outline

In the following thesis, we will discuss two seismic data reconstruction methods we have developed to reconstruct sparse seismic data. The main method of interest is a kinematic wavefront method, which we call nonlinear beamforming (NLBF) seismic data reconstruction. This method is based on the existing framework of NLBF for seismic signal enhancement [Sun et al., 2022]. The second method, a byproduct of this thesis, is a bootstrap alternating projection onto convex sets (BP) method. This method is an adapted version of the alternating projection onto convex sets (APOCS) algorithm and will be used as a control method for the NLBF reconstruction.

Therefore, in this thesis we attempt to answer the following question: can the NLBF framework be used for seismic data reconstruction, and how do these results compare to some conventional reconstruction methods? To answer this question, we present our research in the following manner:

- Chapter 2: *Alternating projection onto convex sets*
  In this chapter, we introduce the APOCS algorithm, a transformed-based signal reconstruction method. We present its potential to reconstruct sparse 2D images and propose modifications to an existing method to further enhance results.

- Chapter 3: *Nonlinear beamforming*
  In this chapter, we present the NLBF framework and its use for seismic signal enhancement. We define the NLBF traveltime operators, introduce the solvers used to estimate them from input data, and illustrate how these are used for seismic signal enhancement.

- Chapter 4: *Methods*
  In this chapter, we present the methodology of the developed seismic data reconstruction methods. First, we present the use of the NLBF framework for seismic data reconstruction. Then, we present the BP method, an adapted version of the APOCS scheme.

- Chapter 5: *Results*
  In this chapter, we present the results of the NLBF data reconstruction method on a sparse synthetic SEAM Arid dataset and a field dataset. We compare these results to those of several control methods, including the introduced BP method.

- Chapter 6: *Discussion and conclusions*
  In this chapter, we further discuss the results, applications, and potential research opportunities of the NLBF and BP reconstruction methods. Finally, we conclude by answering our research question.

Chapter 2

Alternating projection onto convex sets

Alternating projections onto convex sets (APOCS) is a powerful method for image and signal reconstruction [Marks II, 1997]. Given certain sets of constraints represented by convex sets, a signal can be reconstructed by iterative projections onto these convex constraints [Marks II, 1997]. In the following chapter, we discuss the method's mathematical framework, present its applications in seismic data restoration, and test our algorithm's efficiency on synthetic seismic reflections and 2D images. The theory presented here forms the basis of the BP method in Chapter 4.

2-1 Projections onto convex sets (POCS) theory

2-1-1 Convex set

To illustrate the principle of convex sets, we use the vector space as it allows a simplified geometric illustration. In this sense, convex sets contain all vectors that share a predefined property. Geometrically, a set A is convex if for any two vectors $\vec{x} \in A$ and $\vec{y} \in A$, Equation (2-1) holds [Marks II, 1997]. Equation 2-1 can be visualized as the line segment between two vectors contained in A (Figure 2-1). As λ goes from 0 to 1, we traverse the line segment. If a piece of the line is not in the set (Figure 2-1B), then the set is not convex. Convex geometrical sets include balls, boxes, lines, line segments, cones, and planes [Marks II, 1997].

$$\lambda\vec{x} + (1-\lambda)\,\vec{y} \quad \in A\,. \tag{2-1}$$

In signal reconstruction, the same principle is applied to signals in the signal space, which is defined as all signals with finite energy [Marks II, 1997]. It contains higher-order convex sets and subspaces that can be visualized as hypervolumes. Signals that form convex sets are usually geometrically related. For example, bounded signals, signals with identical middles, signals with a constant area, band-limited signals, or signals with a constant phase.

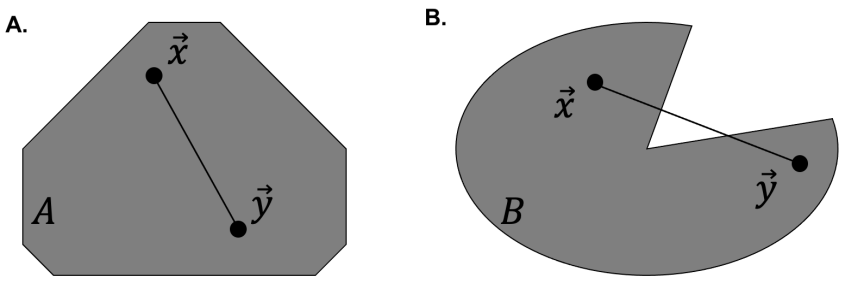

Figure 2-1: Geometric illustration of (A) a convex set, and (B) a non-convex set. A set is convex if every line segment with endpoints in the set is totally subsumed in the set.

2-1-2 Projections

The concept of signal projection is simplified using the vector analogy once more. Given a vector $\vec{u}$ within the convex set A, its projection onto A is just itself. If $\vec{u}$ is not in the convex set A, then its projection onto A must satisfy all the desired constraints of A [Marks II, 1997]. Geometrically, the projection is performed such that the distance between $\vec{u}$ and its projection $\vec{v}$ is minimized (Figure 2-2) (Equation 2-2). The projection is usually obtained by forcing the signal to conform to the constraint in the simplest way. This concept is easily scalable to higher dimensions in the signal space [Marks II, 1997].

$$min||\vec{u} - \vec{v}|| \qquad \vec{v} \in A \ . \tag{2-2}$$

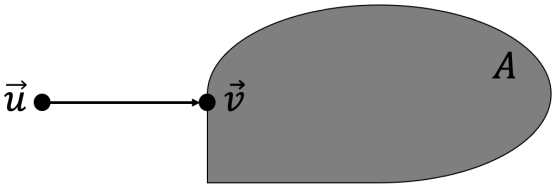

Figure 2-2: A is a convex set and the vector $\vec{u}$ is not included in A. The projection of $\vec{u}$ onto A is $\vec{v}$, the unique element in A closest to $\vec{u}$.

For two or more convex sets, alternating the projections between these sets - namely APOCS - aims to reconstruct a signal conforming to all the constraints of these convex sets. However, the convergence outcome depends on how the convex sets intersect. If the convex sets intersect in a single point, the alternating projections will converge to a unique solution (Figure 2-3A). If the intersection area is more than a point - a line, a hyperplane, or a hypersphere volume in general - the solution is non-unique (Figure 2-3B). It then depends on the starting point. If the sets do not intersect, convergence can lead to a limit cycle or greedy limit cycle (Figure 2-3C), only partially conforming to the desired convex set properties [Marks II, 1997].

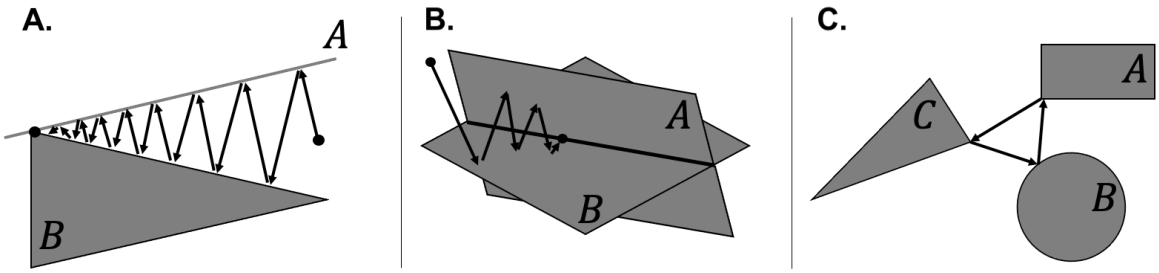

Figure 2-3: Alternating projections between convex sets with different intersections. (A) Two sets intersect at a point, so alternating projections converge to a unique solution. (B) Two sets intersect along a line, so the APOCS solution is non-unique. (C) Three convex sets do not intersect, so the alternating projections end up in a limit cycle.

2-2 Alternating POCS (APOCS) in seismic data reconstruction

The APOCS method has shown effective sparse seismic data reconstruction [Gao et al., 2010]. Abma and Kabir [2006] demonstrate successful interpolations of irregular seismic data using a single-pass APOCS (SP) method. The convex sets A and B used are represented by Equations 2-3 and 2-4. Set A contains the known ground-truth traces at the correct locations; in set B all $f-k$ components below the defined threshold β are set to 0s.

$$A = \{u(x,y) \mid u(x,y) = v(x,y) \text{ for pre-defined } (x,y) \text{ locations } \}, \quad u, v \in R \ . \tag{2-3}$$

$$B = \{u(x,y) \mid u(x,y) = 0 \text{ for } |(u(x,y)| < \beta\}, \qquad u \in C \ . \tag{2-4}$$

where for each (x,y) trace position, $u(x,y)$ are the signals contained in sets A and B, $v(x,y)$ are the ground truth signals, R is the set of all real numbers, and C the set of all complex numbers.

Projection of a signal $S(x,y)$ onto A is performed by reinserting the original data $u(x,y)$ at known (x,y) locations, and preserving the remaining signals $s(x,y)$ at unknown (x,y) locations (Equation 2-5). Projection of a signal $S(x,y)$ onto B is performed by a 2D Fourier transformation, multiplication with the $f-k$ thresholding mask T leaving only the strong amplitude components (Equation 2-7), and a 2D inverse Fourier transformation [Gao et al., 2010] (Equation 2-6). The $f-k$ threshold used by Abma and Kabir [2006] and Gao et al. [2010] is initially strong and decreases between each set of projections. Gao et al. [2010] suggest that an exponential decrease of the threshold results in a faster interpolation of missing data. Contrary to the cited literature, the APOCS algorithm used in this thesis adds a convergence condition for the alternating projections (Equation 2-8), and the $f-k$ threshold only updates when convergence between sets A and B is reached for the current threshold. This new version of the APOCS algorithm is called convergent POCS (CP), and it follows the workflow outlined in Figure 2-4. Research by Hirsch [2022] presents the advantages of the CP method compared to the conventional SP method in greater detail.

$$P_A(S(x,y)) = \begin{cases} u(x,y) & \text{if (x,y) is a pre-defined location} \\ s(x,y) & \text{otherwise} \end{cases} , \tag{2-5}$$

$$P_B(S(x,y)) = F_y^{-1}F_x^{-1}TF_xF_yS \quad , \quad \text{T is a pre-defined 0/1 mask matrix} \quad , \tag{2-6}$$

$$T_k(x,y) = \begin{cases} 1, |S_{k-1}(f,k)| \geq p_k \\ 0, |S_{k-1}(f,k)| < p_k \end{cases} , \quad p_k \in p \quad , \tag{2-7}$$

$$\|S_{k+1} - S_k\| < \alpha \cdot \|S_k\| \quad , \tag{2-8}$$

where S_{k-1} denotes the $f-k$ spectrum of reconstructed data at the k-1 th iteration, p the N-dimensional threshold set, $p = p_1, p_1, ..., p_N$ meets $p_1 > p_2 > ... > p_N$, and N denotes the maximum thresholding iteration [Gao et al., 2010]. The convergence constant α is predefined and depends on the desired degree of accuracy (Equation 2-8). According to our study [Hirsch, 2022], we recommend to use $\alpha = 0.0001$ and L_2 norm in Equation 2-8.

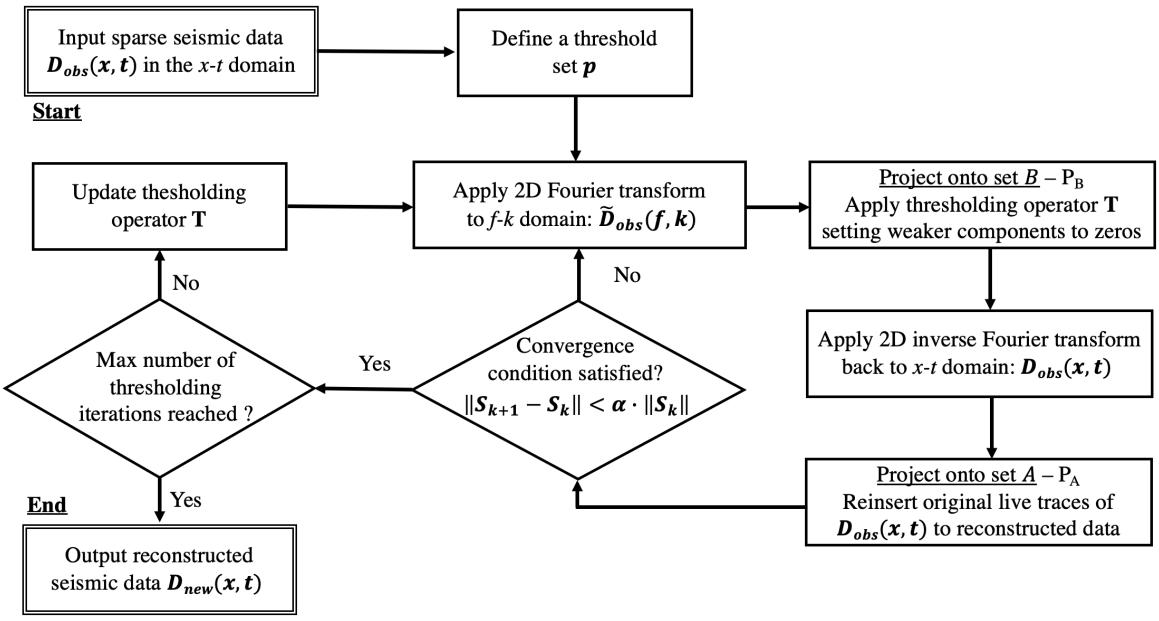

Figure 2-4: Flowchart of the convergent POCS (CP) algorithm.

2-3 Convergent POCS (CP) Application

2-3-1 Synthetic seismic data

Application of the CP algorithm shows successful interpolation of irregularly sparse synthetic seismic reflections. The algorithm was tested on synthetic seismic reflections with linear and curved events with 65% and 85% randomly generated sparsity, for 100 and 1000 thresholding iterations (Figures 2-5 and 2-6). Parameter tests performed by Hirsch [2022] reveal that image reconstruction quality plateaus around 1000 thresholding iterations and shows different trends for curved and linear events.

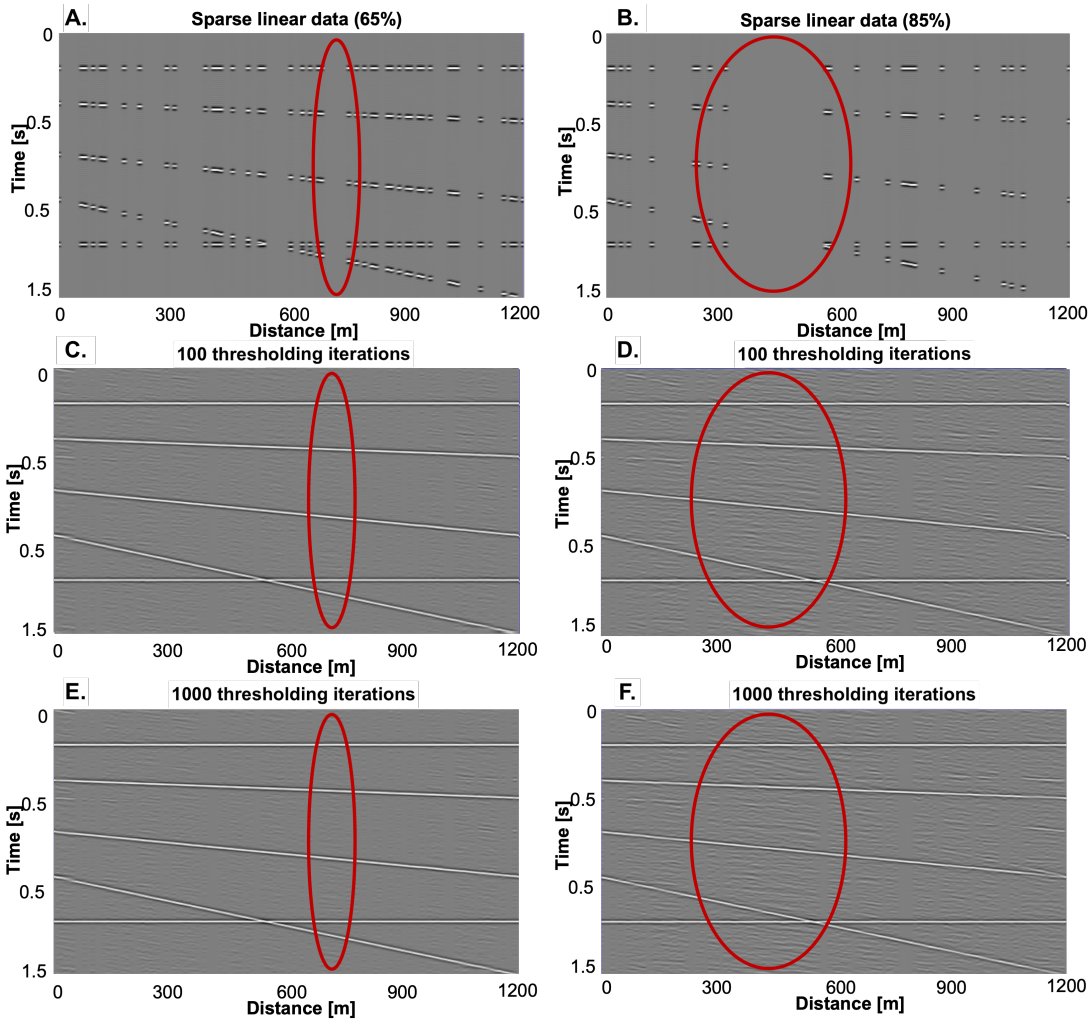

Figure 2-5: CP reconstruction results of 65% (A,C,E) and 85% (B,D,F) sparse linear synthetic seismic reflections. (C-D) Results for 100 CP thresholding iterations, and (E-F) results for 1000 CP thresholding iterations. Red ellipses highlight some differences.

For linear events - horizontal and dipping - larger numbers of thresholding iterations yield better reconstruction results [Hirsch, 2022]. Figure 2-5 illustrates better reflection continuity and reduced background noise for 1000 thresholding iterations compared to 100 thresholding

iterations. These features are highlighted by the red ellipses. For curved events (Figure 2-6), larger amounts of thresholding iterations appear to better reconstruct the reflections. However, we also observe enhanced background noise for 1000 thresholding iterations compared to 100 iterations, which can obstruct the reflections of interest (Figure 2-6).

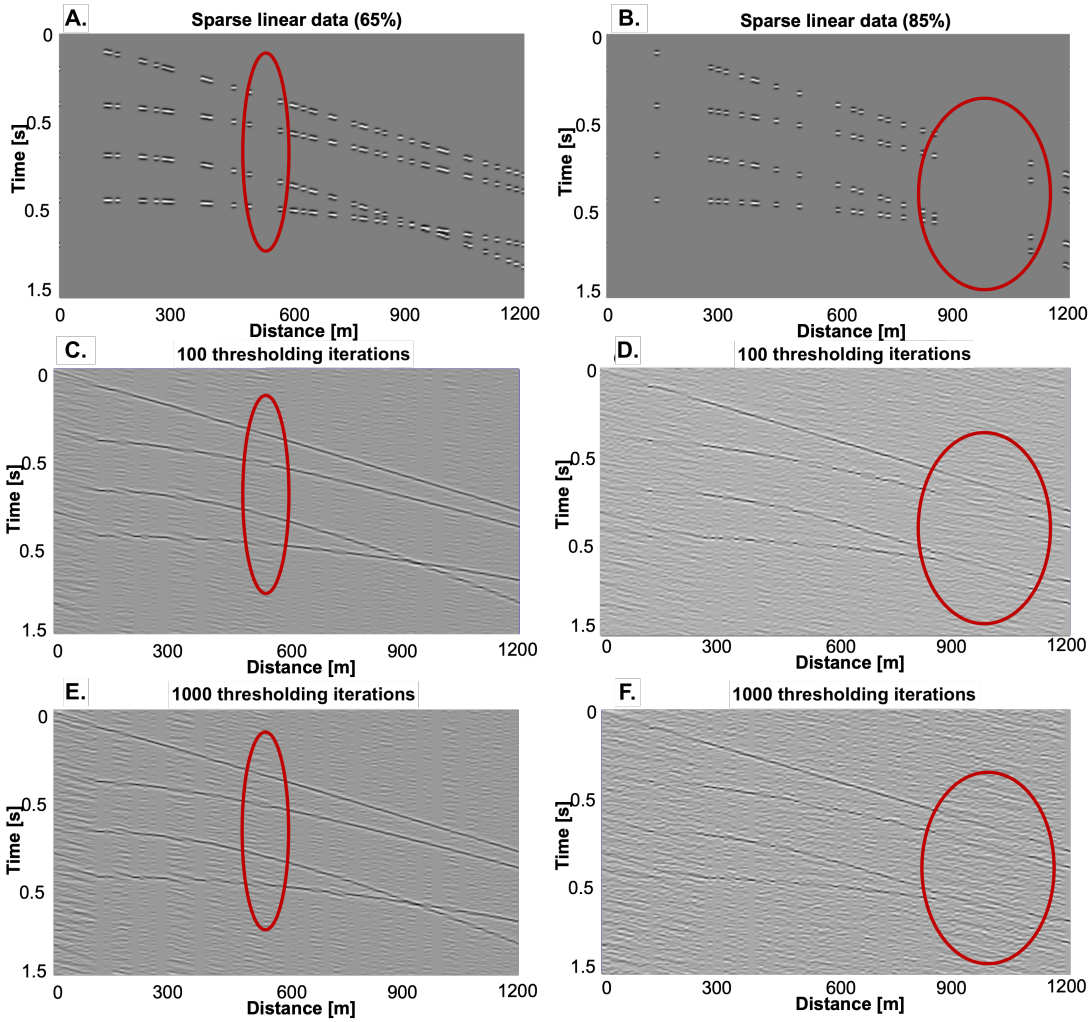

Figure 2-6: CP reconstruction results of 65% (A,C,E) and 85% (B,D,F) sparse curved synthetic seismic reflections. (C-D) Results for 100 CP thresholding iterations, and (E-F) results for 1000 CP thresholding iterations. Red ellipses highlight some differences.

Overall, the interpolation of linear events requires fewer thresholding iterations compared to curved events and exhibits superior reconstruction quality (Figures 2-5 and 2-6). In fact, linear reflections are interpolated fast and display little added noise (Figure 2-5). Curved seismic events generally require more thresholding iterations to resolve the sparse traces and are subject to increased noise. Additionally, both linear and curved reconstructed events show increased levels of noise in the reconstruction at higher degrees of sparsity (Figure 2-5 and 2-6). These observations must be taken into account when attempting to reconstruct sparse, geometrically complex, seismic wavefronts.

2-3-2 2D image reconstruction

The CP algorithm was also tested on two grayscale images. Analogously to the previous section, different percentages of randomly decimated pixels and thresholding iterations were used. Hirsch [2022] shows the full extent of our parameter testing to achieve the best possible reconstruction. Figures 2-7 and 2-8 illustrate results for 75% and 90% sparse images, and the associated signal-to-noise ratio (SNR) to assess the quality of the CP reconstruction. The SNR in question is presented in detail in Chapter 5. The CP algorithm shows impressive image reconstruction abilities. For high levels of data sparsity (90%), we observe successful image reconstruction around 1500 thresholding iterations. Similarly to the CP results in synthetic seismic data, we find that higher levels of data sparsity require more thresholding iterations to successfully reconstruct the image and can introduce increased background noise.

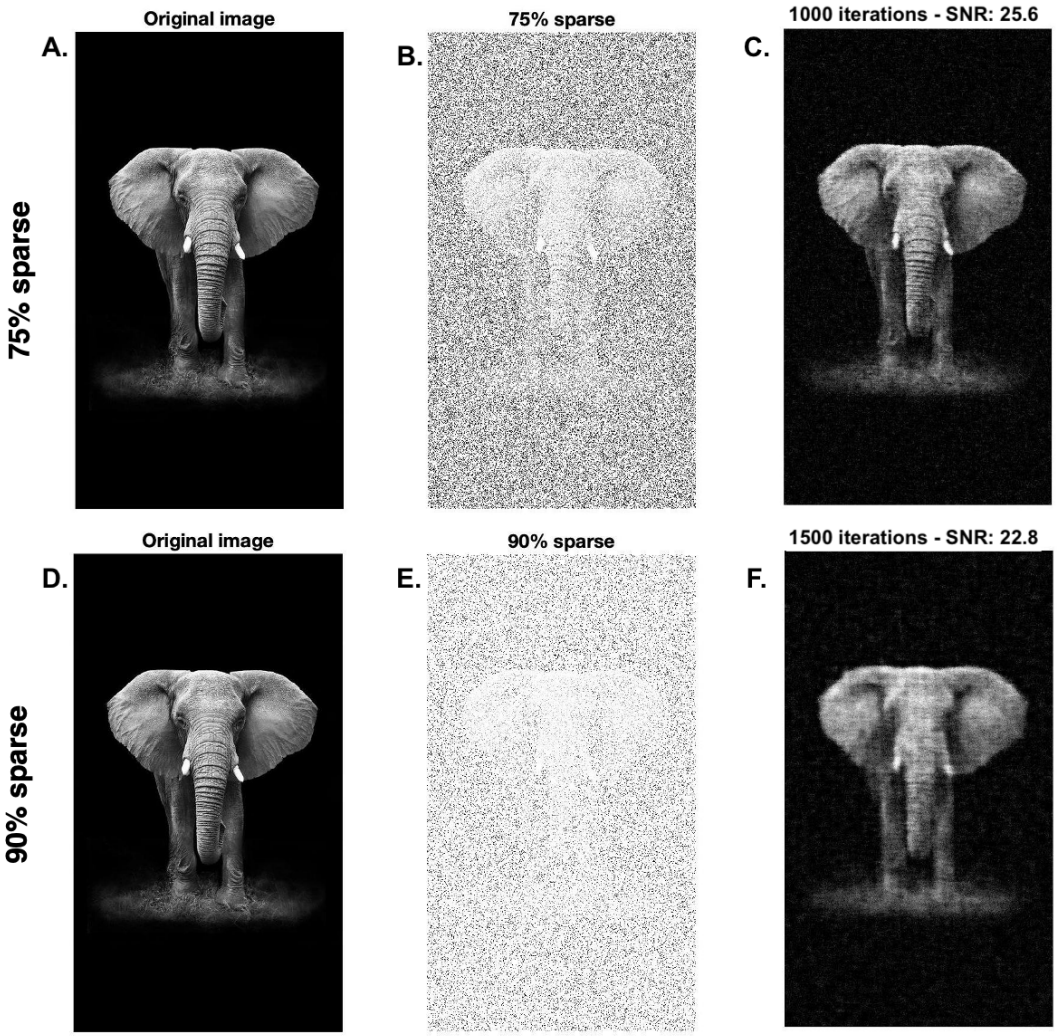

Figure 2-7: CP reconstruction on 75% (A-C) and 90% (D-F) sparse images. (A,D) The original images, (B,E) the sparse images, and (C,F) results after 1000 (SNR = 25.6) and 1500 (SNR = 22.8) thresholding iterations.

Image reconstruction quality in Figure 2-7 appears superior to Figure 2-8, but the SNR values are in fact lower. These differences can be explained by the superior complexity in the gray scales and geometric shapes in Figure 2-8. Overall, CP results are promising as they show successful image reconstruction and adequate SNR values. These results reveal the potential of the CP method on real seismic data as we will see in the following chapters. However, it is important to note that the CP algorithm was only tested on randomly decimated images. Chapter 4 highlights a major limitation of the CP method, which fails to reconstruct regularly sparse data. There we introduce an adaptation to the CP algorithm for images with regular decimation. More on this in Chapter 4.

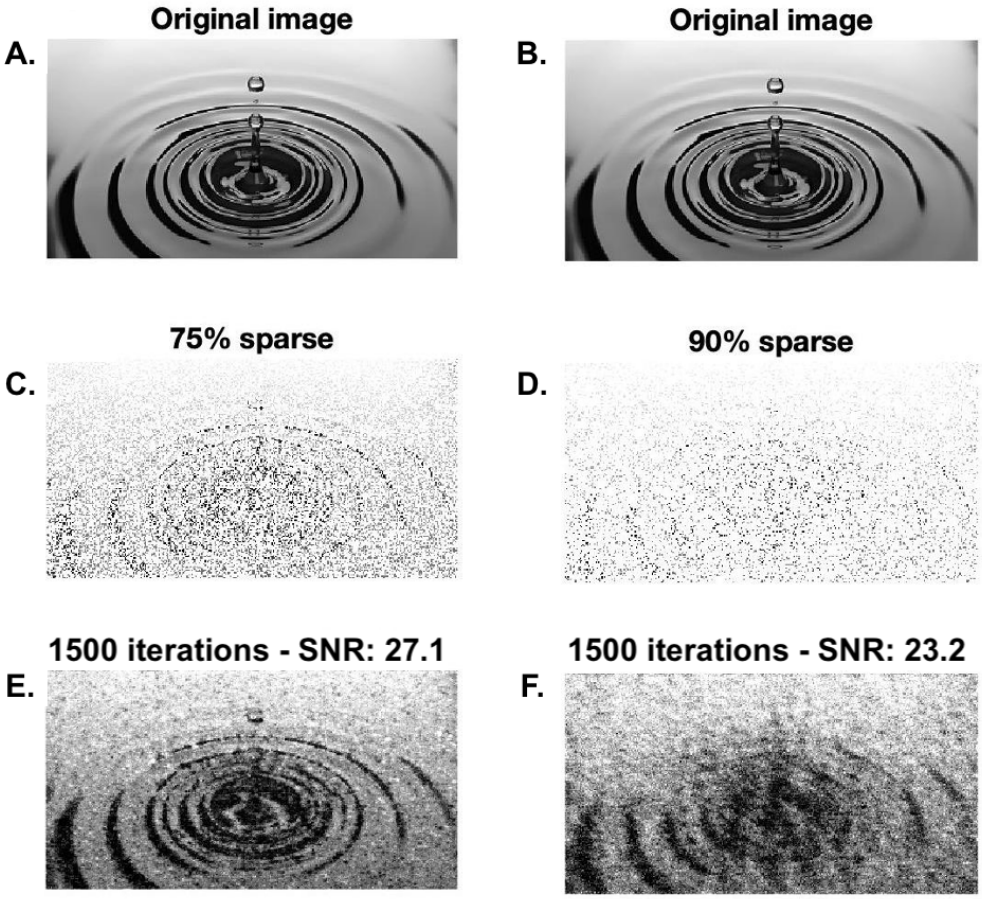

Figure 2-8: CP reconstruction on 75% (A,C,E) and 90% (B,D,F) sparse images. (A,B) The original images, (C,D) the sparse images, and (E,F) results after 1500 thresholding iterations ($SNR_E = 27.1$ and $SNR_F = 23.2$).

Chapter 3

Nonlinear beamforming

Nonlinear beamforming (NLBF) is a modern 3D seismic signal-enhancement method that uses multi-dimensional trace stacking to improve the signal-to-noise ratio (SNR) of prestack data [Bakulin et al., 2020]. Modern land seismic acquisition is moving away from large source and receiver arrays to dense high channel count and single sensor geometries [Pecholcs et al., 2012; Regone et al., 2015]. These methods allow detailed sampling of wavefields but lead to weaker signals and stronger noise compared to conventional acquisition geometries. Existing preprocessing methods such as deconvolution and velocity analysis are unable to robustly deal with such data [Sun et al., 2022], so a new suite of tools is needed. This chapter presents the NLBF framework and its use for enhancing the sought-after seismic signal. In the larger context of this thesis, we use the NLBF framework presented hereafter for a new purpose, seismic data reconstruction. More on this in Chapter 4.

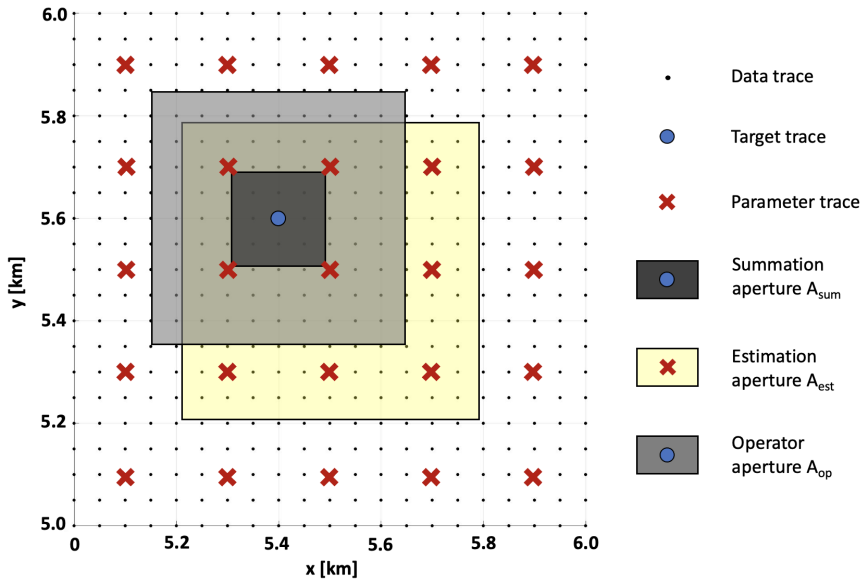

Figure 3-1: Example trace geometry in the spatial domain. Black dots show the data traces, red x's the parameter traces, and the blue dot the enhanced target trace. The summation, estimation, and operator apertures are indicated in the legend.

The NLBF framework can be used to enhance the SNR of existing traces by a local weighted summation of traces along local operators. The term "operator" here defines a wavefront, determined at parameter trace locations from the data itself (Figure 3-1) [Bakulin et al., 2020]. Figure 3-1 illustrates the parameter trace locations and the multiple user-specified apertures involved in the NLBF signal-enhancement process. The estimation aperture (A_{est}) defines the traces used to determine the operator's parameters, which are used to calculate the operators (kinematic wavefronts). These NLBF operators can be used in the signal-enhancement process, and are also the basis of our reconstruction method presented in Chapter 4. For the signal-enhancement process, there are two more operators engaged: the operator aperture (A_{op}) determines the number of surrounding operators used to enhance a specified target trace; the summation aperture (A_{sum}) defines the number of neighboring data traces used in the local summation along the operators to produce the SNR-enhanced output trace [Bakulin et al., 2020]. NLBF can be applied in any user-specified domain such as common-shot, common-receiver, or common-midpoint [Bakulin et al., 2020; Hoecht et al., 2009]. The cross-spread domain (CSD) is commonly used for its regular and dense trace distribution [Bakulin et al., 2020]. In the following sections, we present the NLBF operator estimation and the use of NLBF operators for seismic data SNR enhancement. We introduce the traveltime operator, the cost function, some existing solvers, the local weighted summation, and present some enhancement results.

3-1 Traveltime operators

The traveltime operator $\Delta t(x, y; x_0, y_0)$ describes the traveltime moveout of a kinematic wavefront with respect to a specified parameter trace time sample (Figure 3-1). It is approximated using a second-order surface:

$$\begin{aligned} \Delta t\left(x, y; x_0, y_0\right) &= t(x, y) - t\left(x_0, y_0\right) \\ &= A\Delta x + B\Delta y + C\Delta x\Delta y + D\Delta x^2 + E\Delta y^2 \quad , \end{aligned} \tag{3-1}$$

where $t(x, y; x_0, y_0)$ is the traveltime for a trace located at (x, y), given an NLBF parameter trace at (x_0, y_0); Δy and Δx are the distances from the position of the current trace to the parameter trace: $(y - y_0)$ and $(x - x_0)$ respectively. The unknown coefficients A, B, C, D and E describe the operator at $(x, y, t(x, y))$ [Sun et al., 2022]. Figure 3-2 illustrates a single operator around the parameter trace time sample at $t = 6$ ms; it extends in both the spatial and temporal directions. The Δy and Δx values from Equation 3-1 are shown in Figure 3-2A. The operator's Δt value at data trace locations is illustrated in Figure 3-2B, and it is calculated using Equation (3-1).

To enhance seismic signals using NLBF, we must first estimate the local kinematic wavefront parameters A, B, C, D and E from the data inside A_{est} for each parameter trace. These parameters define the traveltime operators and must be estimated for each time sample of each parameter trace (Figures 3-1 and 3-2). In the second step, a target trace is enhanced by a weighted summation of traces in A_{sum} along operators in A_{op}. The summation is further discussed in Section 3-4.

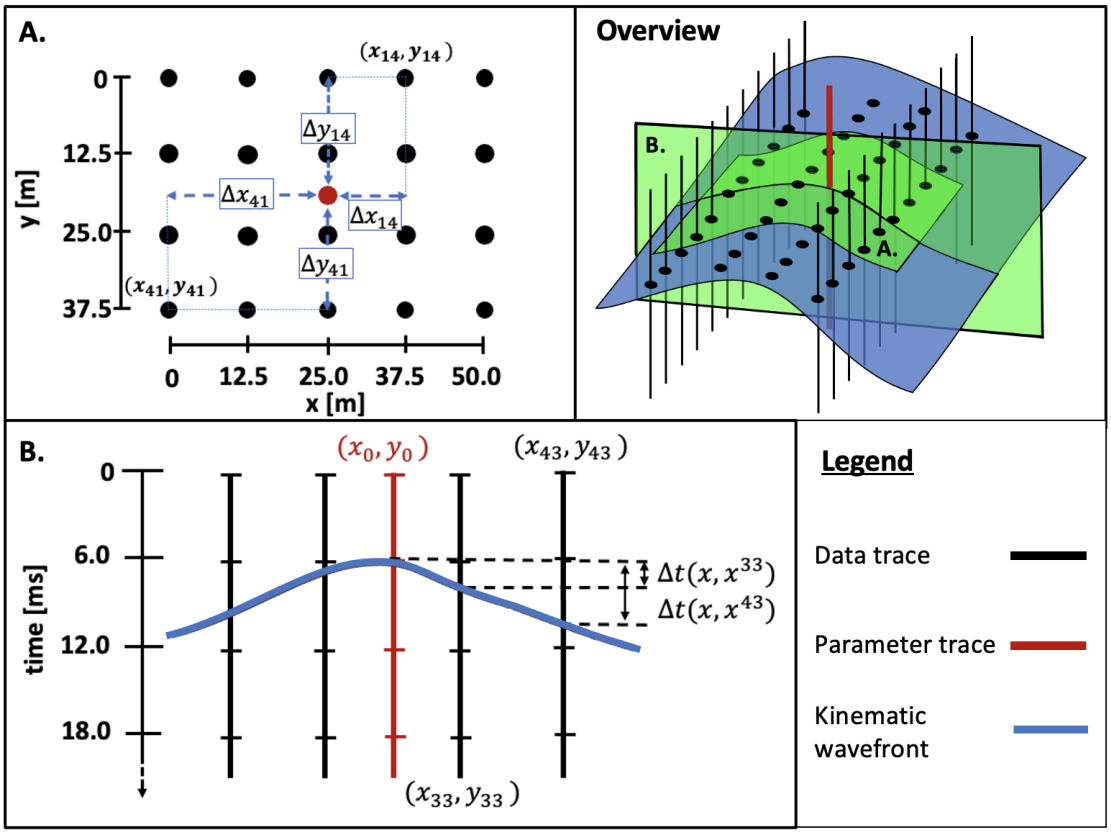

Figure 3-2: Schematic illustration of the operator construction Equation (3-1) in 3D (x, y, t). The overview inset shows the 3D operator around the red parameter trace, cross-cutting the data traces (black dots). (A) Top view of an operator section, highlighting Δy and Δx values for two data trace positions. (B) Cross-section highlighting traveltime operator values for two data trace positions.

3-2 Cost function

The traveltime operator parameters A, B, C, D and E estimated at parameter trace locations are dependent on the surrounding data traces, and therefore on the specified aperture size A_{est} [Bakulin et al., 2020; Sun et al., 2022]. A semblance-type cost function is optimized for these parameters, and it aims to maximize the semblance inside a chosen aperture such as:

$$S(x_0, y_0) = \frac{\sum_{j=1}^{N} \left\{ \sum_{i=1}^{M} u\left[x_i, y_i; t_j(x_0, y_0) + \Delta t(x_i, y_i; x_0, y_0)\right] \right\}^2}{M \sum_{j=1}^{N} \sum_{i=1}^{M} \left\{ u\left[x_i, y_i; t_j(x_0, y_0) + \Delta t(x_i, y_i; x_0, y_0)\right] \right\}^2} \quad , \tag{3-2}$$

where $u(x_i, y_i, t)$ represents a time sample of the trace located at (x_i, y_i) in seismic data, M is the total amount of traces inside the spatial aperture A_{est}, and N the total amount of samples inside the temporal aperture of a local traveltime operator [Sun et al., 2022].

The cost function, Equation (3-2), is highly nonlinear, therefore solving for coefficients A, B, C, D and E at $t(x_0, y_0)$ by maximizing Equation (3-2) is non-trivial. The search ranges for these coefficients are usually determined by local geology and must be specified by the user [Bakulin et al., 2020]. Several solvers, such as the 2+2+1 local search method, the 5D brute force method [Sun et al., 2022], sequential dips and curvatures estimation, and the efficiency-improved Genetic Algorithm (eGA) method, have proven their efficacy [Bakulin et al., 2020; Sun et al., 2022]. In the following section, we present possible solvers, notably the NLBF+eGA method used in this thesis.

3-3 Solvers

Estimation of local kinematic parameters A, B, C, D and E is performed by scanning different wavefronts and finding the best coherency based on the maximum semblance value [Bakulin et al., 2020]. The 5D brute force method is stable and accurate in finding the maximum semblance value but it is computationally very expensive [Sun et al., 2022]. The 2+2+1 method is less expensive but presents risks of convergence to local maxima [Bakulin et al., 2020]. In the 2+2+1 approach, a local solver breaks down the 5D optimization problem into three separate problems with smaller dimensions. Parameters A and D are estimated in the x direction using a subset of the ensemble data, B and E in the y direction using another subset of the ensemble data, and C in both directions using the complete data [Buzlukov and Landa, 2013]. Overall, efficiency is improved but results are sub-optimal.

Genetic algorithms (GAs) have found successful applications in the field of geophysics as nonlinearity is a reoccurring challenge [Sambridge and Drijkoningen, 1992; Sun and Verschuur, 2014]. Acuna and Sun [2020] propose the efficiency-improved GA (eGA), which offers higher convergence speeds over other GAs. It demonstrates systematic efficiency and robustness in the global optimization of nonlinear CRS stacking problems and has shown its potential to estimate nonlinear NLBF kinematic wavefront parameters [Sun et al., 2022]. The eGA uses a modified island model to secure a robust global search, a self-adaptive differential evolution fine-tuning scheme to guarantee a good local search, and a local exhaustive search method to handle the premature convergence [Acuna and Sun, 2020]. Seismic events of interest often lie in the far field, where wavefronts smoothly vary with distance. Therefore, traveltime operator parameters between nearby parameter traces must also change gradually (Figure 3-1). The NLBF+eGA method exploits this "spatial consistency" property to both enhance computational efficiency and improve final data quality [Sun et al., 2022]. It uses estimated coefficients at parameter traces as initial values for neighboring parameter traces. The NLBF+eGA method has proven to be very efficient in yielding high-quality estimations of NLBF parameters A, B, C, D and E in noisy datasets with less computational effort compared to other methods.

Figure 3-3 illustrates an example of operator parameters A, B, C, D and E and the resulting semblance for several parameter traces, calculated using the NLBF+eGA method [Sun et al., 2022]. These operators were calculated for signal enhancement purposes using a finite difference simulation on a velocity and a density model from a synthetic SEAM Arid dataset (Figure 5-1). The acquisition geometry is shown in Sun et al. [2022]. Resulting data trace spacing is $\Delta x = \Delta y = 0.625$ m, time sampling rate is 4 ms, and parameter trace spacing is $\Delta x = \Delta y = 140$ m with an A_{op} of 400 m x 400 m. Sun et al. [2022] highlights the superiority

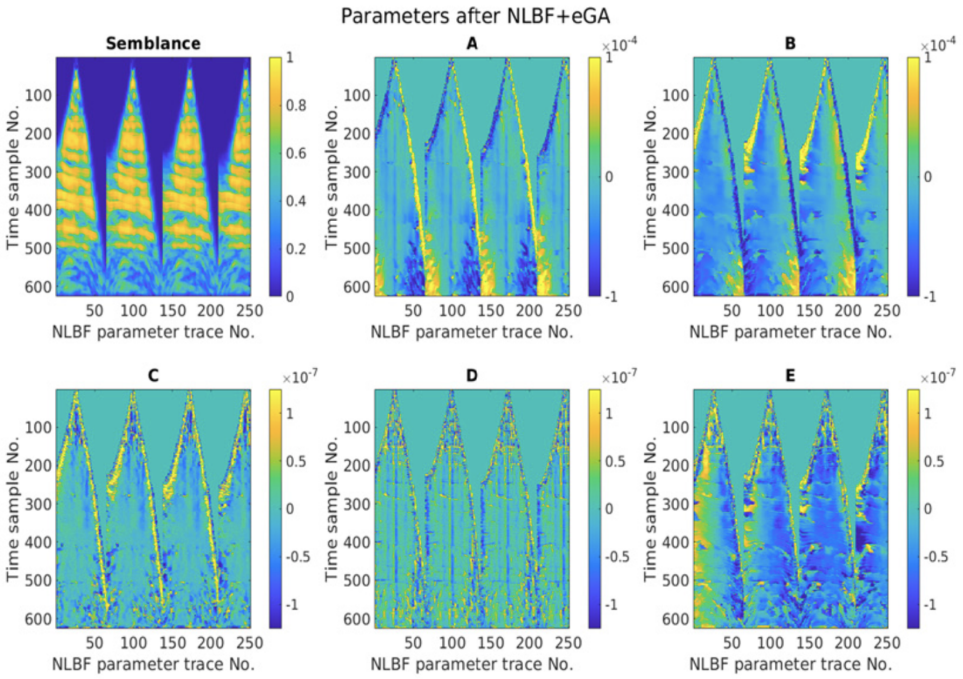

Figure 3-3: Estimated semblance and parameters A-E for NLBF parameter traces obtained by the NLBF+eGA method on the SEAM Arid dataset, adapted from Sun et al. [2022]

of the NLBF+eGA solver by the higher semblance values compared to the other solvers previously discussed. In this thesis, the traveltime operator parameters have been estimated for the datasets of interest using the NLBF+eGA method, and are directly ready to use.

3-4 Local summation

NLBF is a wavefront operator-based signal enhancement method that exploits multi-dimensional trace stacking, where weak signals are enhanced by summation along coherent wavefronts [Bakulin et al., 2020; Sun et al., 2022]. The last step to seismic signal enhancement using NLBF is the weighted summation of data traces in A_{sum} along the operators defined by A_{op} [1] (Figure 3-1). The NLBF local summation method can be written in the general form:

$$u\left(\mathbf{x}_0, t_0\right) = \sum_{\mathbf{x} \in A_{sum}} w\left(\mathbf{x}, \mathbf{x}_0\right) u\left(\mathbf{x}, t_0 + \Delta t\left(\mathbf{x}, \mathbf{x}_0\right)\right) \quad , \tag{3-3}$$

where u represents the data traces in the original gather with location $\mathbf{x} = (x, y)$, whose values (black dots in Figures 3-1 and 3-2) are summed along the wavefronts in the defined window A_{sum}; the beamforming weights $w(\mathbf{x}, \mathbf{x_0})$ are used to better preserve signals and to suppress noise; $\mathbf{x_0} = (x_0, y_0)$ represents the position of the resulting enhanced trace [Bakulin et al., 2020].

[1] The inclusion of several operators in A_{op} is only for the operator oriented scheme; in the target-oriented scheme a single operator is used therefore A_{op} does not need to be defined [Bakulin et al., 2020].

Different summation methods such as the operator oriented (OO) and trace oriented (TO) approaches can be used [Bakulin et al., 2020; Hoecht et al., 2009]. The TO method estimates kinetic wavefront parameters at all target trace locations, and these are used to construct operators for the target time samples [Bakulin et al., 2020]. The TO scheme results in one unique operator per output sample. The OO method, represented in Figures 3-1 and 3-2, estimates, and builds traveltime operators on a regular grid of parameter traces (Figure 3-2). During the summation, enhanced target traces can receive contributions from several surrounding operators defined in A_{op}. The OO scheme offers significant computational advantages compared to the TO method. Notably, computational performance is improved by using a sparser grid of parameter traces [Bakulin et al., 2020].

3-5 Signal enhancement examples

The NLBF signal enhancement results illustrated in Figure 3-4 were obtained by Sun et al. [2022] using the synthetic SEAM Arid dataset and operator parameters from Figure 3-3. These results illustrate the potential of the NLBF method for seismic SNR enhancement. Figure 3-4 shows several NMO corrected shot gathers before and after the NLBF+eGA enhancement. The original gather (Figure 3-4A) exhibits strong scattering noise from the complex near-surface highlighted by the red oval [Sun et al., 2022]. Figure 3-4B illustrates the results after the enhancement with NLBF+eGA. There is a clear improvement in the quality of the prestack data, and events show better continuity and less background noise as highlighted by the red oval. Further details regarding the signal enhancement results are available in Sun et al. [2022], as seismic signal enhancement is not the main focus of this thesis.

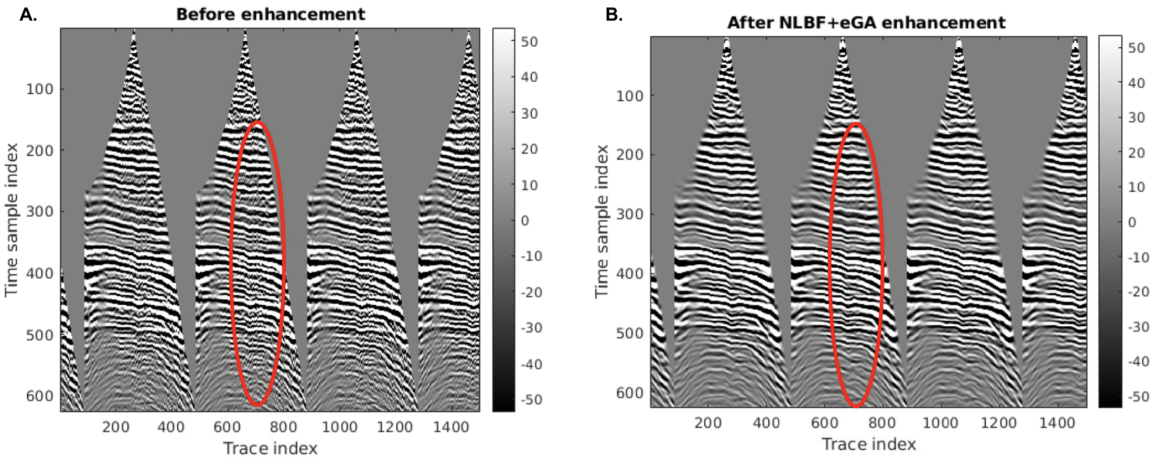

Figure 3-4: (A) NMO-corrected raw gathers from synthetic data. (B) NLBF+eGA corrected signal enhancement of (A). Adapted from Sun et al. [2022]

3-6 The NLBF framework for data reconstruction

In the context of this thesis, the NLBF framework presented in Chapter 3 will be used for seismic data reconstruction instead of signal enhancement. In this sense, traveltime operators will be constructed using the introduced NLBF+eGA method. These wavefronts will then be used to reconstruct sparse data according to the method outlined in Chapter 4. Instead of a weighted summation of data traces along wavefront operators to enhance target traces, sparse target traces will be reconstructed by interpolation of data trace values along these traveltime operators. The use of NLBF for seismic data reconstruction is extensively covered in Chapter 4.

Chapter 4

Methods

In this chapter, we present two methods we developed to reconstruct sparse seismic data. The first, the main method of interest, is the use of the NLBF framework for seismic data reconstruction. Figure 4-1 highlights the important difference between the use of NLBF for signal enhancement (Chapter 3) versus reconstruction (Chapter 4). The second method is the BP method, which overcomes the issue encountered by many reconstruction methods with a regular pattern of data sparsity. In this thesis, we will use the BP data reconstruction method presented hereafter as a control method for the main method of interest, NLBF reconstruction. More on this in Chapters 5 and 6.

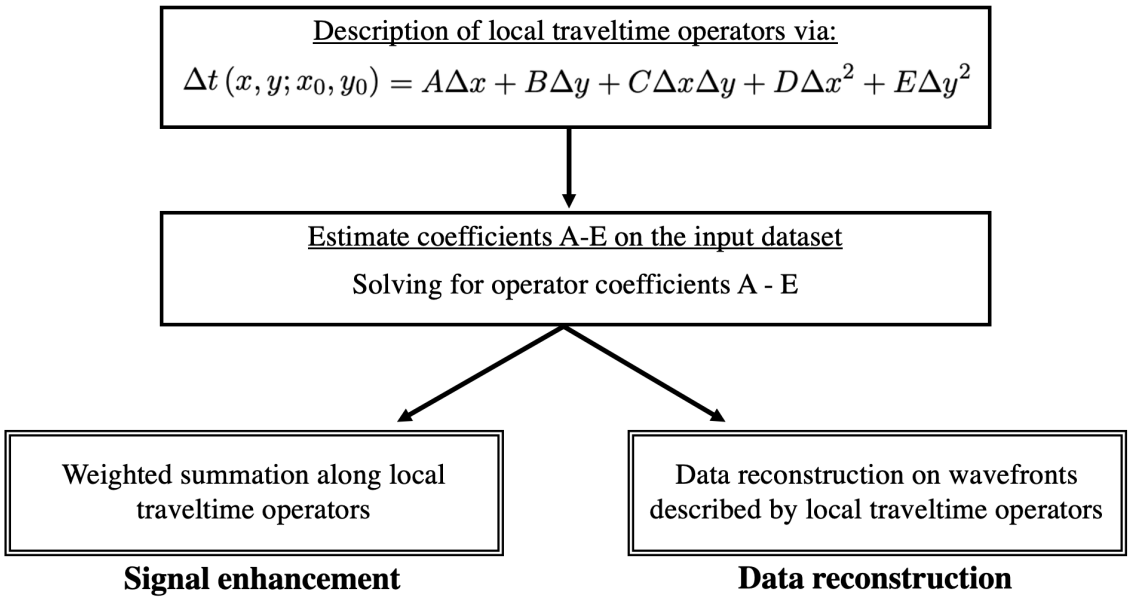

Figure 4-1: A flowchart illustrating the difference and similarities in the use of the NLBF framework for data reconstruction versus signal enhancement.

4-1 NLBF seismic data reconstruction

In the following section, we illustrate the NLBF data reconstruction methodology on a seismic dataset with the acquisition geometry shown in Figure 4-2. Data traces are the ones obtained during the acquisition, and target traces are the ones to be reconstructed. The gather we consider contains a total of 104720 data traces; the inline direction has dense data trace spacing with $\Delta x = 12.5$ m, and the crossline direction has coarser data trace spacing with $\Delta y = 37.5$ m. To achieve equal data trace distribution in both the inline and crossline directions in Figure 4-2, two target traces must be reconstructed between data traces in the crossline direction (gray circles). The data is therefore 2/3 sparse in the crossline direction. Traces have a time sampling rate of Δt_s[1] $= 6$ ms for a total of 690 time samples. Figure 4-3 illustrates a sparse crossline section (A - B from Figure 4-2), where gaps between data traces are apparent. The objective of the NLBF data reconstruction method is to reconstruct all target traces in order to obtain the desired spatial resolution, and for this dataset the goal is to have $\Delta x = \Delta y = 12.5$ m. The NLBF reconstruction steps carried out in this chapter are easily scalable to seismic gathers with different geometries and levels of data sparsity.

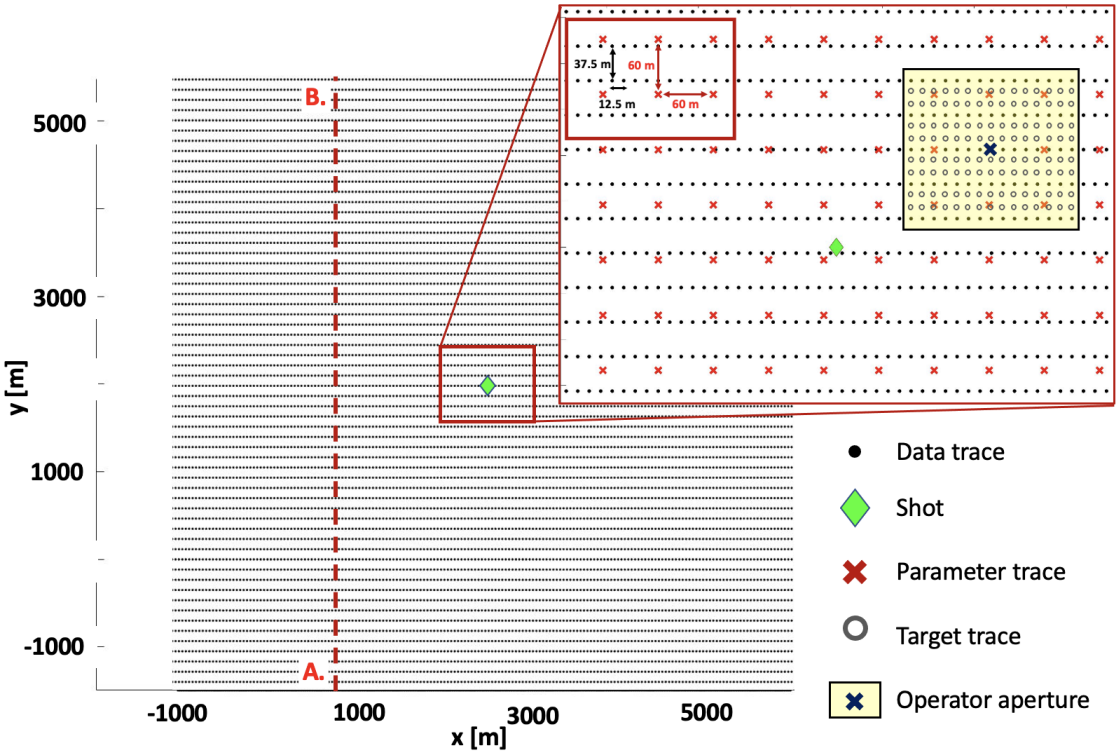

Figure 4-2: Schematic section of the synthetic SEAM Arid data geometry, with relative positions of data traces, target traces, parameter traces, and the shot. Inline and crossline trace spacing are highlighted in the red box, and the operator aperture is indicated in yellow. Note that target traces are only drawn in the operator aperture for figure clarity, but are in reality all throughout the gather.

[1]Note: we distinguish Δt_s as the time sampling rate; and Δt from Equation (3-1) as the traveltime of points on a wavefront.

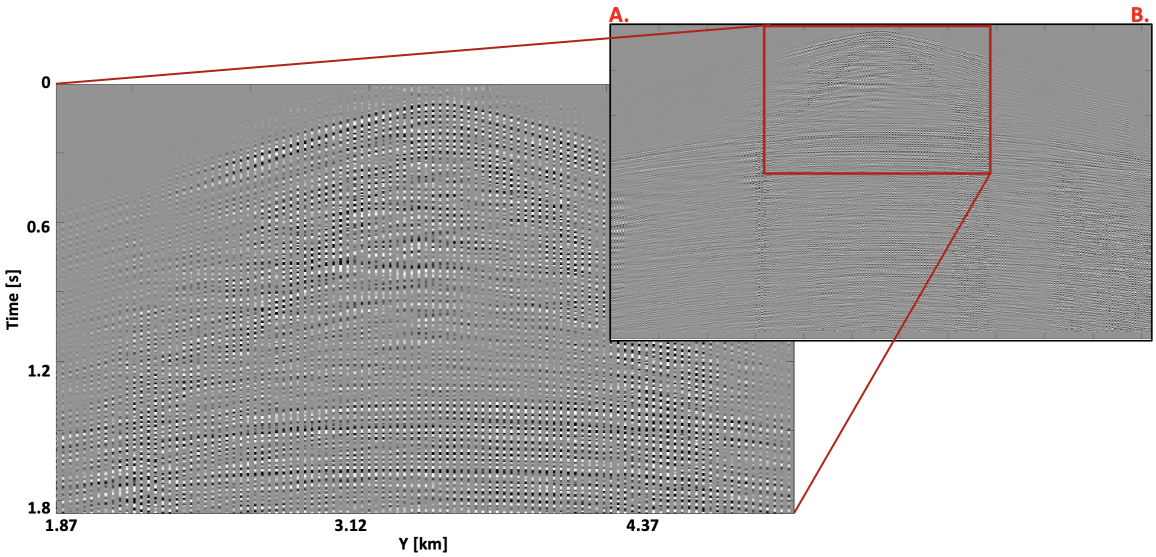

Figure 4-3: A 2/3 sparse synthetic SEAM Arid data section at $x = 750$ m (A-B in Figure 4-2). The red box in the inset highlights the close-up section of the figure. Data trace spacing is $\Delta y = 37.5$ m, separated by two null traces.

4-1-1 Traveltime operators

The first step of the NLBF process is to define traveltime operators we wish to use to describe wavefronts, and then estimate these operators from the dataset of interest (Figure 4-1). The method is identical to the one presented in Chapter 3. We choose to define our traveltime operator as the second-order surface described by Equation (3-1). These operators are defined at parameter traces, which are regularly distributed across the gather (Figure 4-2). The spatial density and distribution of parameter traces define the number of operators involved in the reconstruction. The geometry in Figure 4-2 shows a parameter trace distribution of $\Delta x = \Delta y = 60$ m. Similar to the data traces, parameter traces have a time sampling rate of $\Delta t_s = 6$ ms for a total of 690 time samples. Each time sample has its own operator defined by the parameters A, B, C, D and E obtained through the NLBF+eGA method (Chapter 3) [Sun et al., 2022]. Figure 4-4 illustrates parameters A-E and the resulting semblance values for five parameter trace lines.

The operator aperture defines how far we extend the wavefront around their respective parameter trace. Figure 4-2 illustrates an aperture of 180 m x 180 m. Using the coefficients A-E, the traveltime operators are calculated inside the aperture according to Equation (3-1). Δt describes the traveltime moveout of the wavefront with respect to the current parameter trace time sample (Figure 3-2B). It is calculated at both data and target trace locations, resulting in 3D wavefront surfaces spatially and temporally discretized to the target grid dimensions. The operator aperture in Figure 4-2 highlights these positions. Figure 4-5 illustrates a close-up 3D section of the sparse gather with four wavefronts around their respective parameter traces. Chapter 5 further explores the effects of varying spacing between parameter traces and aperture sizes on the data reconstruction quality.

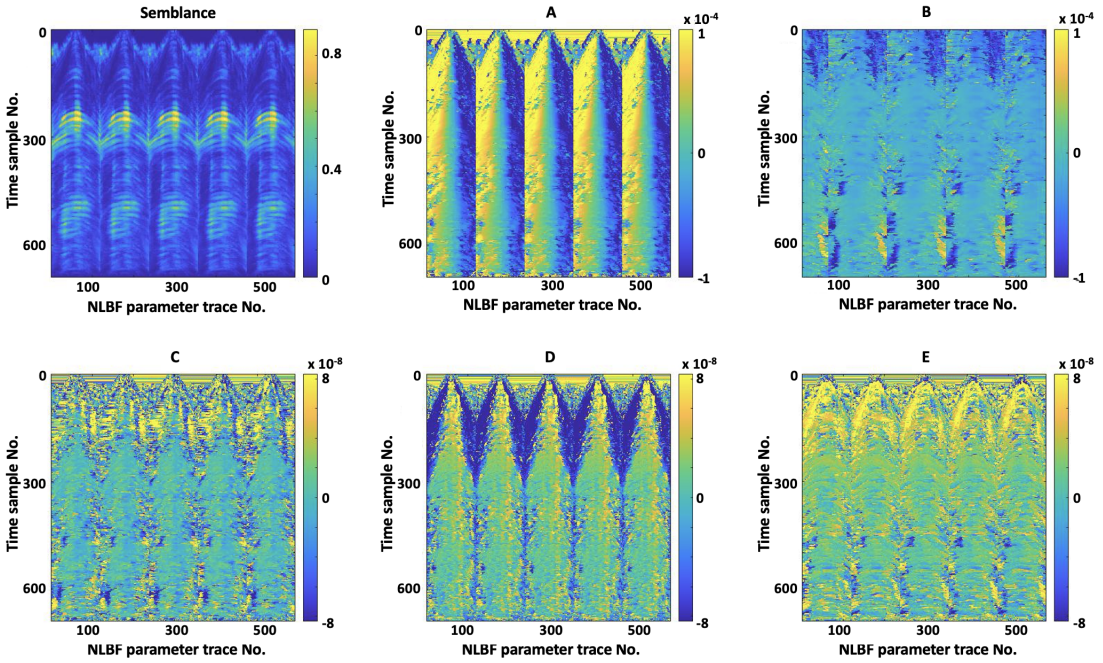

Figure 4-4: NLBF+eGA estimated semblance values and parameters A-E for selected parameter traces (A-B in Figure 5-2) in the synthetic SEAM Arid dataset.

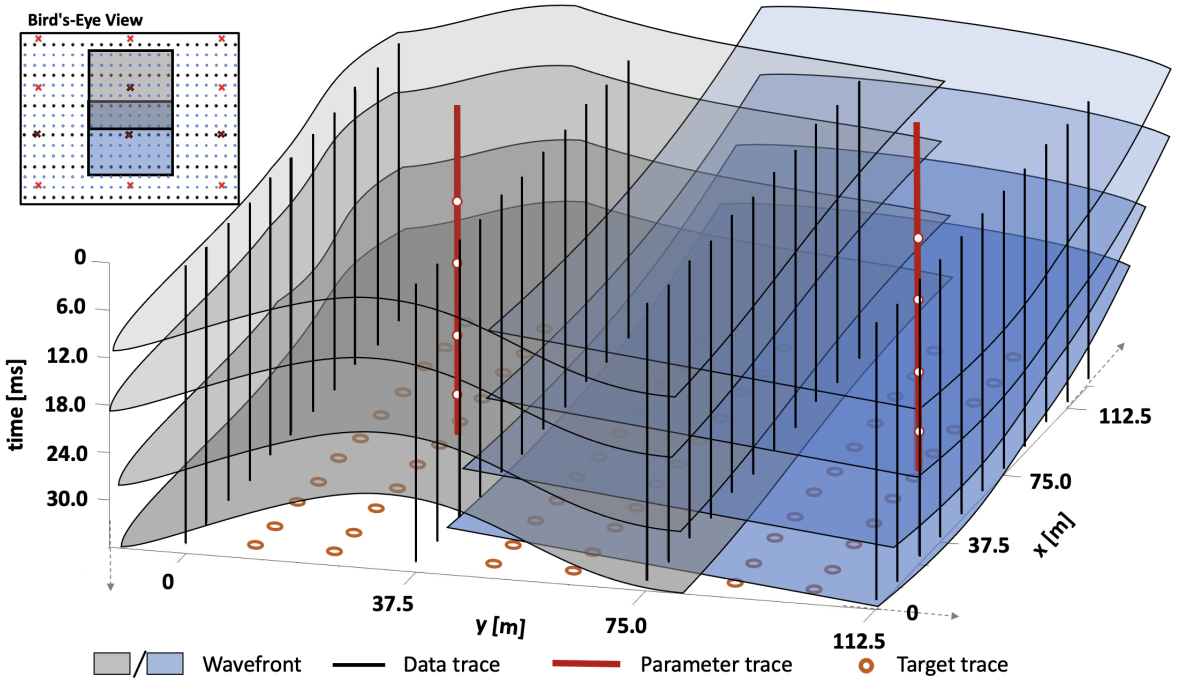

Figure 4-5: Close-up 3D illustration of four wavefronts around two parameter traces (red lines); white dots indicate the operator-parameter trace intersection, and black lines indicate the data traces. The bird's-eye view displays a top-down view of the aperture extent.

4-1-2 Methodology

We use the sparse data shown in Figure 4-2 to demonstrate our reconstruction method step by step. For seismic data reconstruction, the local coherency between traces along kinematic wavefronts allows for the reconstruction of missing target traces in between data traces. The first step to reconstruct sparse seismic data is to reorganize the original sparse data on a denser target grid to the appropriate reconstruction dimensions. The original 2/3 sparse data has grid dimensions of 187 x 560 and is sparse in the crossline direction. A denser target grid is established with the appropriate dimensions. It contains the known data traces, and sparse target traces initially defined as 0-traces. The new target grid dimensions for the synthetic dataset are 559 x 560 with $\Delta x = \Delta y = 12.5$ m in both inline and crossline directions (inside the aperture of Figure 4-2). Parameter traces are kept on a separate parameter grid as discretized trace positions may differ. Traveltime operators are calculated for each parameter trace time sample inside the operator aperture to target grid dimensions.

The reconstruction of target traces within the aperture is performed by repeating the following steps for each operator along all parameter traces. Figures 4-6, 4-7 and 4-9 illustrate these steps in 2D on a single operator for simplicity, but these in reality are applied in the 3D space.

- First, the operator traveltime surface is extended by a multiple of the time sampling rate Δt_s in the temporal dimension in both the positive and negative direction by increments of $\pm\Delta t_s$, as illustrated in Figure 4-6. The time window size is set by the user and ultimately influences the amount of reconstructed target time samples per operator. In our case, these are regular time samples of a $\Delta t_s = 6$ ms time sampling rate along target traces. Only target time samples within this window are reconstructed. Figure 4-6 illustrates a time window of $4\Delta t_s$, and the original operator at 12 ms on the parameter trace (red) is extended by $2\Delta t_s$ in the positive and negative time directions. Chapter 5 further explores the effects of the time window size on reconstruction quality.

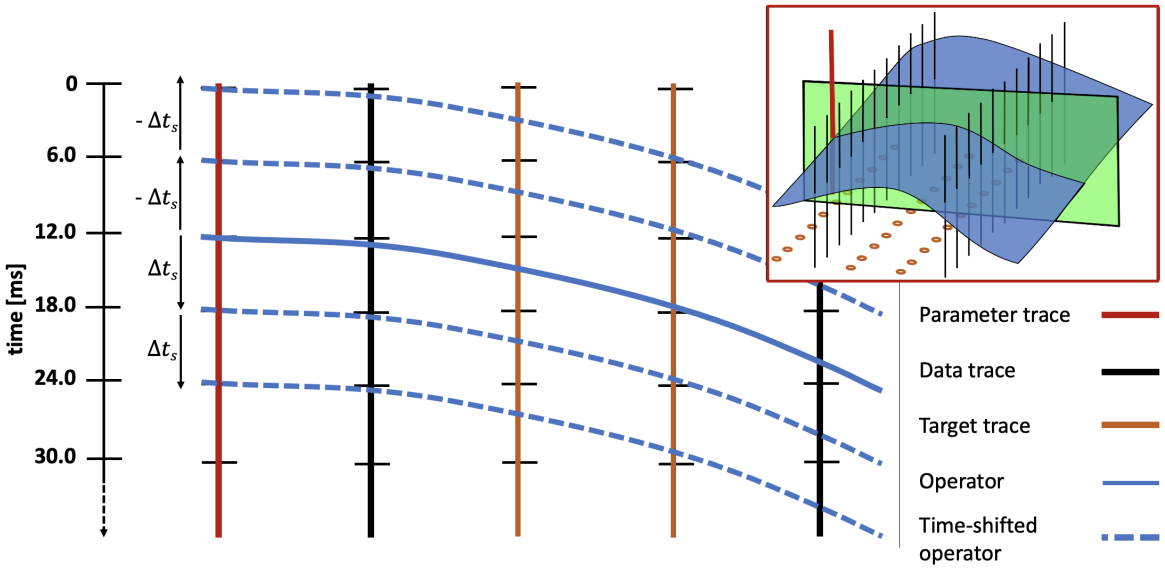

Figure 4-6: Extension of the operator at 12 ms to the desired time window of $4\Delta t_s$.

- Subsequently, the corresponding wavefield amplitude values of the extended traveltime window are linearly interpolated along data traces in the temporal direction (Figure 4-7). This is possible as we know the data trace amplitude at discretized time samples, and the traveltime operator values Δt at data trace locations via Equation (3-1). Figure 4-7 illustrates the interpolated amplitude points as the black dots and the interpolation direction as the black arrow. The inset of Figure 4-7 only shows one wavefront for clarity, but in reality, the amplitudes of all wavefronts separated by Δt_s in the traveltime window (dotted lines) at data trace locations are also calculated. This results in multiple 2/3 sparse wavefront surfaces with regular decimation.

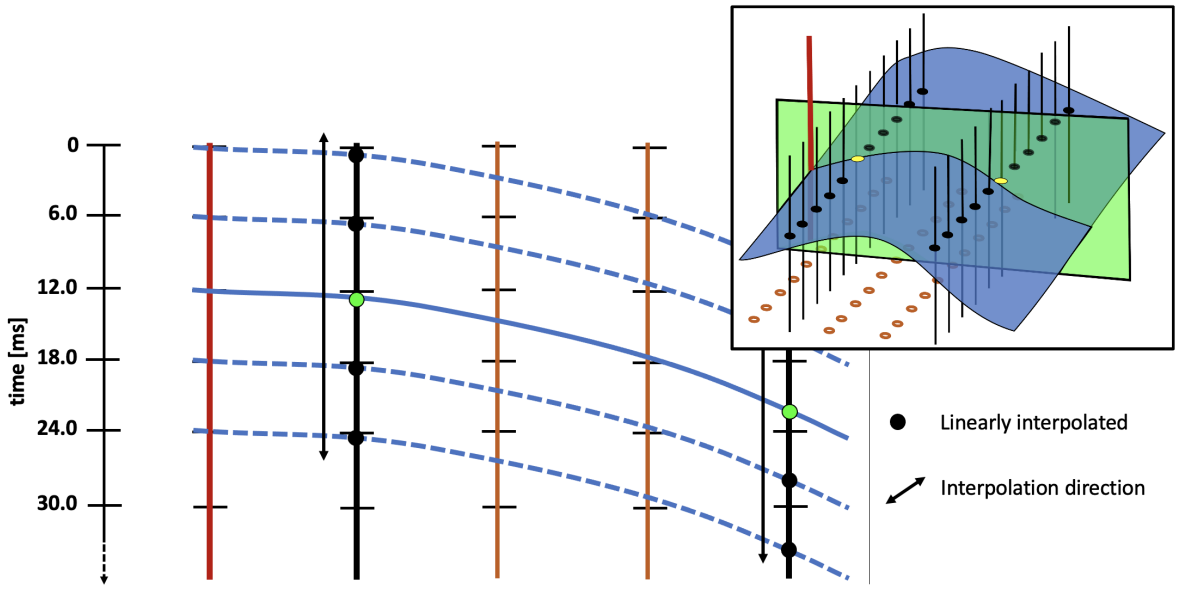

Figure 4-7: Illustration of the linear interpolation along the data traces to obtain the operator amplitudes (black dots). The inset only contains the full (non-dotted) line operator in the 2D section for clarity; the yellow dots indicate the interpolated operator values in the inset cross-section. See trace legend in Figure 4-6.

- The interpolation of wavefronts amplitude values at target trace locations is performed via both linear and bilinear interpolation along the wavefronts, using the previously calculated amplitude at data trace locations (Figure 4-8). Each target trace time sample receives equal (50/50) contributions from both linear and bilinear interpolation results, and this allows the removal of the interpolation direction bias. Figure 4-8 illustrates both linear and bilinear interpolation separately. Linear interpolation is performed for all lines in the crossline direction to obtain target trace amplitudes, as illustrated by the black arrows and orange dots in Figure 4-8A. Bilinear interpolation uses results from the linear interpolation to interpolate in the inline direction, as illustrated by the black arrow and blue dots in Figure 4-8B. This linear-bilinear interpolation is performed along all wavefronts in the time window (Figure 4-6), resulting in multiple reconstructed wavefronts to the target dimensions. This wavefront interpolation method allows us to obtain fast results. However, the BP method explained hereafter shows potential for a higher quality interpolation at the expense of computational efficiency. More on this in Chapter 5.

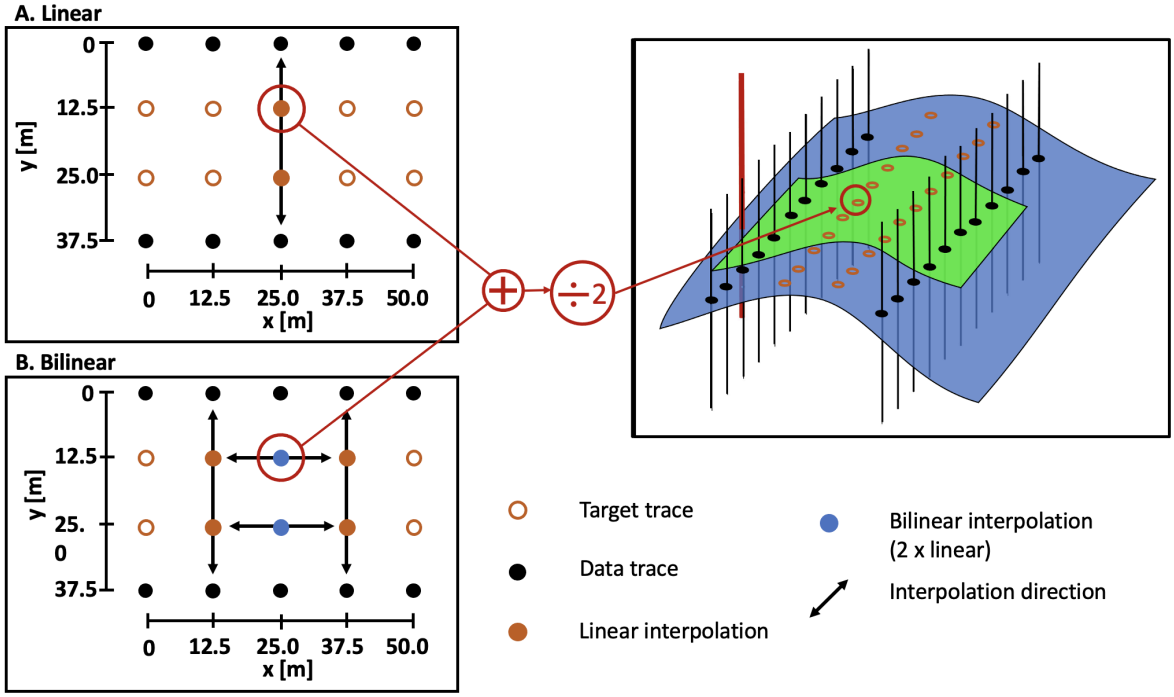

Figure 4-8: Linear and bilinear interpolation of operators at target trace positions. (A) Linear interpolation in the crossline direction. (B) Bilinear interpolation in the inline direction.

- Finally, using the interpolated wavefront amplitudes at target trace locations (blue dots in Figure 4-9), amplitudes at the desired target time samples are interpolated (yellow dots in Figure 4-9). A cubic-spline interpolation is used to better conserve trace continuity, as the interpolated value is based on a piecewise cubic interpolation of the values at neighboring grid points. Figure 4-9 illustrates the cubic-spline interpolation at desired target time intervals of the original input data.

These steps are repeated for all operators along the parameter traces. For each operator, the operator time window covers multiple target time samples (Figure 4-6), and these may overlap between neighboring operators. Interpolated amplitudes at target time samples are summed and weighed by the number of contributions. In this sense, target time samples may receive contributions from several operators. This process is repeated for all parameter traces. These target time samples from individual parameter trace sections may also coincide where two operator apertures overlap. The amplitude values are again summed and weighted to create the total reconstructed wavefield, which allows for a seamless assembly of the reconstructed parameter trace sections. The amount of overlap is influenced by the size of the operator aperture and the distance between neighboring parameter traces. For the parameters used in this chapter (60 m parameter trace spacing; 180 m x 180 m operator aperture), we expect a crossover of 120 m. A sufficient crossover is required to fully reconstruct the entire gather. Chapter 5 covers results of varying cross-over distances on the reconstructed gathers.

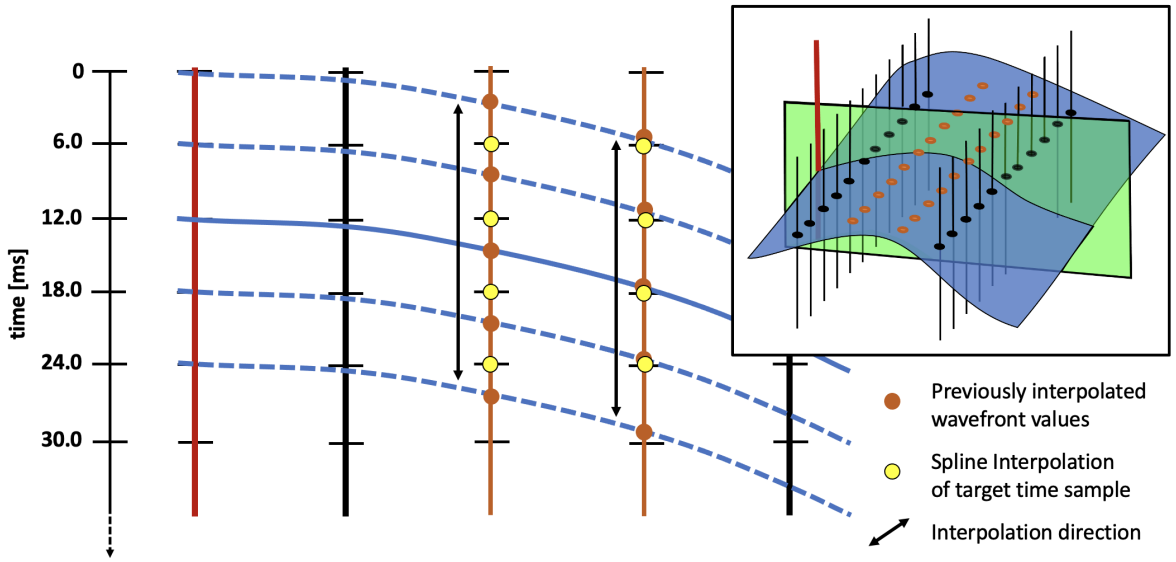

Figure 4-9: Cubic spline interpolation for one operator at the desired discretized target trace time samples, and the yellow dots indicate the interpolation locations. See trace legend in Figure 4-6.

4-2 Bootstrap POCS (BP)

In this section, we present the bootstrap POCS (BP) method. We propose this method to overcome the limitations of the conventional CP scheme on regularly sparse data (Chapter 2). We define regular sparsity as a repeating pattern of missing traces or data points. Our research shows a high-quality image reconstruction potential with applications beyond the scope of this thesis. More on this in Chapter 6.

4-2-1 Regularly decimated data

The CP method, an improved version of Abma and Kabir's POCS method [Abma and Kabir, 2006] (Chapter 2), showed impressive image reconstruction results [Hirsch, 2022]. However, the application of $f-k$ transform-based reconstruction methods on regularly sparse images poses a challenge. A regular pattern of dead traces is very predictable and is retained in the output unless the pattern is forcibly removed [Abma and Kabir, 2006]. Figure 4-10 shows the results of the CP algorithm on the 2/3 regularly decimated drop image. It is apparent that the CP reconstruction does not converge to the desired reconstructed output. It has a signal-to-noise ratio (SNR) of 1.61 which will be further discussed in Chapter 5.

In fact, the regularity of the decimation in the space-time $(x-t)$ domain leads to aliasing in the $f-k$ domain. Figure 4-11B shows the $f-k$ spectrum of the original image (Figure 4-11A), and Figure 4-11D shows the $f-k$ spectrum of the 2/3 sparse image (Figure 4-11C). The repeated spectra in Figure 4-11D illustrates the aliasing effect of the regular decimation. This aliasing affects the $f-k$ thresholding of the CP method (Chapter 2) and therefore leads to the poor reconstruction result we see in Figure 4-10B.

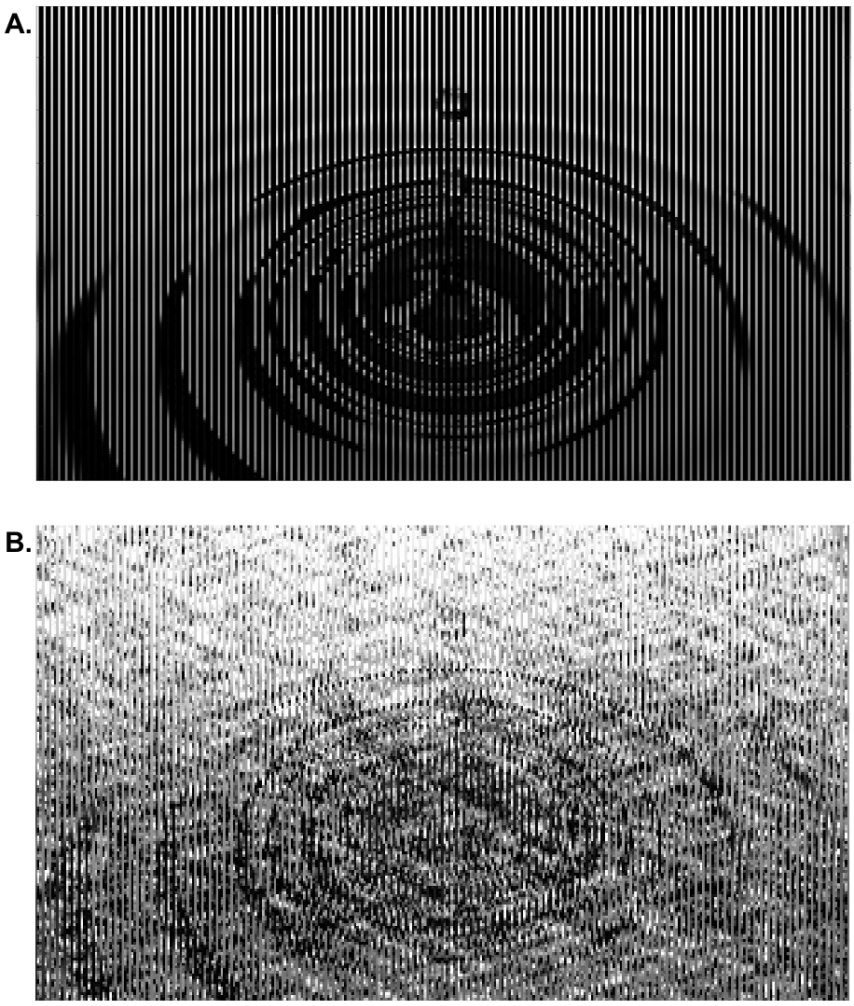

Figure 4-10: (A) A $2/3$ regularly decimated image of drop image from Chapter 2. (B) Results of the CP reconstruction after 100 thresholding iterations using parameters from Hirsch [2022] (SNR = 1.61).

4-2-2 Bootstrap POCS (BP) methodology

The BP method is able to but is not limited to reconstruct regularly sparse images. The first step of BP is to reconstruct the regular sparse image using an interpolation method compatible with a regular decimation pattern. Here, we use a combination of linear and bilinear interpolation, identical to the method presented in Figure 4-8, to obtain contributions from data points in all directions. This is the starting point of the BP method, i.e., the bootstrapping step. The bootstrapping step provides support to the CP method's limitations with regular decimation. Once the image has been coarsely reconstructed by the bootstrap step, we randomly decimate some reconstructed data points and then use the CP method to reconstruct this newly created sparse image. In a nutshell, the BP method uses the CP method as the major reconstruction scheme, and it uses the bootstrapping step to compensate for the CP method's limitation with regular data decimations.

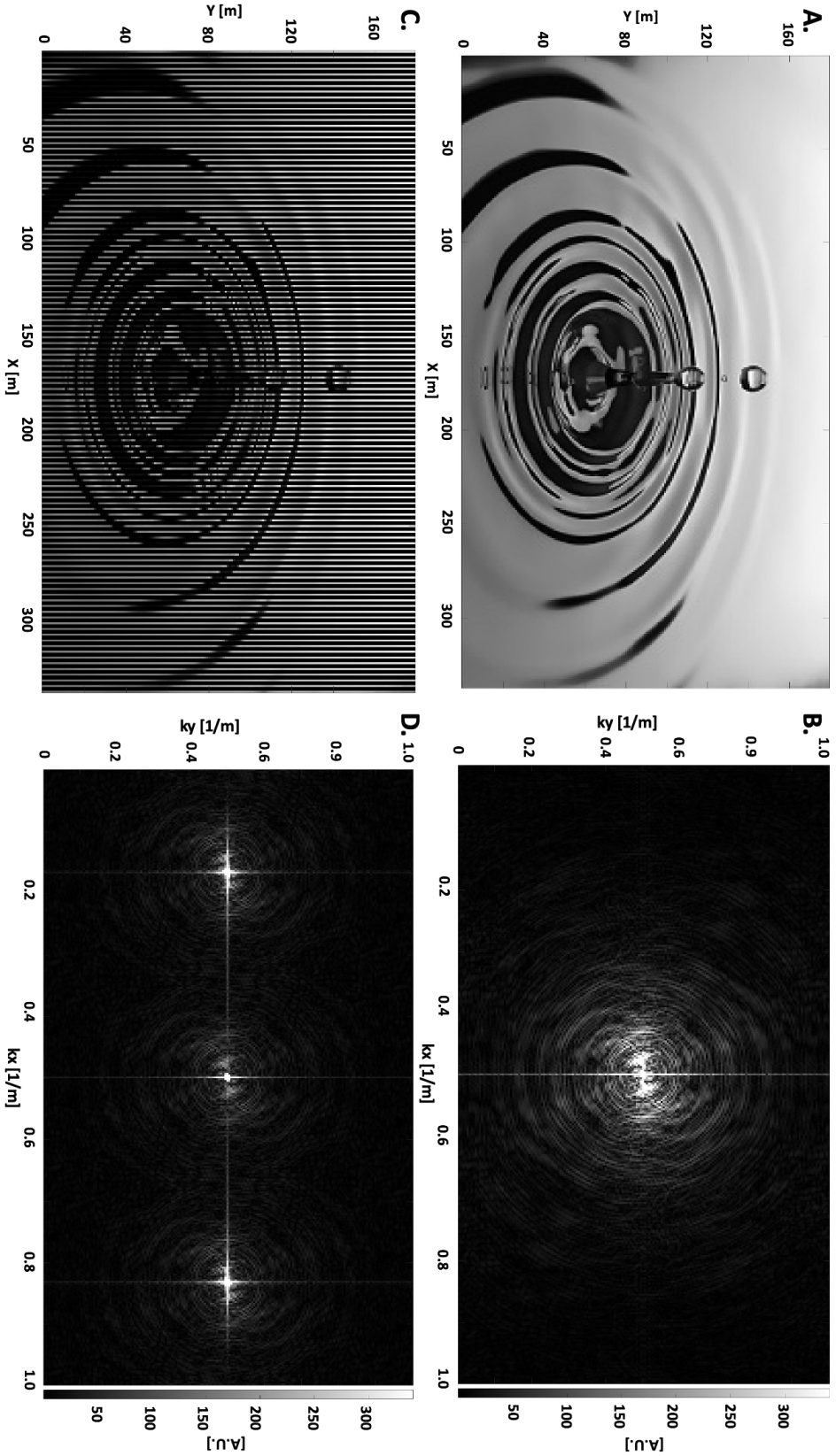

Figure 4-11: (A) Original and (C) $2/3$ sparse drop images with their respective $f-k$ spectra (B) and (D).

Figure 4-12 demonstrates the BP method on a small schematic grid with 2/3 sparsity in the crossline direction, similar to our drop image. After reorganizing our sparse image to a target grid with the appropriate reconstruction dimensions, we create the linear-bilinear bootstrap image. Figure 4-12A shows the original pixels in gray, and the linear-bilinearly interpolated pixels in blue. We then randomly decimate a user set percentage of these coarse linear-bilinearly reconstructed pixels as illustrated by the white pixels in Figure 4-12B. Subsequently, we reconstruct these white pixels using the CP method, resulting in the red pixels in Figure 4-12C. We say these red pixels are BP reconstructed as they use the bootstrap - linear-bilinear - image as support for the CP method. This process is repeated until all the linear-bilinear interpolated pixels have been replaced by the BP reconstructed ones (Figure 4-12F). The gray ground truth pixels are preserved throughout the process.

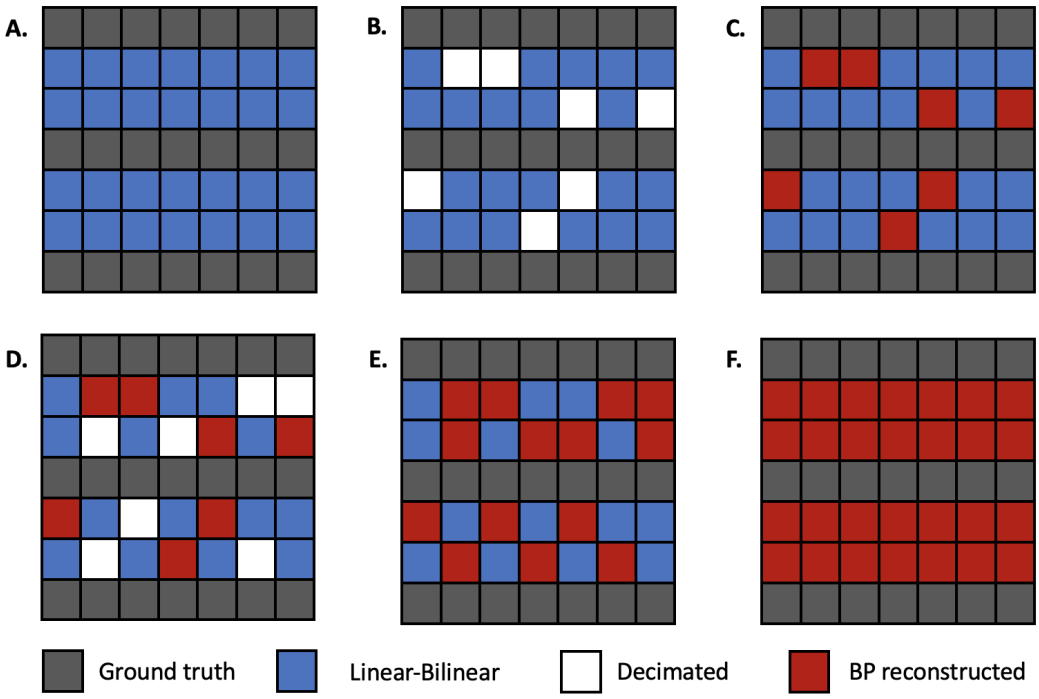

Figure 4-12: Demonstration of the BP method. (A) The linear-bilinear interpolated image, (B) 25% random pixel decimation, (C) reconstruction using the CP method, (D-E) the process is repeated, and (F) all linear-bilinear pixels are BP reconstructed.

To illustrate the potential of the BP method, we apply the presented methodology to the 2/3 sparse drop image in Figure 4-10. We use a 5% pixel decimation value for each BP iteration, 1000 thresholding iterations, and the POCS parameters presented in Hirsch [2022]. Figure 4-13 illustrates the resulting BP reconstruction. We use an SNR to quantitatively assess the quality of the reconstruction (Equation 5-1). Visual comparison of the BP and CP (Figure 4-10B) reconstruction results to the original image (Figure 4-11A), highlights the closer resemblance of the BP reconstruction. The superior reconstruction quality is confirmed by the SNR values, as the BP reconstruction has an SNR of 9.24, and the CP reconstruction has an SNR of 1.61. From this result, it is clear that the BP method presents impressive image reconstruction abilities. In the following chapter, we use the BP method as a control method for the NLBF reconstruction on seismic data.

Figure 4-13: A BP reconstructed drop image (SNR = 9.24), showing a high fidelity to the original image Figure 4-11A.

Chapter 5

Results

In the following chapter, we present the results of the NLBF seismic data reconstruction method on the synthetic SEAM Arid and field datasets. First, we will introduce the two sparse seismic datasets used to test our data reconstruction methods, and the control methods. Then, we present the parameter tests and results of the NLBF reconstruction.

5-1 The data

5-1-1 SEAM Arid dataset

The first dataset is a synthetic seismic gather generated from a SEAM Arid model [Oristaglio, 2015]. Figure 5-1 shows a 2D slice from the 3D SEAM Arid velocity model.

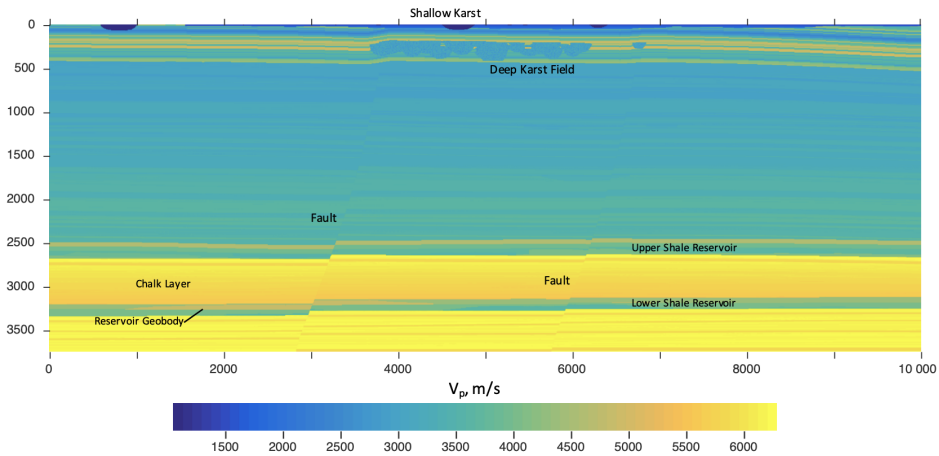

Figure 5-1: A velocity-model section through the center of the 3D SEAM Arid model [Oristaglio, 2015]. The corresponding density model is not shown.

A 3D finite-difference simulation engine was used to generate the synthetic data from the SEAM Arid model, and $f-k$ based pre-processing was performed to remove strong ground-roll noise. Figure 5-2 illustrates the data acquisition geometry in the shot domain. The gather contains a total of 104720 traces organized on a 187 x 560 grid, with inline spacing $\Delta x = 12.5$ m and cross-line spacing $\Delta y = 37.5$ m, resulting in 2/3 spatial sparsity in the crossline direction. We aim to reconstruct the data to target grid dimensions 559 x 560, such that $\Delta x = \Delta y = 12.5$ m. NLBF parameters A-E were estimated by the NLBF+eGA method (Chapter 4) using a time window size of $7\Delta t_s$ and an operator aperture A_{op} of 600 m x 600 m [Sun et al., 2022]. The parameter traces are regularly distributed in the inline and crossline directions such that $\Delta x = \Delta y = 60$ m. Figure 4-4 (Chapter 4) illustrates five parameter trace lines from the SEAM Arid dataset highlighted by the blue A-B lines in Figure 5-2. Both data and parameter traces have a time sampling rate of $\Delta t_s = 6$ ms for a total of 690 time samples.

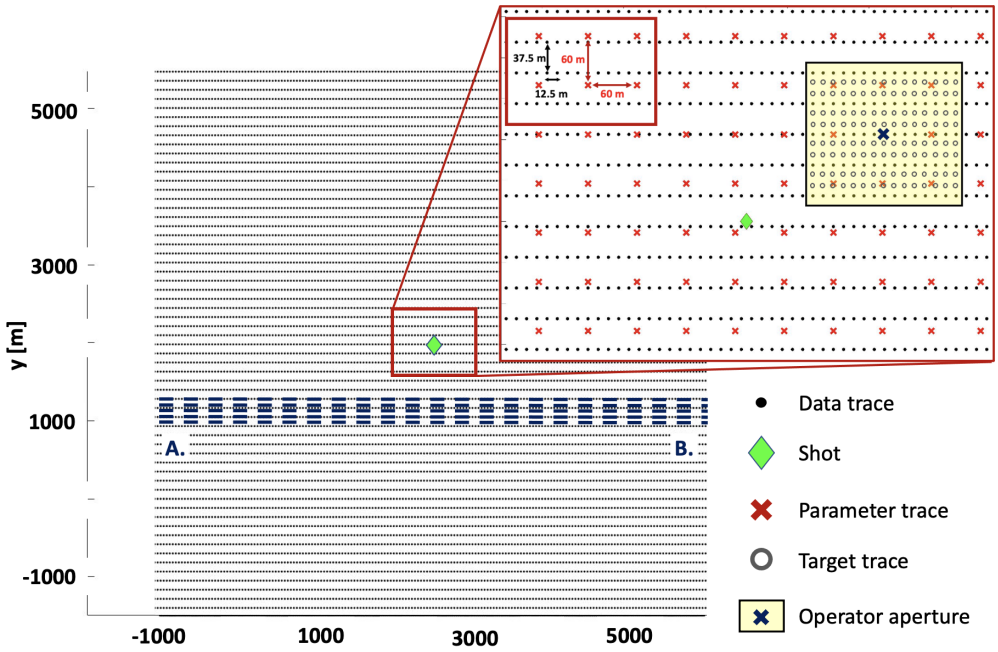

Figure 5-2: Synthetic data from a SEAM Arid velocity model. Trace geometry is highlighted in the inset. Target traces are only drawn in the operator aperture for figure clarity. The dark blue A-B lines indicate the parameter trace positions in Figure 4-4.

5-1-2 Field dataset

The second dataset is a land field dataset, of which the acquisition is characterized by a high channel count and small receiver arrays. Preprocessing followed the standard onshore data denoising workflow, including $f-k$ based ground-roll removal. Figure 5-3A illustrates the acquisition geometry, composed of two orthogonal lines of 111 sources and 194 receivers. The resulting cross-spread domain gather in Figure 5-3B contains a total of 29770 traces, and gaps up to 240 m in the crossline direction. We regularize the data to a 124 x 180 grid using $\Delta x = \Delta y = 60$ m, and we average binned traces to a total of 22320 traces with an increased fold. The dataset is 3/4 sparse, i.e., 1/2 sparse in both the inline and crossline directions ignoring data gaps. We aim to reconstruct the data to a 247 x 359 grid using $\Delta x = \Delta y = 30$ m. Target traces are shown in the operator aperture of Figure 5-4.

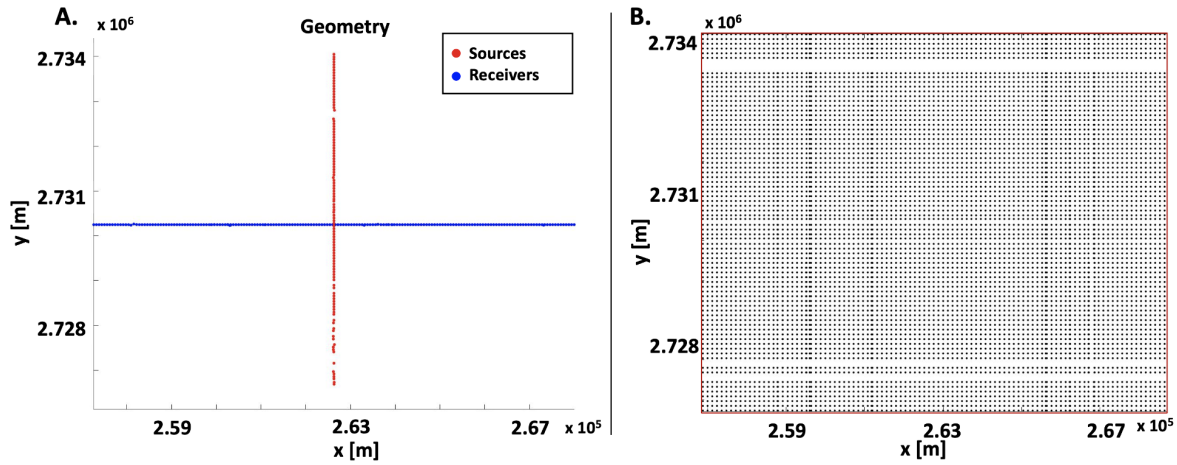

Figure 5-3: Field data acquisition geometry. (A) Orthogonal source and receiver lines, and (B) the resulting cross-spread domain trace geometry.

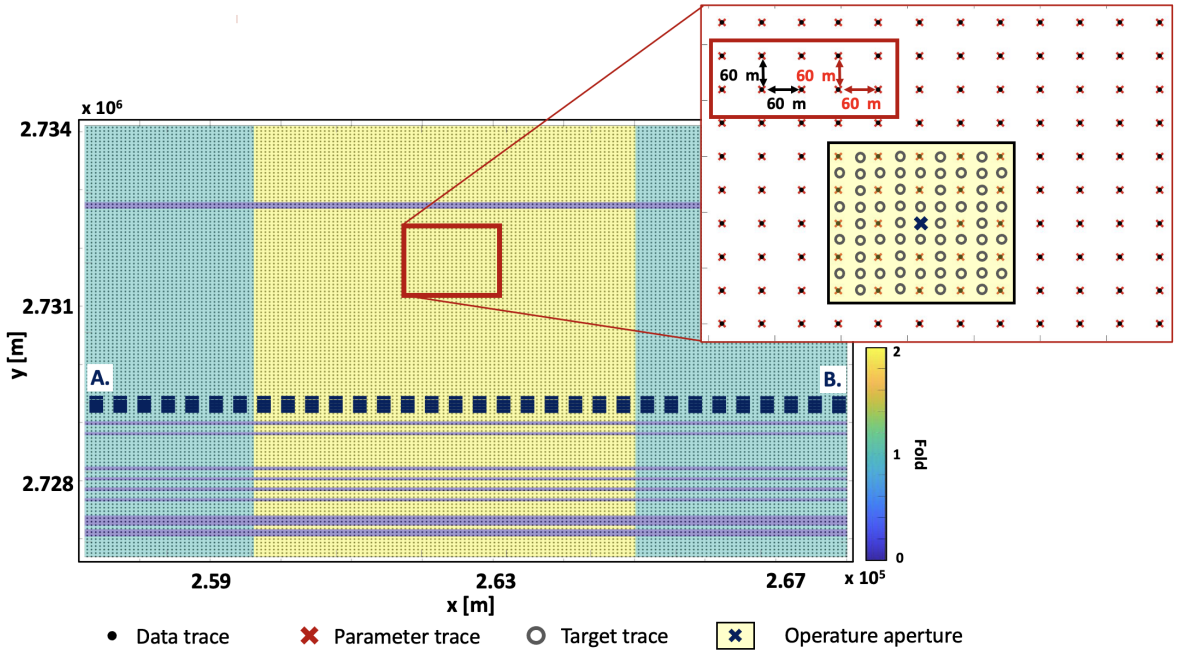

Figure 5-4: The regularized cross-spread geometry for the field data. Trace geometry and fold are indicated. Target traces are only in the operator aperture for clarity. The dark blue A-B lines indicate the parameter trace line positions in Figure 5-5.

The NLBF parameters A-E were estimated by the NLBF+eGA method (Chapter 4), using a time window size of $11\Delta t_s$ and an operator aperture of 600 m x 600 m [Sun et al., 2022]. Parameter traces are regularly distributed in the inline and crossline directions at the same positions as data traces such that $\Delta x = \Delta y = 60$ m, as illustrated in the inset of Figure 5-4. Figure 5-5 illustrates five example parameter trace lines from the field dataset highlighted by the A-B lines in Figure 5-4. Both data and parameter traces have a time sampling rate of $\Delta t_s = 4$ ms for a total of 1250 time samples.

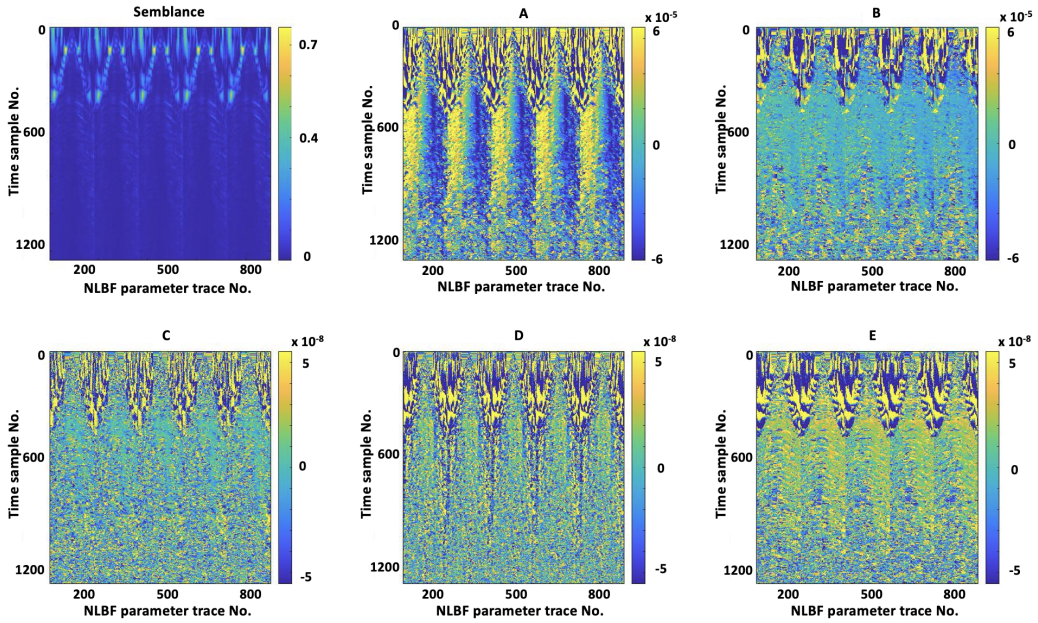

Figure 5-5: NLBF+eGA estimated semblance and parameters A-E for selected parameter trace lines (A-B in Figure 5-4) in the field dataset.

5-2 Parameter tests and control methods

The NLBF data reconstruction method was carried out on the sparse synthetic SEAM Arid and field datasets with varying control parameters. We performed parameter tests on the time window size, the operator aperture size, and the parameter trace interval distance to observe the effects on the data reconstruction quality. Table 5-1 presents the control parameters tested for both datasets. The control parameter test ranges were chosen based on the NLBF parameters, grid size, and data trace density. Additional information regarding control parameters is available in Appendix B.

	SEAM Arid synthetic dataset	Field dataset
Time window [Δt_s]	7, 9, 12, 15, 18	7, 9, 12, 15
Operator aperture [m]	200, 300, 400, 600	300, 400, 500, 600
Parameter trace interval [m]	60, 120, 180	60, 120, 180

Table 5-1: NLBF reconstruction control parameters tested for the SEAM Arid and field datasets.

In order to demonstrate the efficacy and efficiency of the NLBF data reconstruction, we introduce several control methods. These are used to better measure the performance of the NLBF reconstruction. The first control method is the CP method presented in Chapter 2, which will be used for both the SEAM Arid and the field datasets. For the synthetic SEAM Arid model, the availability of the ground truth allows us to directly compare our reconstruction results with the target seismic gather, so the ground truth will serve as the second control method for the SEAM Arid dataset and be used to quantify the quality of the reconstruction. For the field dataset, no ground truth data is available to assess the reconstruction quality, therefore the BP method (Chapter 4) is used as the second control.

In addition to these control methods, we also use a numerical signal-to-noise ratio (SNR) as a quantitative measure when the ground truth is available. The following SNR formula from Wang [2016] is used:

$$\text{SNR} = 20 \log_{10} \frac{\|d_0\|_2}{\|d_{\text{rec}} - d_0\|_2} \quad , \tag{5-1}$$

where d_0 is the complete ground-truth data and d_{rec} is the reconstructed data. Natural wavefield scattering noise in the original data gather d_0 is characterized by strong amplitudes that cannot be suitably reconstructed because of the smooth nature of the NLBF operators [Sun et al., 2022]. Therefore, in our SNR calculations, we do not consider the contributions of this strong scattering noise. It is also important to note that SNR can be defined in many ways, and the SNR collapses the reconstructed gather quality into a single value. Thus it is only a relative average indicator of reconstruction quality, used to differentiate between parameter test results, and does not indicate the quality of the reconstruction in all locations.

Using these control methods, we hope to convey the excellent quality of the NLBF reconstruction and find the best possible set of reconstruction parameters for the datasets of interest. First, we will find the best set of parameters for the SEAM Arid dataset using the SNR. Based on these results and knowledge of parameter interplay (Appendix B), we will choose a set of parameters for the field data, and compare the results to those from the control methods.

5-3 Results

In this section, we present our NLBF reconstruction results for the synthetic SEAM Arid and field datasets and compare them to the control methods. We briefly discuss the rationale behind our parameter decision and their influence on the reconstruction quality. Detailed presentation of the tests and the interplay among parameters are available in the Appendices.

5-3-1 SEAM Arid results

For the synthetic SEAM Arid dataset, the best SNR value was obtained using the following parameters: a time window of $12\Delta t_s$ (Δt_s = 6 ms), an operator aperture of 600 m x 600 m and a parameter trace interval of 60 m. The resulting gather has an SNR value of 5.716. Again, it is important to emphasize that the SNR is calculated for the entire gather regardless of local reconstruction quality, and therefore is only an average indicator. A full presentation of SNR values for all control parameter tests is available in Appendix B. Figures 5-6 and 5-7 illustrate these results for a time section at t = 3.006 s and an inline section at y = 3587.5 m respectively, and compare them to their sparse and ground truth equivalents. They demonstrate the successful reconstruction of the 2/3 sparse target traces in the crossline direction. The reconstructed time section shows clear wavefront reconstruction between the available data traces with an SNR of 12.94 (Figure 5-6C); and the inline section shows a fully reconstructed line of target traces with an SNR of 12.41 (Figure 5-7C). Figures 5-6D and 5-7D illustrate the differences between the NLBF reconstruction and the ground truth. They show the influence of high amounts of scattering noise, emphasizing the good reconstruction quality of the NLBF method. These results and the associated SNR values display the impressive reconstruction ability and the denoising properties of the NLBF method. Additional result sections, including example crossline sections, are available in Appendix A-1.

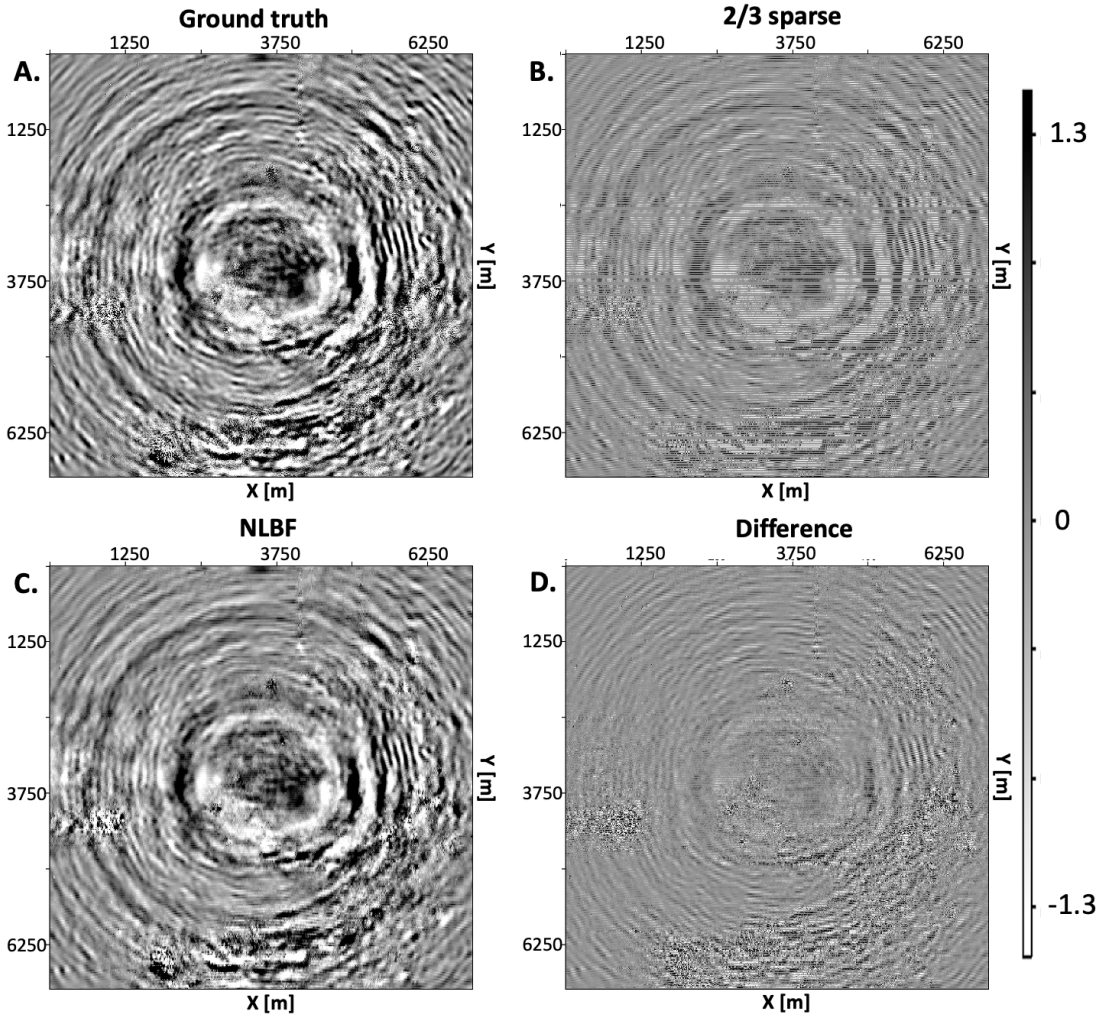

Figure 5-6: NLBF reconstruction for a time section at $t = 3.006$ s on the SEAM Arid dataset. (A) The ground truth, (B) the 2/3 sparse input data (SNR = 1.32), (C) the NLBF reconstructed section (SNR = 12.94), and (D) the difference between (A) and (C).

Figures 5-8 and 5-9 compare our NLBF reconstruction results to the CP control method, for a time section at $t = 3.0$ s and a crossline section at $x = 1500$ m, in the $t - x$ and $f - k$ domains. Visually, the improvement in reconstruction quality of the NLBF method is striking. For both sections, the $t - x$ domain of the CP reconstruction is barely recognizable, whereas the NLBF reconstruction resembles the ground truth. This is confirmed by our SNR values. The SNR of the reconstructed time sections is 1.32 for the CP method and 12.94 for the NLBF method (Figure 5-8). The SNR of the reconstructed crossline sections is 0.96 for the CP method and 11.71 for the NLBF method (Figure 5-8). In the $f-k$ domain, the NLBF reconstruction shows good fidelity to the ground truth, whereas the CP spectrum shows high amounts of additional aliased energy. In fact, the $f - k$ domain of the NLBF reconstruction shows good reconstruction of dominant central frequencies (Figures 5-8E and 5-9E). The white-noise-like spatial frequencies, corresponding to the mentioned scattering noise, are less well reconstructed, which is understandable. In the time section, the faint repeating circular pattern hints at slight aliasing in the crossline direction (Figure 5-8E), which is caused by the

small periodic amplitude mismatch between the reinserted and reconstructed traces. Overall, the $f-k$ spectrum of the NLBF reconstruction shows good quality preservation of the signal's energy. The CP reconstruction, however, shows strong added energy and aliasing for the time section (Figure 5-8F), and poor energy preservation and strong aliasing for the crossline section (Figure 5-9F). It is clear from Figures 5-8 and 5-9 and the associated SNR values, that the NLBF method achieves a successful, more accurate reconstruction of the synthetic SEAM Arid dataset compared to our CP control method.

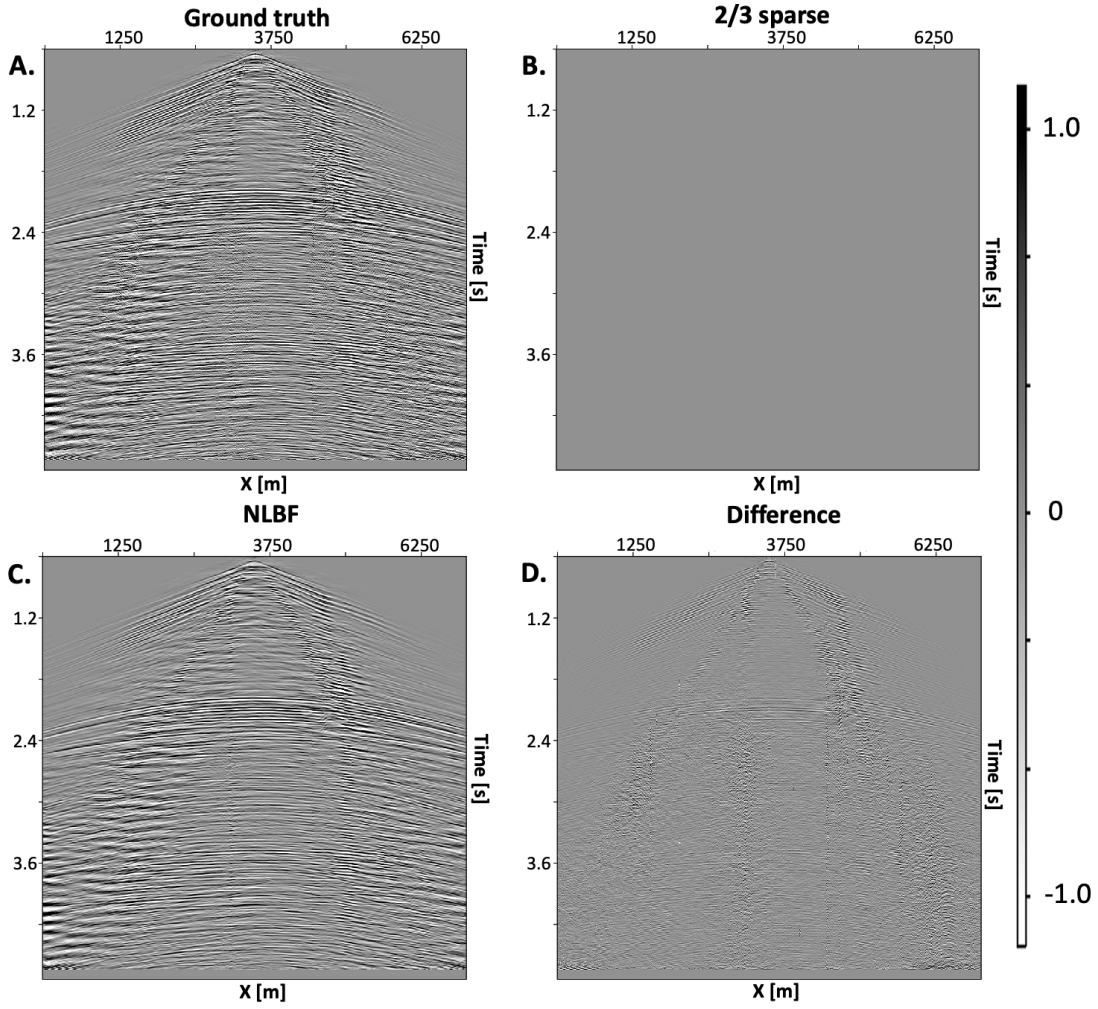

Figure 5-7: NLBF reconstruction for an inline section at $y = 3587.5$ m on the SEAM Arid dataset. (A) The ground truth, (B) the missing section (SNR = 0.0), (C) the NLBF reconstructed section (SNR = 12.41), and (D) the difference between (A) and (C).

Two individual NLBF reconstructed traces in the time domain are shown in Figures 5-10 and 5-11, and compared to the ground-truth traces. The NLBF method shows a high-quality phase reconstruction and good amplitude reconstruction. Completely correct reconstruction of target amplitudes is challenging for all reconstruction methods, therefore results here are satisfying. The trace located at $[x, y = 404, 257]$ has an SNR of 9.68, and the trace located at $[x, y = 83, 420]$ has an SNR of 7.22. These SNR values and the small difference between the NLBF reconstructed and the ground truth traces in Figures 5-10C and 5-11C, confirm that the individual trace quality of the NLBF method is excellent.

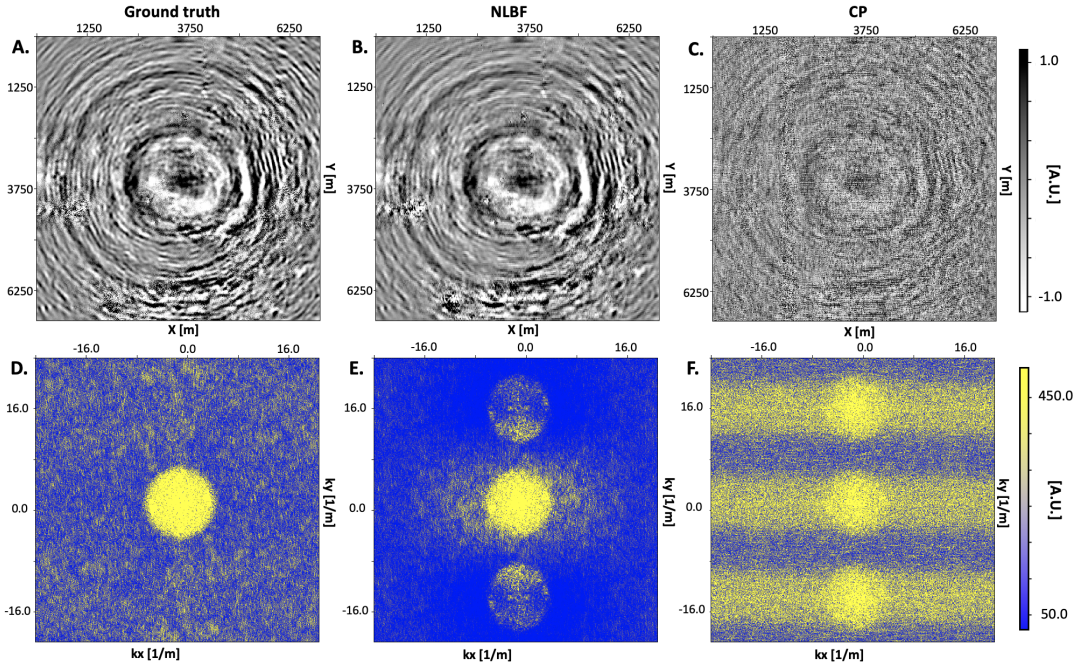

Figure 5-8: Comparison of the NLBF reconstruction with the CP control method in the $x - y$ and $k_x - k_y$ domains for a time section at $t = 3.0$ s on the synthetic SEAM Arid dataset. (A, D) The ground truth, (B, E) the NLBF reconstruction (SNR = 12.94), and (C, F) the CP reconstruction (SNR = 1.32).

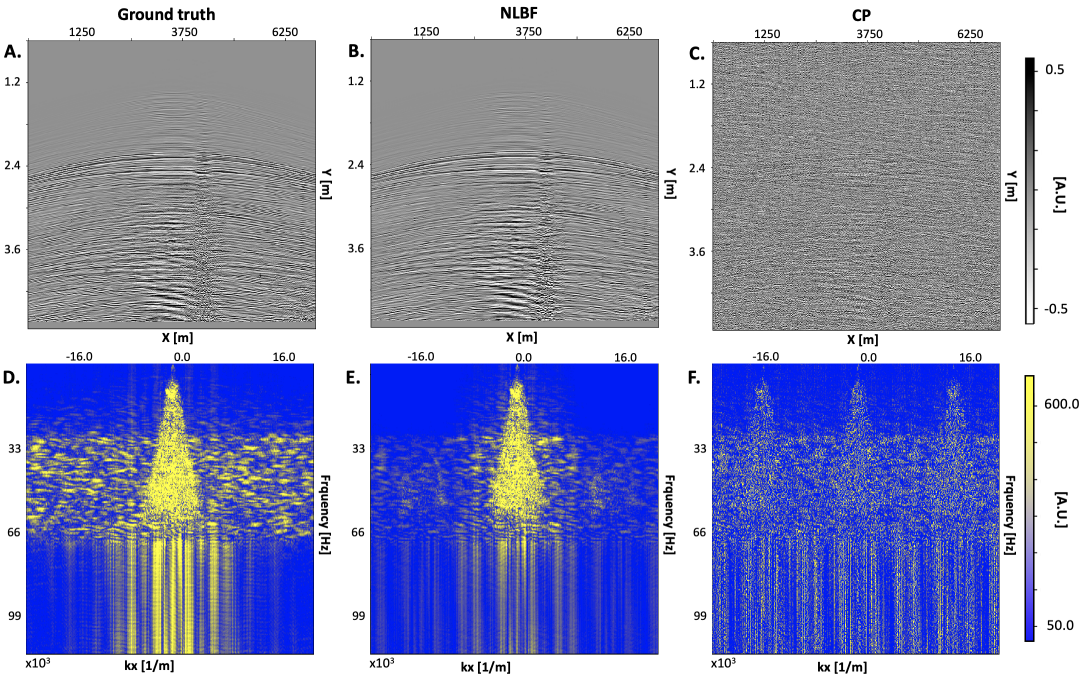

Figure 5-9: Comparison of the NLBF reconstruction with the CP control method in the $t - x$ and $f - k$ domains for a crossline section at $x = 1500$ m on the synthetic SEAM Arid dataset. (A, D) The ground truth, (B, E) the NLBF reconstruction (SNR = 11.71), and (C, F) the CP reconstruction (SNR = 0.96).

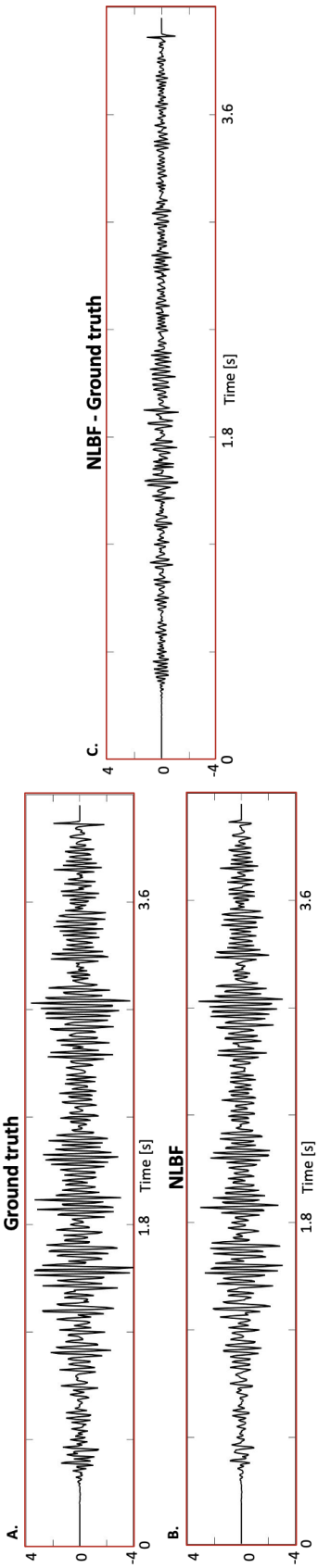

Figure 5-10: NLBF and ground-truth trace comparison at $[x, y = 404, 257]$ on the synthetic SEAM Arid dataset. (A) The ground truth, (B) NLBF reconstructed trace (SNR = 9.69), and (C) the difference between (A) and (B).

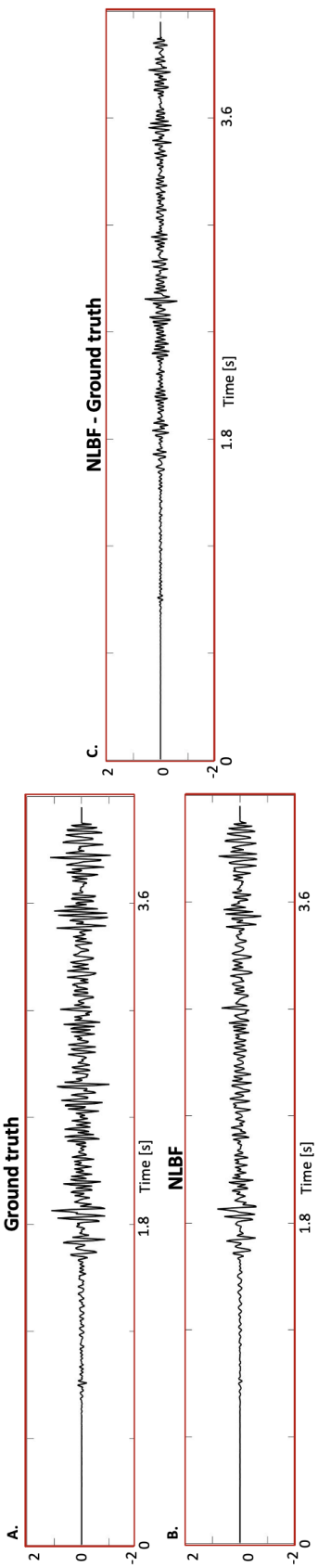

Figure 5-11: NLBF and ground-truth trace comparison at $[x, y = 83, 420]$ on the synthetic SEAM Arid dataset. (A) The ground truth, (B) NLBF reconstructed trace (SNR = 7.22), and (C) the difference between (A) and (B).

5-3-2 Field results

For the field dataset, we use the experience gained on the synthetically controlled SEAM Arid dataset to guide our parameter selections. Optimal parameters are data-dependent as gather geometry plays an important role. For example, critical considerations include a sufficient overlap between neighboring parameter trace operators and sufficient aperture size to cover gaps in the data. Appendix B presents the effects of parameters on reconstruction quality in greater detail. Therefore, for the field dataset we selected the following parameters: a time window of $12\Delta t_s$ ($\Delta t_s = 4$ ms), an operator aperture of 600 m x 600 m, and a parameter trace interval of 60 m. The NLBF reconstruction results are compared to those from the CP and BP control methods as there exists no ground truth for this field dataset. Figures 5-12, 5-13, and 5-14 respectively illustrate a time section at $t = 1.92$ s, an inline section at $y = 3000$ m, and a crossline section at $x = 2190$ m. Additional result sections are available in the Appendix A-2.

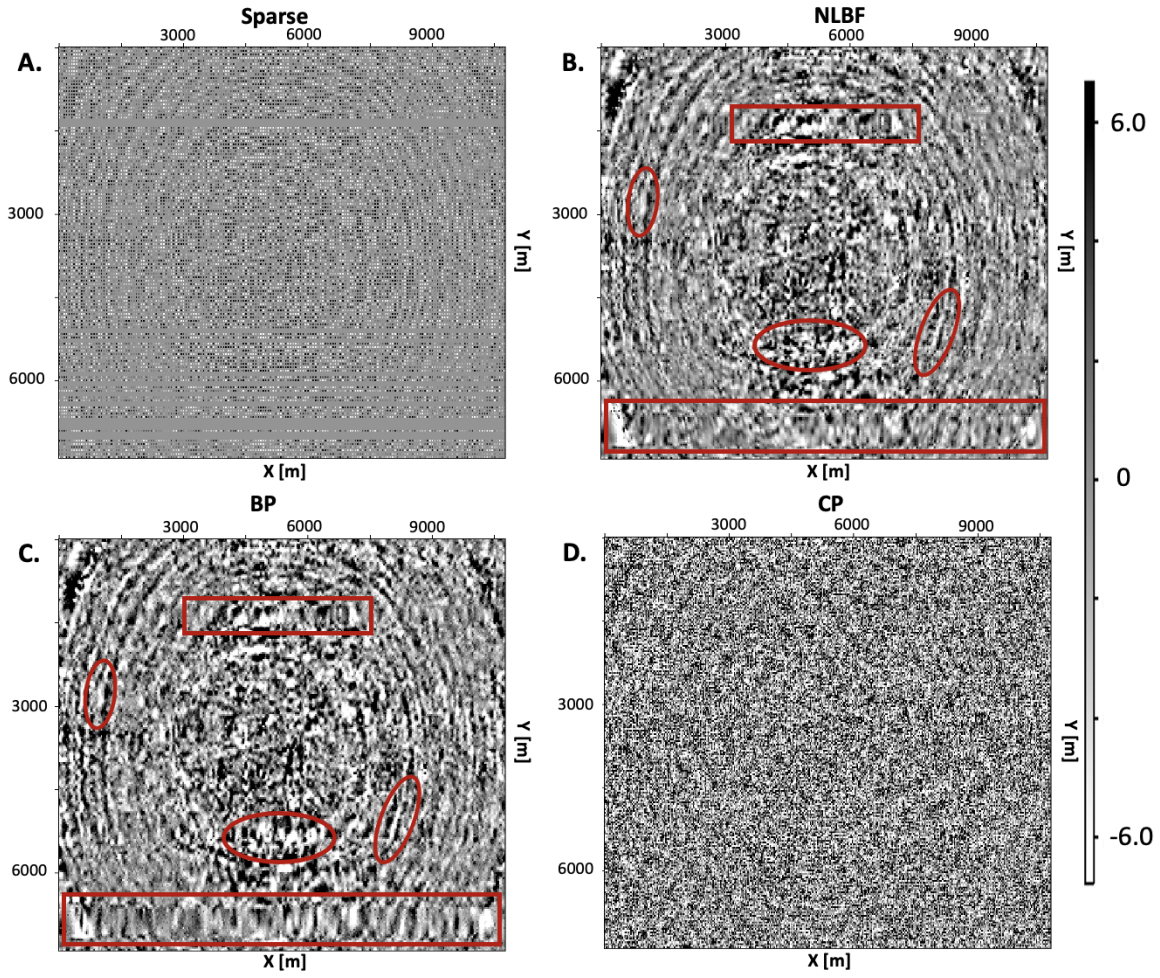

Figure 5-12: Field data reconstruction results of a time section at $t = 1.92$ s. (A) The sparse section, (B) the NLBF reconstruction, (C) the BP reconstruction, and (D) the CP reconstruction. Red circles indicate wavefield differences; red boxes highlight some differences in reconstruction in the gaps.

The NLBF and the BP reconstructed results of the three separate sections (Figures 5-12, 5-13 and 5-14) resemble the expected geometry. The CP method shows very poor reconstruction results for all three sections as there are no distinguishable seismic data features. Figure 5-12 illustrates the field data reconstruction results for a time section at $t = 1.92$ s. Comparing the NLBF and BP results in detail, we first observe that the NLBF wavefields have a smoother change than those from the BP method. By this, we mean that the transition of the wavefield from positive to negative amplitude is more progressive for the NLBF reconstruction compared to the BP reconstruction. Some of these areas are highlighted by red circles in Figure 5-12. These observations are also confirmed in the $x-t$ sections (Figures 5-13 and 5-14), where the NLBF reconstructed curved seismic reflections are more continuous than the BP reconstructed ones, without compromising our ability to distinguish them.

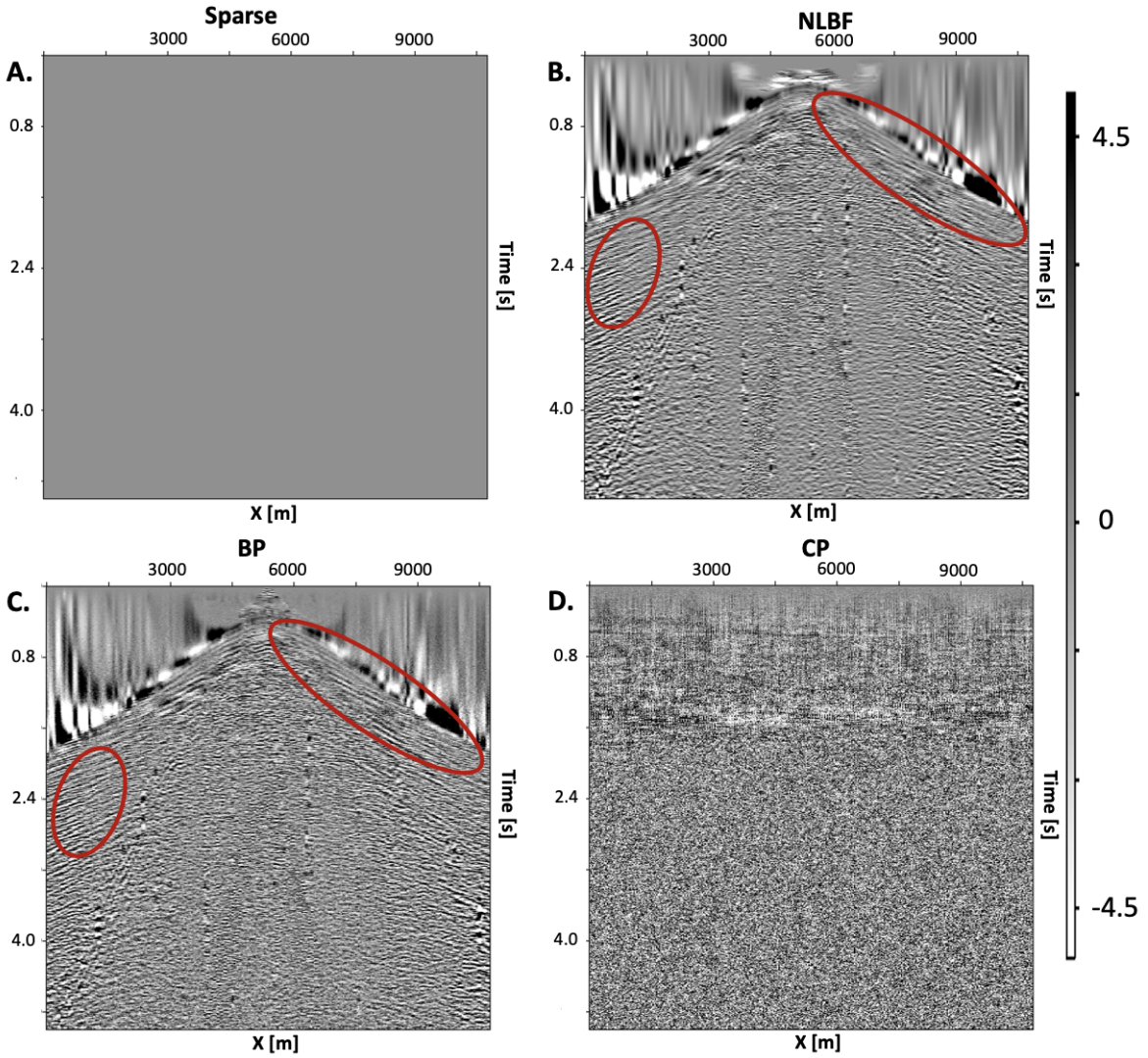

Figure 5-13: Field data reconstruction results of a missing inline section at $y = 3000$ m. (A) The missing section, (B) the NLBF reconstruction, (C) the BP reconstruction, and (D) the CP reconstruction. Red circles indicate some reconstruction differences.

Additionally, a closer look at the reconstructed wavefield in the data gaps shows very different results for NLBF and BP, which are highlighted by the red boxes in Figures 5-12 and 5-14. For example, in the gaps of the BP time section around $y = 6750$ m (Figure 5-12C), we notice that data appears blocky and elongated in the crossline direction. In the NLBF reconstruction (Figure 5-12B) the wavefield is more continuous with its bordering data. We can also observe finer details in the wavefield variations, which are highlighted by the red circles in Figure 5-12. In the crossline section gaps, the BP reconstructed reflections are horizontal and appear strongly artificial, which further deteriorates the reflection quality. Since such gaps are common in all field data acquisitions, we believe that the NLBF method presents advantages to reconstruct field seismic data.

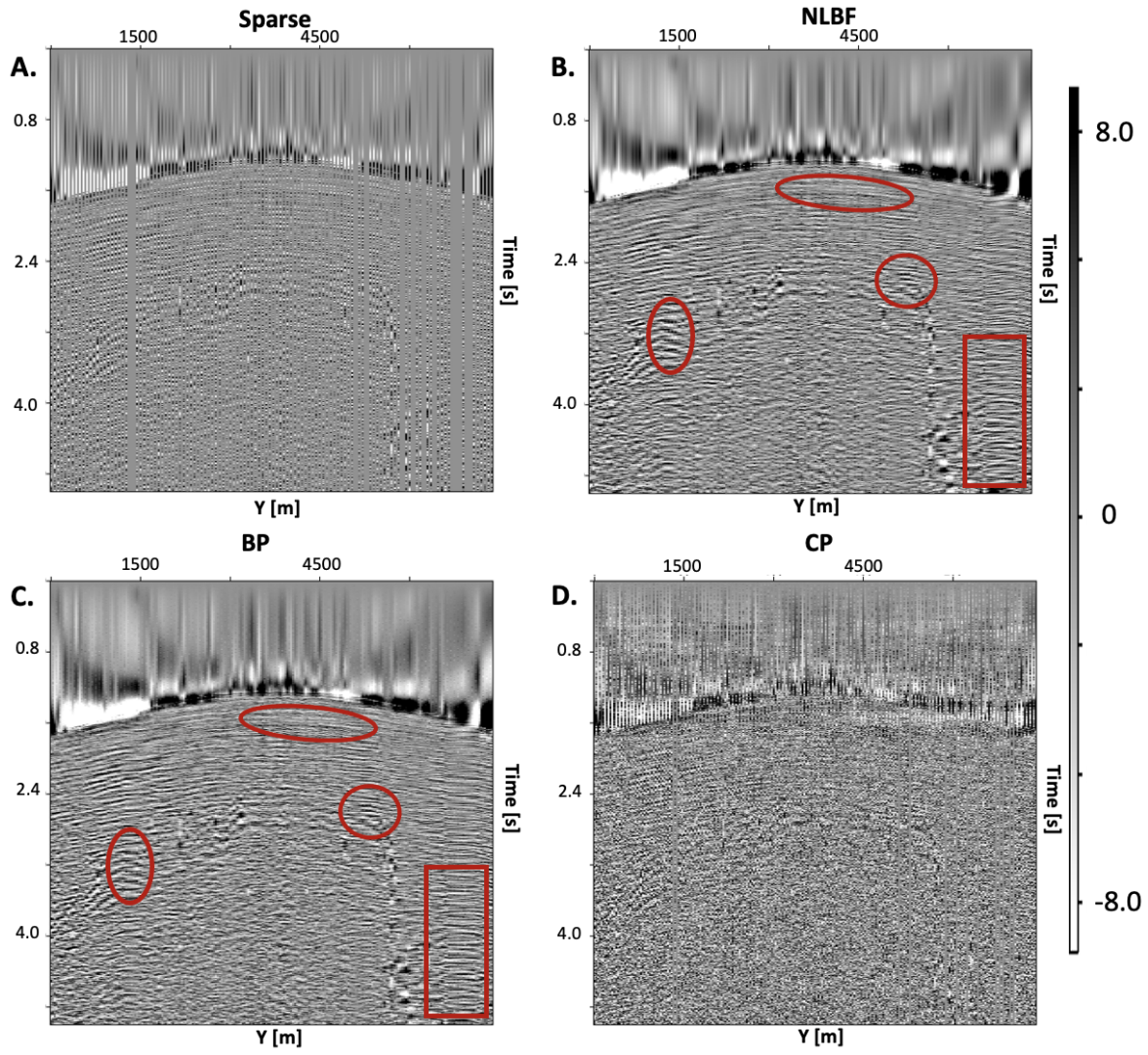

Figure 5-14: Field data reconstruction results of a crossline section at $x = 2190$ m. (A) The sparse section, (B) the NLBF reconstruction, (C) the BP reconstruction, and (D) the CP reconstruction. Red circles indicate some wavefield differences; red boxes highlight some differences in reconstruction in the gaps.

Finally, based on our experience from the synthetic SEAM Arid data to the field data, we expect the NLBF reconstruction to suppress the scattering noise. The reduced amount of overall scattering noise of the NLBF reconstruction compared to the BP reconstruction is indeed visible in Figures 5-12, 5-13 and 5-14. In comparison, the BP results present a grainier texture, which deteriorates reflection quality. Nevertheless, the BP reconstruction results are still impressive considering its algorithmic simplicity compared to the NLBF reconstruction method. Overall, for the cited reasons, on the datasets used, and with the selected parameters, our efforts to apply the NLBF framework for seismic data reconstruction are seen as a success.

Chapter 6

Discussion and conclusions

6-1 Discussion

Throughout this thesis, we have presented a novel seismic data reconstruction method using the NLBF framework. Results have shown high-quality reconstructions for both a 2/3 crossline sparse synthetic SEAM Arid dataset, and for a 3/4 sparse field dataset. However, in the limited scope of this thesis, only a small number of tests and conclusions could be accomplished, leaving many open questions and future research opportunities.

Technical and economic constraints of seismic data acquisition lead to sparse data acquisition geometries as discussed in Chapter 1. This sparsity is the driving force behind the development of seismic signal enhancement and reconstruction. In this thesis, we have successfully used the NLBF method to reconstruct single gathers up to 3/4 sparsity. Field data are typically composed of many gathers with levels of sparsity reaching up to 5/6 sparsity. Our NLBF algorithm is easily scalable to higher levels of data sparsity. We expect that higher levels of data sparsity would also yield impressive reconstruction results using our NLBF reconstruction method, at the expense of computational time. Using the NLBF reconstruction method on datasets with higher degrees of sparsity has yet to be tested and demonstrated.

In addition, the importance of the parameter selection for NLBF reconstruction has been reiterated several times. The bottom-line conclusion from our tests indicates that these parameters are data dependent, as the NLBF method depends on the input data geometry, the gaps in data coverage, the sparsity, and the factors highlighted in Appendix B. In this thesis, we manually selected these parameters for each dataset. Future developments on the NLBF reconstruction method could propose a self-adaptive algorithm that would automatically select a good set of parameters for each dataset. In addition, a better understanding of the local wavefield conditions in input datasets could lead to self-adaptive parameters within the same NLBF reconstruction. Parameters would be chosen to optimally reconstruct the wavefield at the localized level. Another interesting observation regarding the influence of parameters on the reconstruction quality is the potential NLBF operator parameter bias. In the scope of this thesis, the NLBF operators used for the reconstruction were pre-calculated

using an A_{op} of 600 m x 600 m and a time window size of $7\Delta t_s$. Results from our parameter tests indicated good SNR values for reconstructed gathers using these parameters as well. To test this hypothesis, further NLBF reconstruction experiments should be performed using different parameters to build the NLBF operators. In particular, we conjecture that a smaller aperture size ($200 - 300$ m) and denser parameter traces should yield higher quality reconstruction results if such parameters can fully cover gaps in the input dataset, but this has yet to be verified.

As a side product of this thesis, the BP methodology has also shown its potential to reconstruct regularly sparse seismic data. Preliminary tests, outside the scope of this thesis, have also revealed that some regularly sparse BP reconstructed images present higher SNR values than linear-bilinear reconstructed images. Such results were typically obtained using 20% or less randomly decimated pixels for each BP iteration (Chapter 4). We suggest that applying the BP methodology to the NLBF reconstructed time sections might improve the reconstruction quality, and could potentially provide a solution to the amplitude mismatch between the ground truth and NLBF reconstructed traces. Furthermore, BP reconstruction could also be applied to sparse wavefronts during the NLBF reconstruction, instead of the current linear-bilinear interpolation. This modification has the potential to yield even higher quality results, at the expense of computational efficiency. In the limited time scope of this thesis, we did not have the time to test this proposition, which offers another direction for further research.

To conclude this discussion, the NLBF method has shown promising results for sparse seismic data reconstruction and has left us with some interesting open questions for further research. We believe that it also has the potential to be applied to other types of data, even in different industries beyond the scope of geophysics.

6-2 Conclusions

Throughout this thesis, we presented the theory and methodology of NLBF reconstruction. Our research objective consisted of using the existing NLBF framework to reconstruct sparse seismic data along wavefronts. Additionally, we aimed to find the best NLBF control parameters to optimize reconstruction and assess the overall reconstruction quality compared to other control methods. We tested our method on a synthetic SEAM Arid dataset and a field dataset, and attempted to answer the following research question: can the NLBF framework be used for seismic data reconstruction, and how do these results compare to some conventional reconstruction methods? Our conclusions are the following:

- Results from the NLBF reconstruction show a successful reconstruction of the sparse synthetic and the field datasets. For the synthetic SEAM Arid data, the availability of the ground truth allowed us to confirm the NLBF reconstruction quality by direct comparisons with the target gather. The SNR and spectral analysis showed excellent preservation of the seismic signals, good continuity of reflections, and suppression of scattering noise. The field data also showed impressive reconstruction results, despite the superior complexity of the data. Similarly, the reconstructed gather exhibited an excellent wavefield continuity. Therefore, we successfully reconstructed sparse seismic data using the NLBF framework.

- Parameter tests performed on both datasets examined the effects of varying time window sizes, operator aperture sizes, and parameter trace interval distances on the reconstruction quality. Using the SNR from the SEAM Arid dataset and studying parameter interdependence, we found that control parameters are "dataset specific" because they ultimately depend on the gather's acquisition geometry. For our two datasets, the parameters that yielded the best NLBF reconstruction results are: a time window of $12\Delta t_s$, an operator aperture of 600 m x 600 m, and a parameter trace interval of $\Delta x = \Delta y = 60$ m.

- Comparison of the NLBF reconstructed gathers to the ground truth, and the reconstructed results from the CP and BP methods confirmed the effectiveness and quality of the NLBF reconstruction method. Both the synthetic SEAM Arid and field datasets showed excellent reconstruction results. For the SEAM Arid data, the NLBF reconstruction method showed high trace fidelity to the ground truth. Subsequent tests on the field data confirmed the high-quality reconstruction of the NLBF reconstruction method. Furthermore, the control tests also demonstrated the superiority of the NLBF reconstruction which showed enhanced wavefield continuity, better-reconstructed gaps, and wavefield amplitude variations closer to those of physical propagating waves.

- Overall, the use of the NLBF framework for seismic data reconstruction is a novel method. The scope of our research was adapted to the limited time available for the IDEA League Applied Geophysics Master's thesis. Although promising, more research is necessary to better understand the NLBF reconstruction method. We have already identified some interesting questions with the potential to further enhance reconstruction quality. Our research has shown the potential for the NLBF seismic data reconstruction method to be applied in industry settings, and beyond.

Appendix A

Additional NLBF results

In this Appendix, we present some additional NLBF reconstruction results from the synthetic SEAM Arid and field datasets to complement the ones presented in the main body of the thesis. For both datasets, we present a different inline, crossline and time section from those in Chapter 5. Appendix A-1 shows additional synthetic SEAM Arid results. We compare the NLBF reconstruction to the sparse section and the ground truth. Visual examination and SNR values highlight the quality of the NLBF reconstruction. Appendix A-2 shows additional field result sections. We compare the NLBF reconstruction to the BP and CP control methods. The respective conclusions reached for both datasets in Chapter 5 are also valid for the results presented in Appendix A.

A-1 SEAM Arid dataset

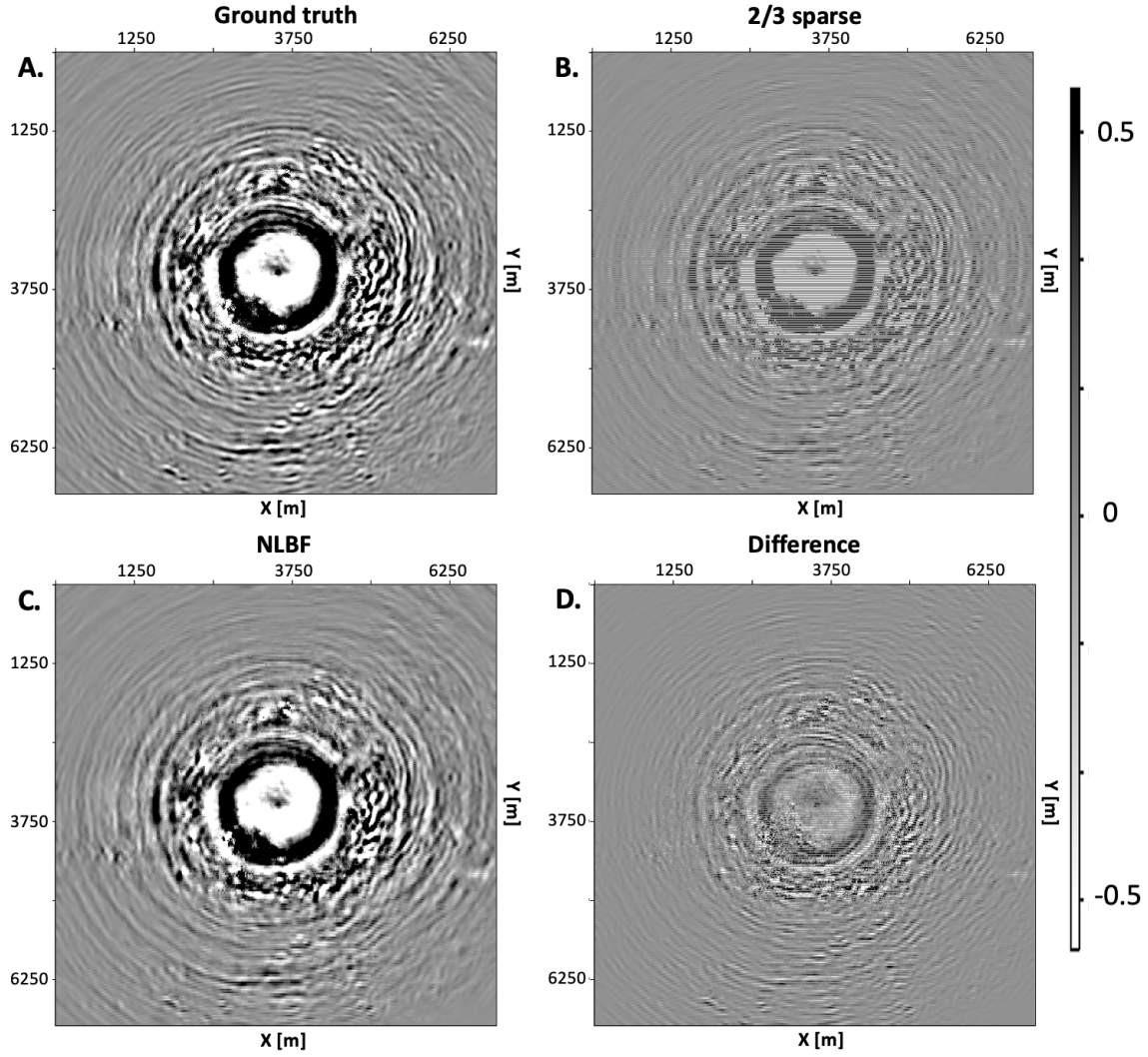

Figure A-1: NLBF reconstruction results for a time section at $t = 1.41$ s of the synthetic SEAM Arid gather. (A) The ground truth, (B) the 2/3 sparse section (SNR = 1.76), (C) the NLBF reconstructed section (SNR = 16.12), and (D) the difference between (A) and (C).

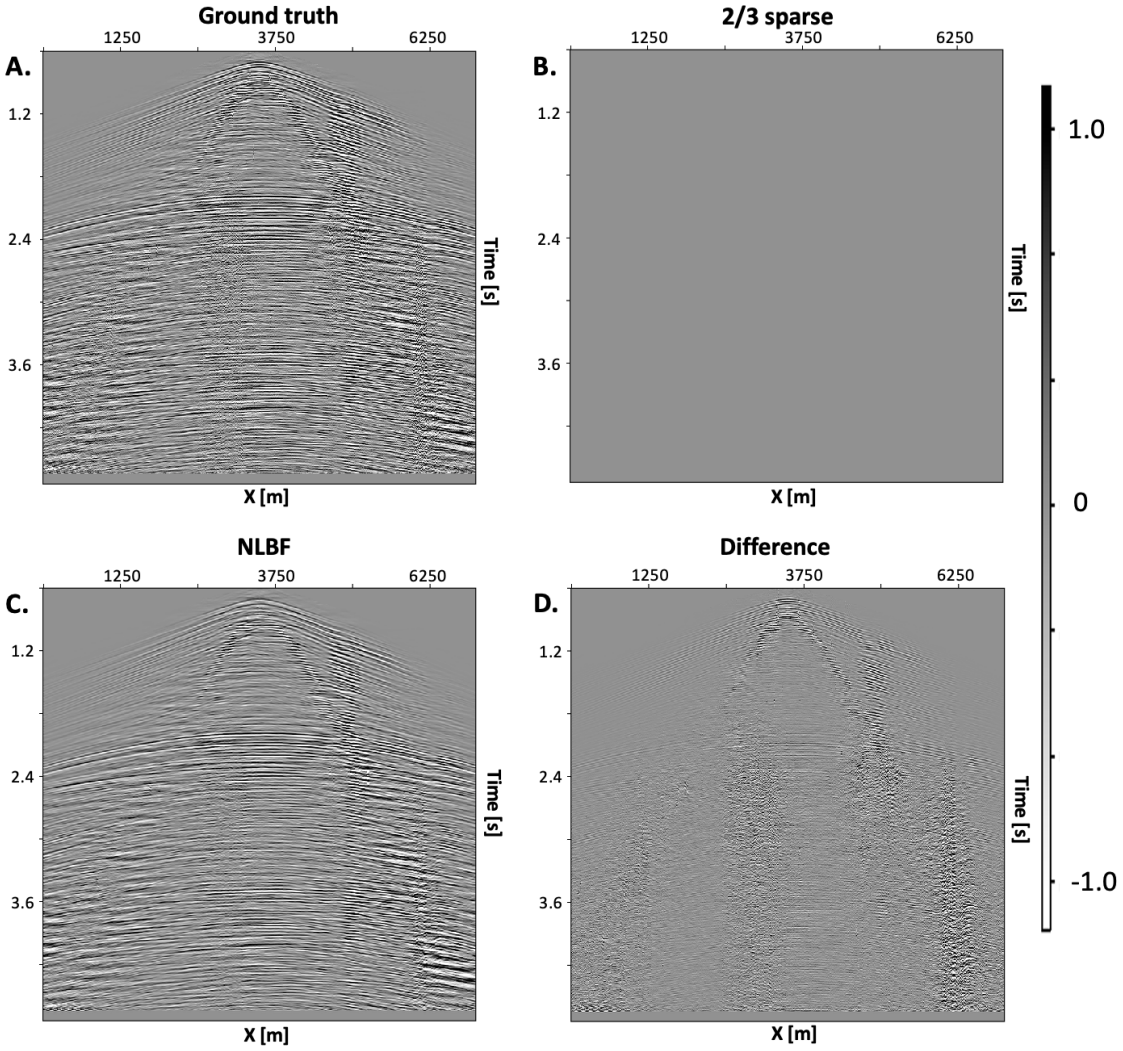

Figure A-2: NLBF reconstruction results for an inline section at $y = 6250$ m of the synthetic SEAM Arid gather. (A) The ground truth, (B) the missing sparse section (SNR = 0.0), (C) the NLBF reconstructed section (SNR = 4.95), and (D) the difference between (A) and (C).

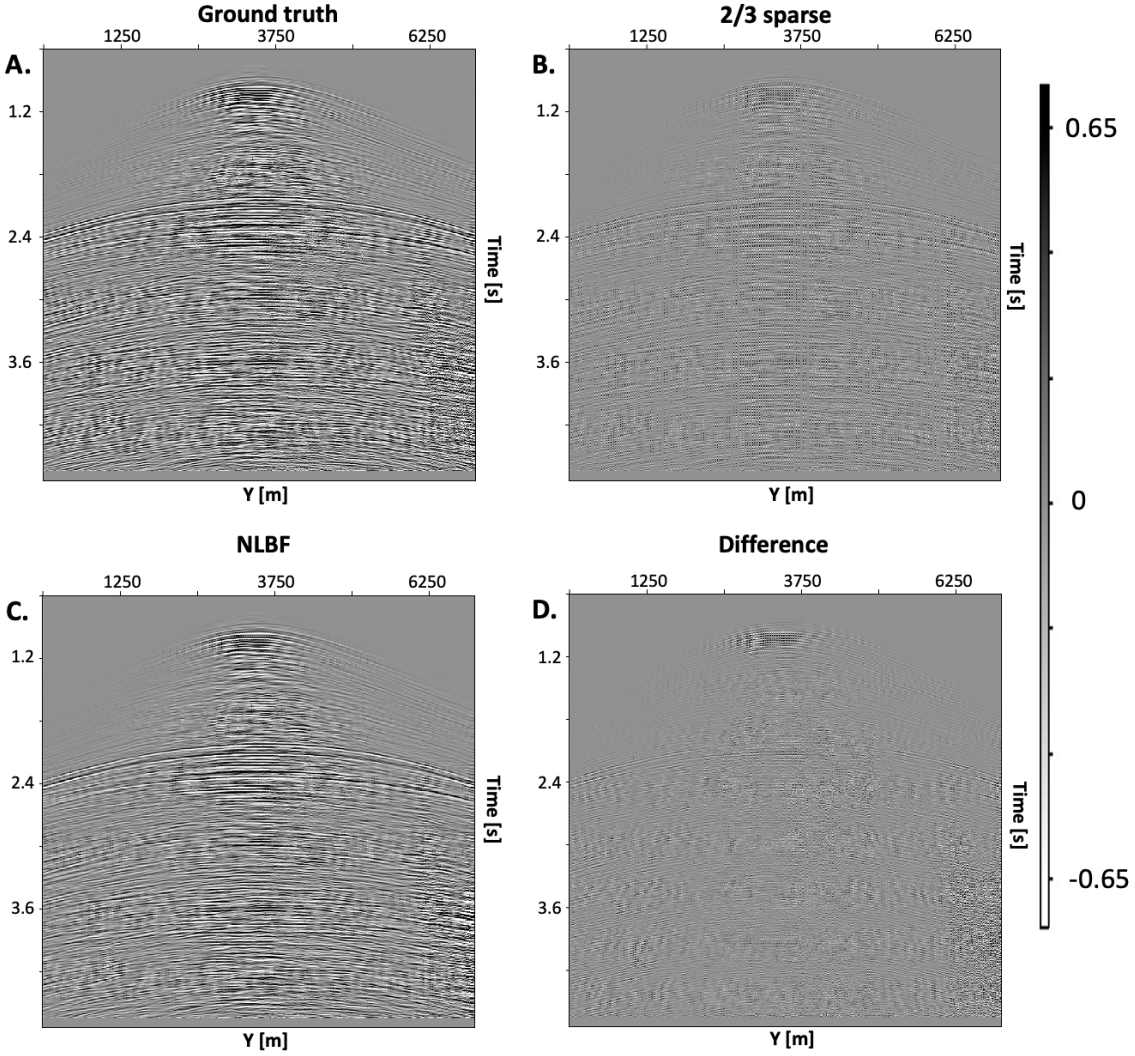

Figure A-3: NLBF reconstruction results for a crossline section at $x = 3750$ m of the synthetic SEAM Arid gather. (A) The ground truth, (B) the 2/3 sparse crossline section (SNR = 1.75), (C) the NLBF reconstructed section (SNR = 9.02), and (D) the difference between (A) and (C).

A-2 Field dataset

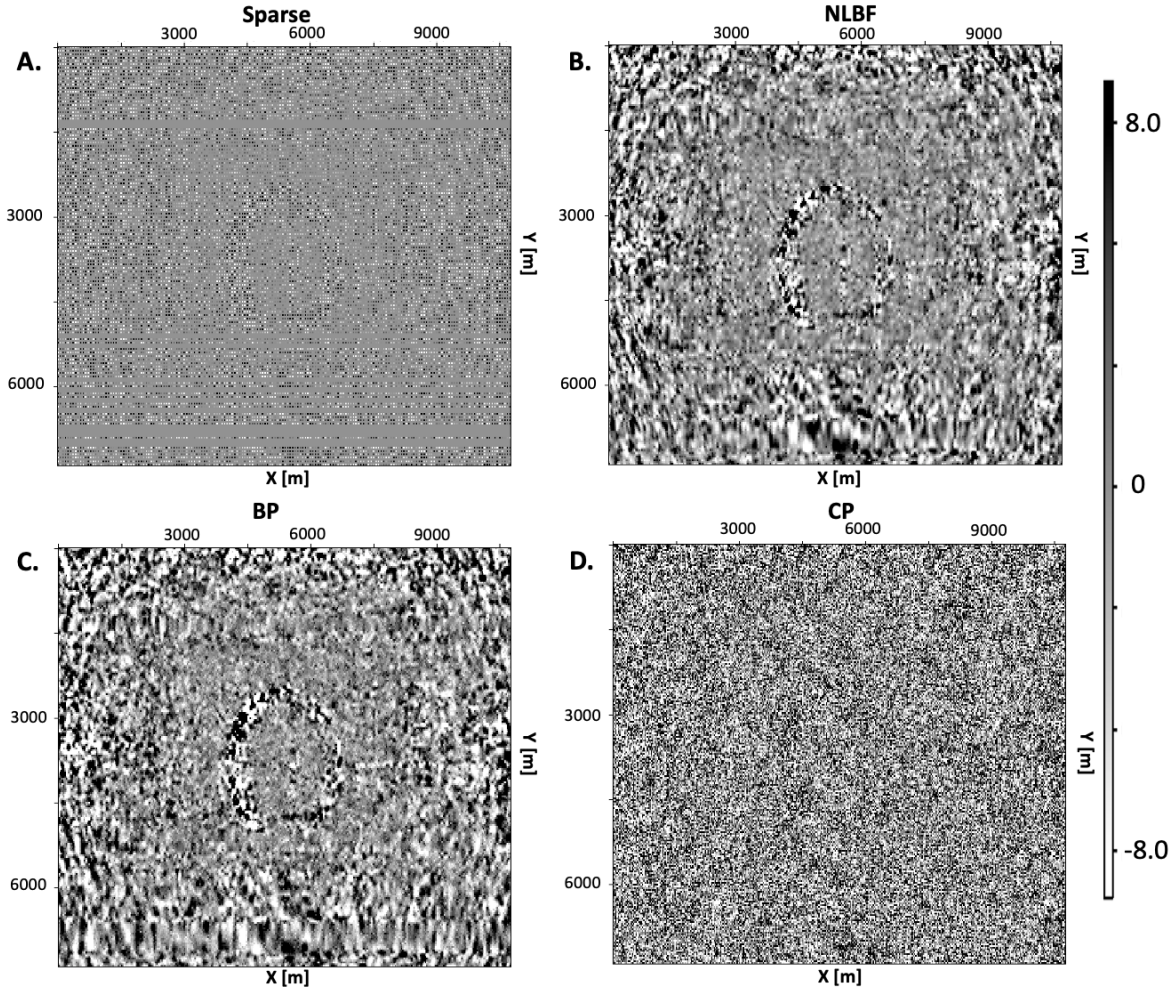

Figure A-4: Field data reconstruction results of a time section at $t = 3.8640$ s. (A) The sparse section, (B) the NLBF reconstructed section, (C) the BP reconstructed section, and (D) the CP reconstructed section.

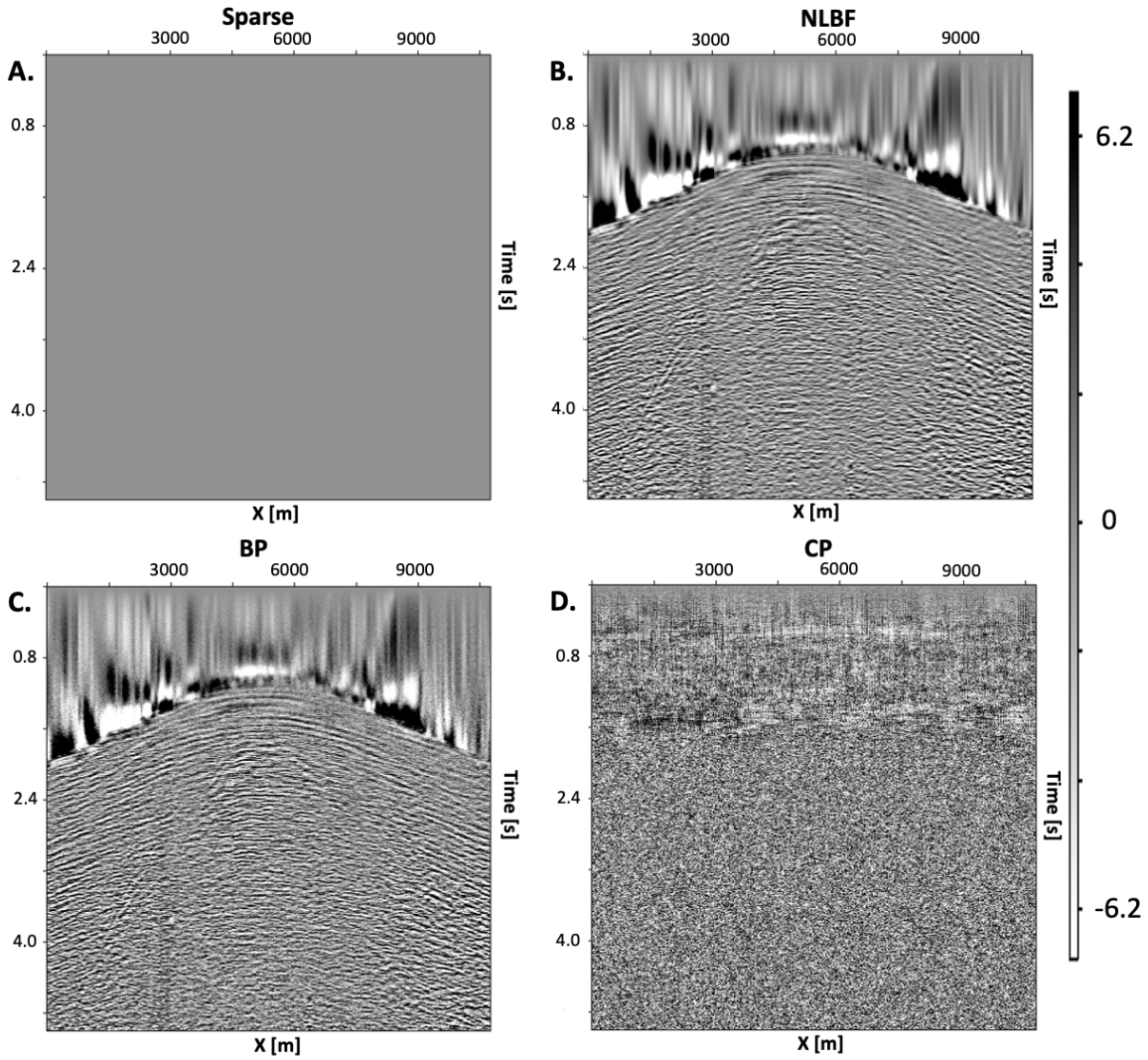

Figure A-5: Field data reconstruction results of an inline section at $y = 1200$ m. (A) The missing section, (B) the NLBF reconstructed section, (C) the BP reconstructed section, and (D) the CP reconstructed section.

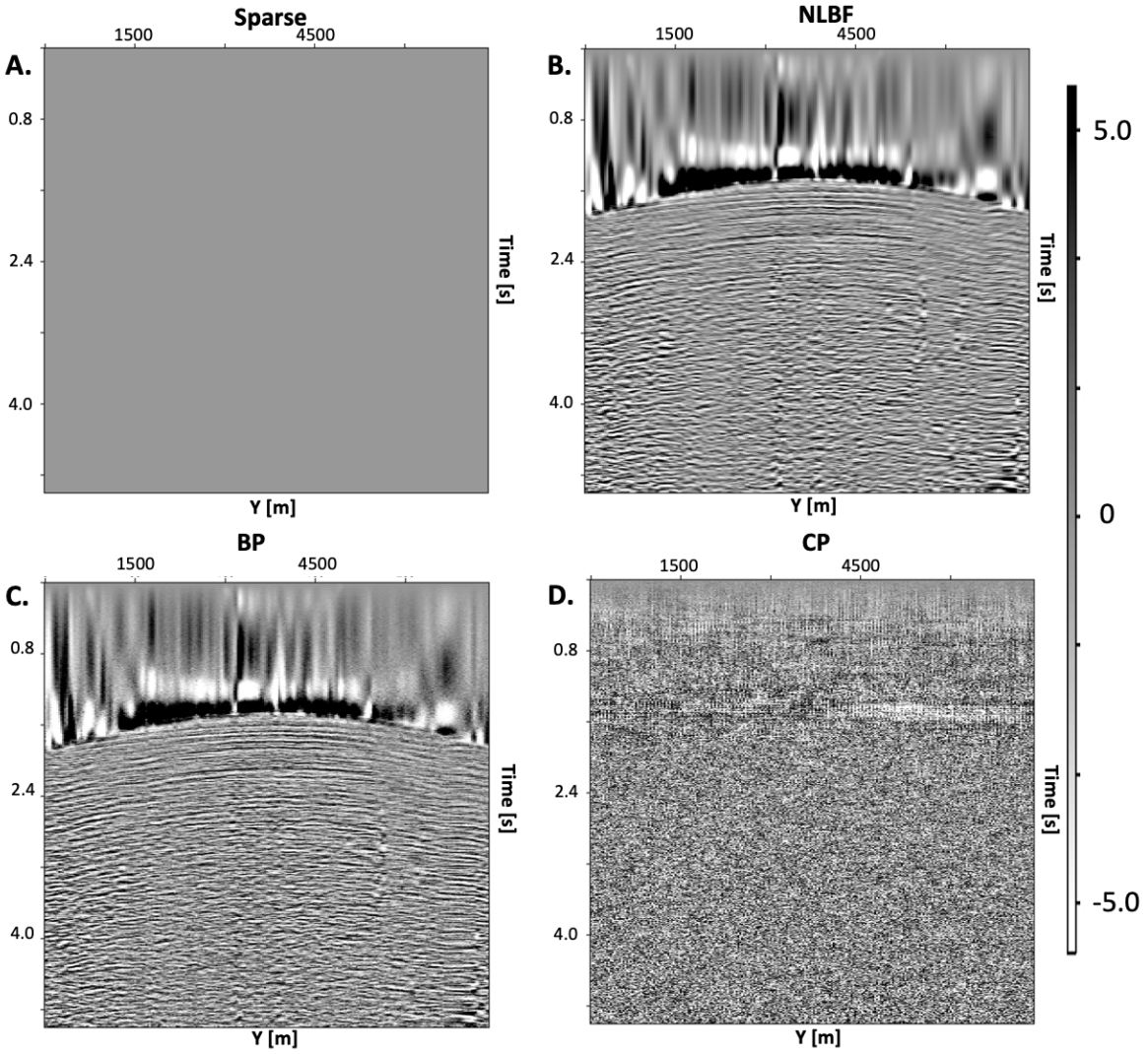

Figure A-6: Field data reconstruction results of a crossline section at $x = 9300$ m. (A) The missing section, (B) the NLBF reconstructed section, (C) the BP reconstructed section, and (D) the CP reconstructed section.

Appendix B

Control parameters for NLBF reconstruction

In this Appendix, we present and discuss in detail the effects and interplay among the NLBF control parameters, and we take a closer look at the results of the parameter tests presented in Table 5-1. The variable parameters of interest are the parameter trace spatial interval, the time window size, and the operator aperture. The detailed NLBF reconstruction theory is available in Chapter 4. Overall, our main and most important finding is that parameter selection is data-dependent, and it should be carefully tailored to each dataset in order to maximize reconstruction quality. In fact, spatial trace distribution (Δx and Δy), trace time discretization (Δt_s), degree of sparsity, gaps in the data, and parameters used to estimate the NLBF operators should all be taken into account. In the next section, we present how these parameters affect each other, and then we will discuss the results of our parameter tests.

B-1 Parameter selection considerations

The choice of the suitable parameters and combination of parameters are crucial, as some might negatively affect the reconstruction. First, the operator aperture must be carefully selected to fit the reconstructed dataset. In particular, our tests have shown that the operator aperture size must account for gaps in the dataset, and be large enough to cover the gaps from both sides. If this fails to happen, the interpolation along the sparse wavefronts during the NLBF reconstruction (Figure 4-7) will not be able to reconstruct target traces within the gap. The resulting reconstruction will show sharp boundaries at these locations and therefore will not achieve the maximum interpolation potential. Figure B-1 illustrates the operator setup for these two scenarios with parameter trace interval $\Delta x = \Delta y = 120$ m and time window size $9\Delta t_s$. Figure B-1A shows the setup for an operator aperture of 300 m x 300 m and the resulting reconstruction; Figure B-1B shows the setup for an operator aperture of 600 m x 600 m and the resulting reconstruction. It is clear that the reconstruction quality of the smaller aperture shows sharp discontinuities; these are highlighted by the red boxes.

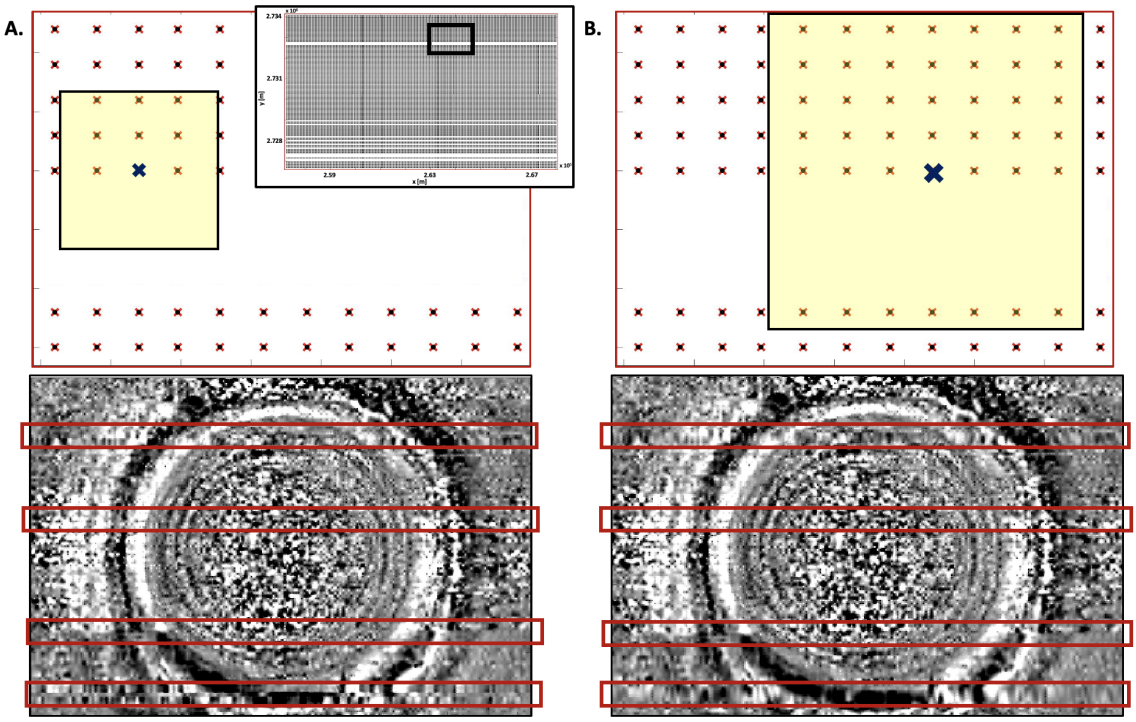

Figure B-1: Schematic illustration of the effects of the operator aperture size on gathers containing gaps, and associated NLBF reconstruction on the field dataset. The inset indicates the position of the close-up geometry. (A) An operator aperture of size 300 m x 300 m doesn't cover the gap. (B) An operator aperture size of 600 m x 600 m that covers the gap. A single aperture is shown for clarity and reconstruction differences are highlighted by the red boxes.

Another important criterion to consider is the interplay between the parameter trace interval and the operator aperture size. Specifically, the resulting "aperture crossover" between neighboring parameter traces is necessary for a complete reconstruction. Figure B-2 illustrates two scenarios, with and without crossover. Figure B-2A shows a parameter trace interval of $\Delta x = \Delta y = 180$ m and an operator aperture of 200 m x 200 m, resulting in no crossover. The associated data example shows the use of these parameters on the SEAM Arid dataset and the resulting gaps highlighted in the red boxes. Figure B-2B schematically shows the parameter trace interval of $\Delta x = \Delta y = 120$ m and an operator aperture of 400 m x 400 m, resulting in an overlap between each adjacent operator. The associated data example shows the absence of gaps in the reconstruction. Crossover is necessary in order to obtain a seamless transition between the reconstructed parameter trace sections. In the next section, we discuss the recommended amount of crossover based on our results.

The next observation found to affect the NLBF reconstruction results is the interplay between the time window size and the operator aperture. The curved nature of the NLBF operators causes them to temporally drift from their originating parameter trace time samples as illustrated in Figure B-3. This deviation is more pronounced at larger offsets, therefore it is common for large operator apertures. The issue encountered when operators deviate significantly is that certain target trace time samples might no longer be covered by these operators

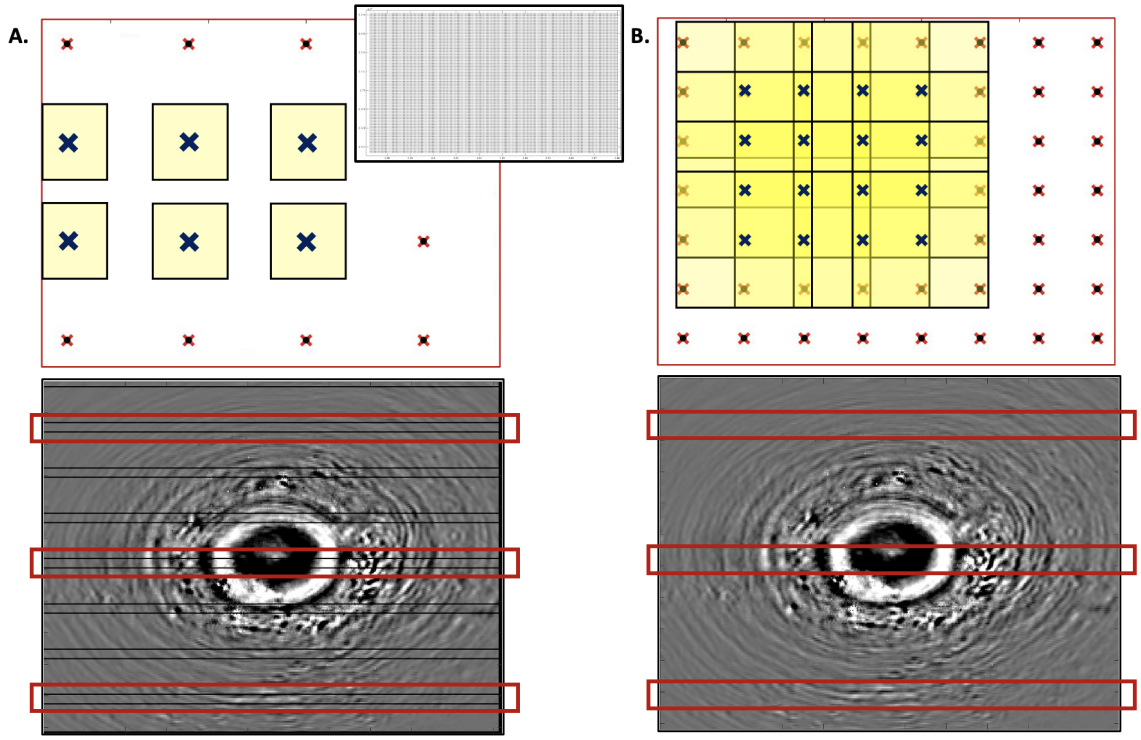

Figure B-2: Schematic illustration of the interplay between operator aperture size and parameter trace interval on the synthetic SEAM Arid dataset. The inset indicates the position of the close-up geometry. (A) A 200 m x 200 m aperture with 180 m parameter trace spacing results in no "crossover" and gaps in the reconstruction. (B) A 400 m x 400 m aperture with 120 m parameter trace spacing results in overlaps, and therefore no gaps exist in the reconstruction. Only a few apertures are drawn for clarity. Some reconstruction differences are highlighted by the red boxes.

in the far field, as illustrated in Figure B-3. Therefore, for larger operator aperture sizes, we need larger time windows in order to reconstruct all target trace time samples. However, a time window that is too large can also affect the reconstruction quantity. In the next section, we discuss the recommended time window and operator aperture dimensions for the SEAM Arid dataset and discuss the results of our parameter tests.

B-2 Parameter tests results

Based on our tests and the limited time scope of this thesis we cannot draw any definite conclusive rules of thumb regarding dataset-dependent parameter selection. But based on the reconstruction results obtained from the tested SEAM Arid gather and the associated SNR results, we are able to give some recommendations.

Figure B-4 displays the SNR results from the grid search style parameter tests on the synthetic SEAM Arid dataset. The parameters tested were: time window $\in [7\Delta t_s,\ 12\Delta t_s,\ 15\Delta t_s,\ 18\Delta t_s]$, $A_{op} \in$ [200 m x 200 m, 300 m x 300 m, 400 m x 400 m, 600 m x 600 m], and parameter trace interval $\Delta x = \Delta y \in [60m, 120m, 180m]$. The SNR information is divided in

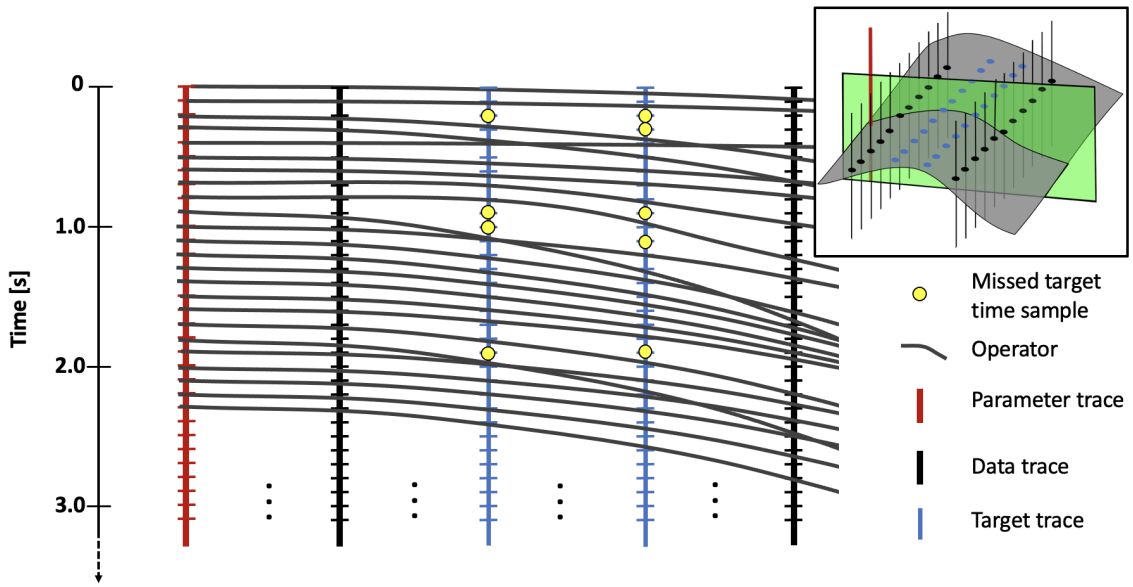

Figure B-3: Schematic illustration of the interplay between operator aperture size and the time window size on a synthetic SEAM Arid crossline slice. At larger offsets from the originating parameter trace, the curved operators are more likely to miss target time samples. These missed samples are highlighted in yellow.

the following way: red dashed lines delimit separate time window sizes in increasing order; within each time section, four colored subdivisions indicate different aperture sizes according to the legend; and within these aperture sizes, different shaped markers indicate the parameter trace intervals. Studying this plot reveals some interesting repeating patterns.

First, we note that for all time window sizes, the SNR of the 200 m x 200 m operator aperture and 180 m parameter trace interval is missing. This is caused by an insufficient aperture crossover distance, leading to gaps in the reconstruction (Figure B-2). We now turn our attention to individual parameters.

- Operator aperture

Figure B-4 shows that within the different time window sizes, the 600 m x 600 m A_{op} results often show higher SNR values. In fact, the SNR seems to consistently increase with increasing operator aperture size. This is true for a time window size of $12\Delta t_s$, $15\Delta t_s$, and $18\Delta t_s$; but not for $7\Delta t_s$. In this last case, we suspect the time window size to be too small for the large 600 m x 600 m aperture, as explained and illustrated in Figure B-3. Therefore, for the synthetic SEAM Arid dataset, an A_{op} of 600 m x 600 m seems to yield the best NLBF reconstruction results.

While we cannot draw any definite conclusions regarding the reasons behind this, we suspect that the parameters used to create the NLBF operators [Sun et al., 2022] influence the optimal aperture size for the NLBF reconstruction. In fact, an aperture of 600 m x 600 m was originally used to estimate the NLBF operators from the data, and we believe the NLBF reconstruction

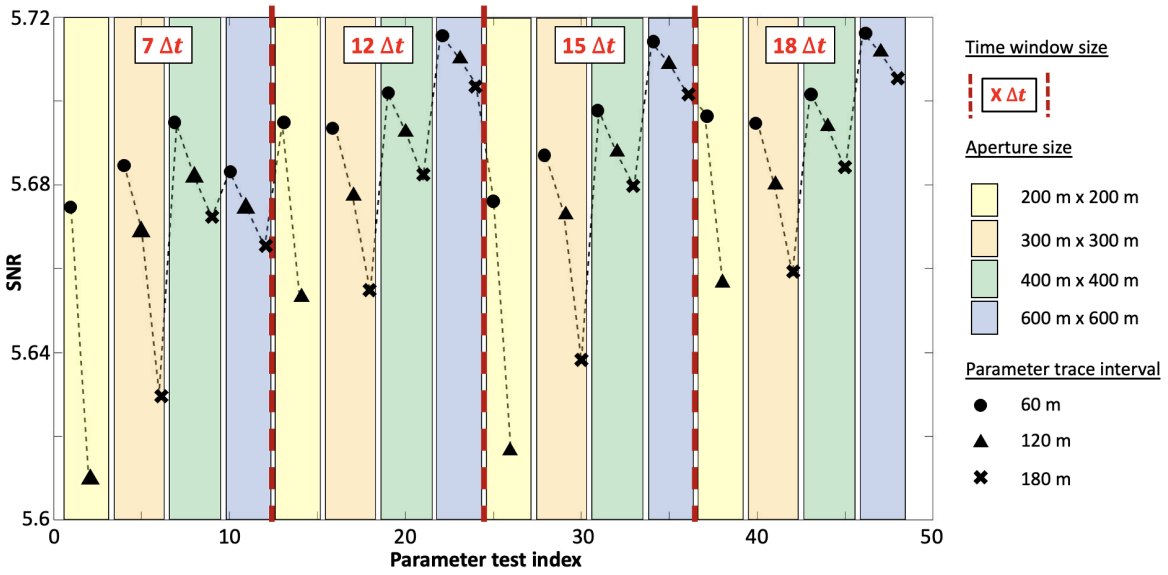

Figure B-4: SNR values from the synthetic SEAM Arid gather NLBF reconstruction parameter tests according to table 5-1.

could be intrinsically influenced by this operator estimation bias. Therefore, we recommend matching the A_{op} for the NLBF reconstruction to the aperture used to estimate the NLBF operators. Additional research regarding the influence of the NLBF operator aperture should be done to draw any further conclusions.

Figure B-5 compares the results from different A_{op} sizes for a time section at $t = 0.9$ s and an inline section at $y = 6250$ m, using a time window of $15\Delta t_s$ and a parameter trace interval of 120 m. Some differences in the reconstructed seismic gathers are highlighted in red.

- Time window

Figure B-4 highlights some interesting patterns regarding the influence of the time window size on the SNR value. The SNR values of gathers reconstructed using $7\Delta t_s$ are consistently smaller than the SNR values of larger time windows using the same A_{op} and parameter trace intervals. Again, as explained in Figure B-3, we theorize that the time window is too small compared to the 600 m x 600 m A_{op} used for the NLBF operator estimation. Larger time windows are able to cover more target trace time samples and therefore present higher SNR values.

Interestingly, the SNR values for the other time window sizes are fairly equivalent across all A_{op} and all parameter trace intervals. Overall, a time window size of $12\Delta t_s$ showed very slight improvements in SNR compared to $15\Delta t_s$ and $18\Delta t_s$. Figure B-6 compares the results from different time window sizes for a time section at $t = 1.41$ s and an inline sections at $y = 3587.5$ m, using an $A_{op} = 200$ m x 200 m, and a parameter trace interval of 120 m. Differences in the reconstructed seismic gathers are highlighted in the red ovals. These results show improvements for larger time window sizes. However, visible improvements between the $12\Delta t_s$, $15\Delta t_s$, and $18\Delta t_s$ windows are hard to see, although there are slight variations. Overall, for the SEAM Arid dataset, we recommend a time window size of $12\Delta t_s$.

- Parameter trace interval

Figure B-4 shows that across all time window sizes and A_{op} sizes, the parameter trace interval related SNR results follow the same pattern. In fact, we see that the SNR decreases for increasing parameter trace interval, and the best SNR is always reached at $\Delta x = \Delta y = 60$ m. Figure B-7 compares the results from different parameter trace interval distances for a time section at $t = 1.5$ s and an inline section at $y = 3587.5$ m, using a time window size of $7\Delta t_s$, and an $A_{op} = 300$ m x 300 m. Some differences in reconstruction quality between the gathers are highlighted by the red ovals. Interestingly, Figure B-4 shows that for $A_{op} = 600$ m x 600 m, the variation in SNR values for different parameter trace intervals is even smaller.

For the SEAM Arid dataset reconstruction results, we have chosen $\Delta x = \Delta y = 60$ m, as the results presented the highest SNR value. However, the resulting slight improvement in SNR also leads to increased computational time because of the increased amount of parameter traces. Therefore there is a tradeoff between reconstruction quality and computational time. In the context of this thesis, we use 60 m as it yields the highest SNR, but in other industry settings, a faster computational time might be more valuable than a slight increase in SNR.

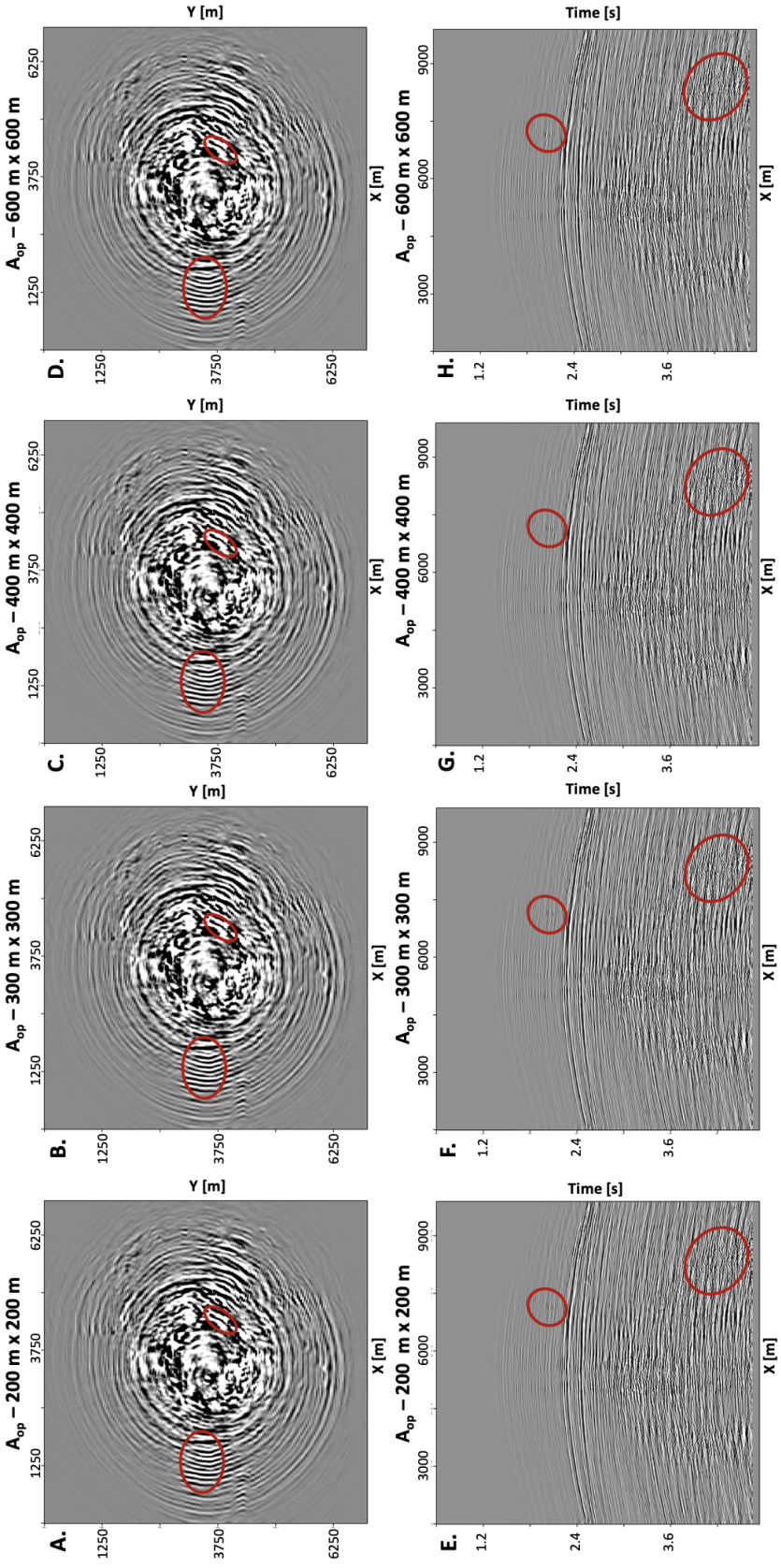

Figure B-5: Results from the operator aperture parameter tests for $A_{op} = 200$ m x 200 m, 300 m x 300 m, 400 m x 400 m, 600 m x 600 m, using a time window of $15\Delta t_s$ and a parameter trace interval of 120 m. (A-D) A time section at $t = 0.9$ s. (E-H) An inline section at $y = 6250$ m

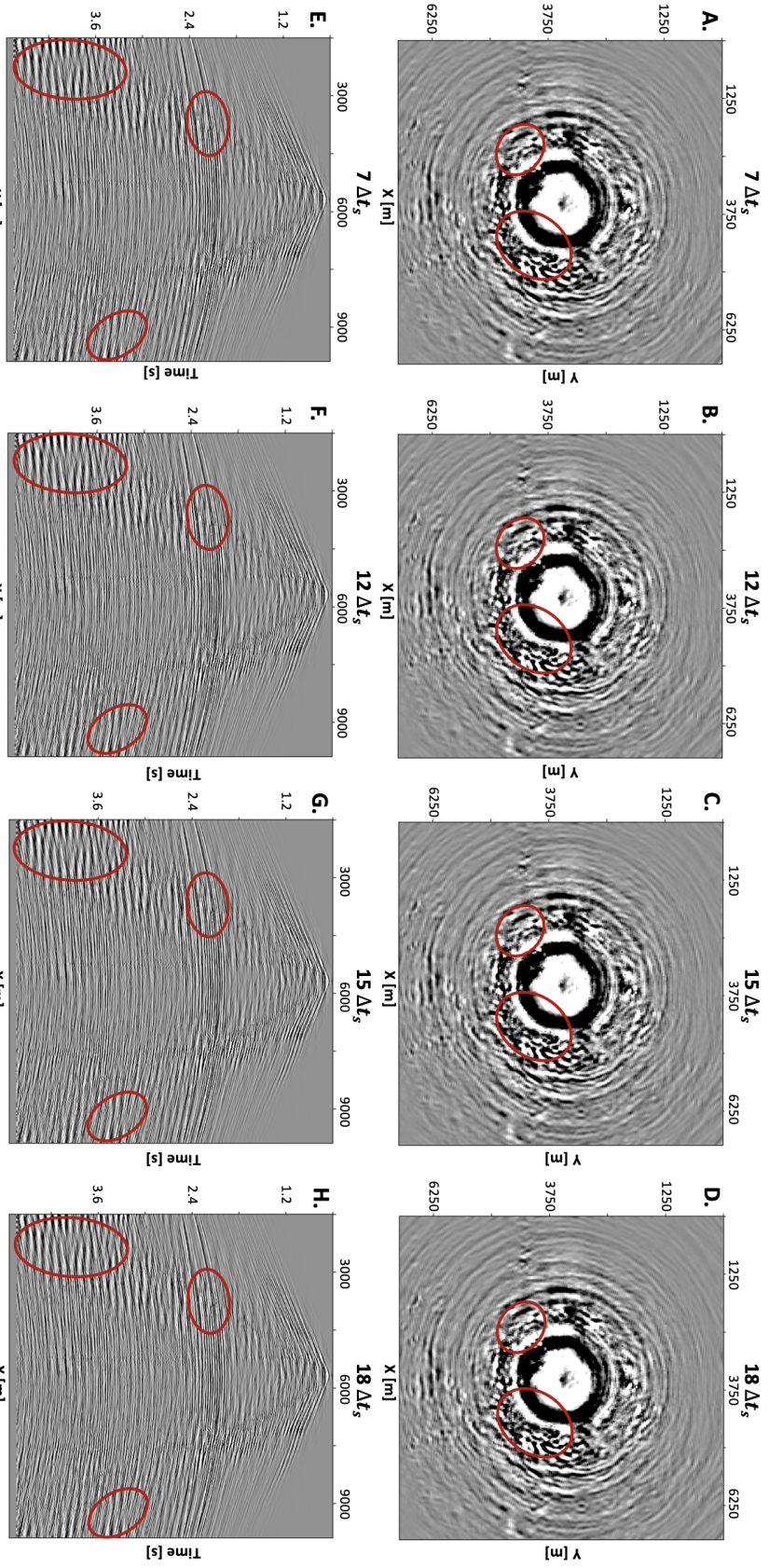

Figure B-6: Results from the time window parameter tests for $7\Delta t_s, 12\Delta t_s, 15\Delta t_s, 18\Delta t_s$, using an $A_{op} = 200$ m x 200 m and a parameter trace interval of 120 m. (A-D) A time section at $t = 1.41$ s. (E-H) An inline section at $y = 3587.5$ m

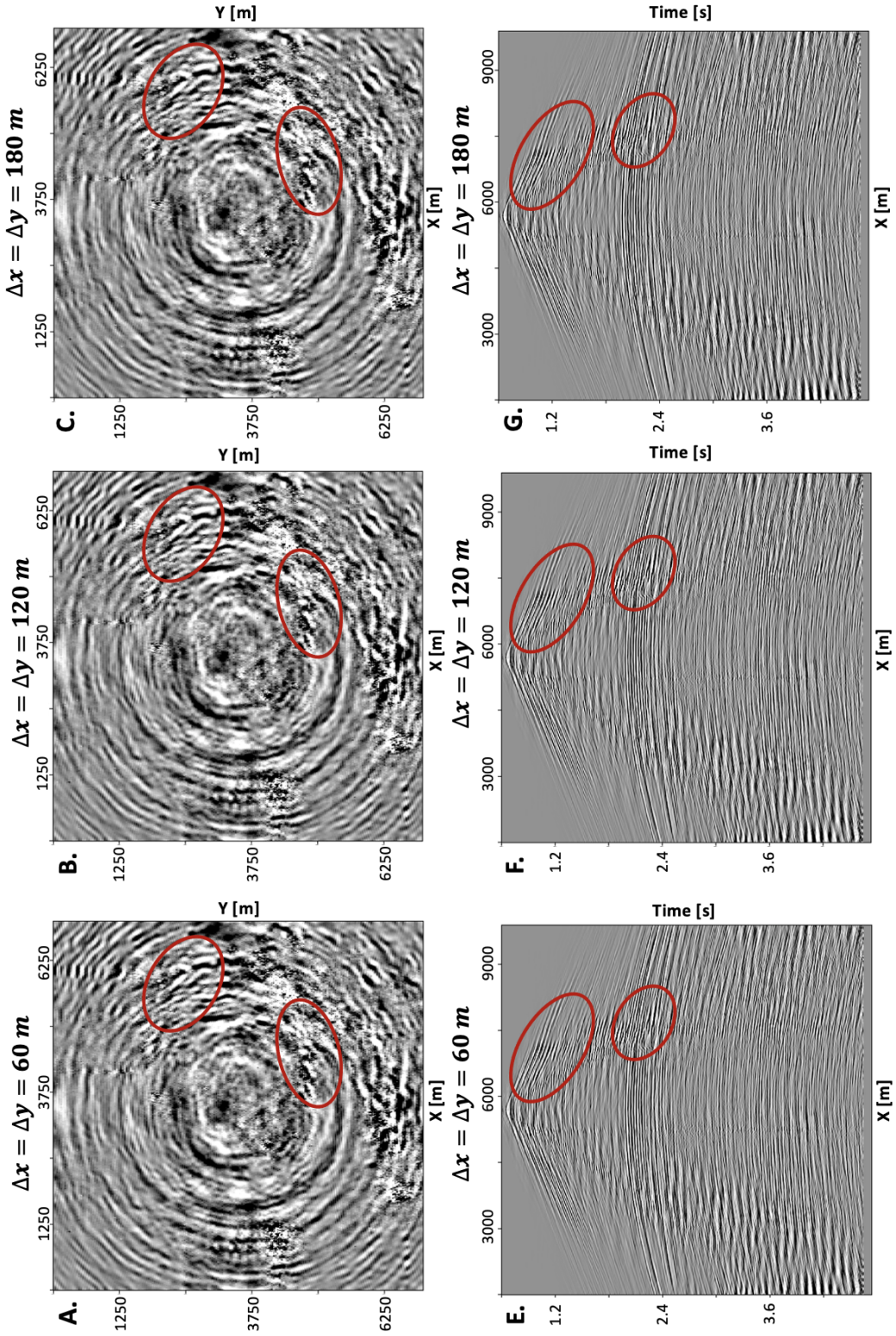

Figure B-7: Results from the parameter trace interval parameter tests for $\Delta x = \Delta y = 60, 120, 180$ m, using an $A_{op} = 300$ m x 300 m and a $7\Delta t_s$ window. (A-D) A time section at $t = 1.5$ s. (E-H) An inline section at $y = 3587.5$ m

Acknowledgements

First of all, I want to thank the entirety of the Aramco Overseas Company Delft office who have welcomed me into their workspace and treated me like one of their own. They have helped me throughout this thesis, we've shared memorable moments and had some engaging helpful discussions.

In particular, my advisor Dr. Yimin Sun has supported me and led me to accomplish my thesis. He offered much of his time through guidance, constructive comments, and immense support. Also, Center Director Roald van Borselen for his kindness and outstanding leadership. I am very grateful to have had them by my side during this process.

Additionally, I would like to thank the entire IDEA League staff including professors, the program board, and teacher's assistants. Without them, the past two years wouldn't have been possible. In particular, my academic advisor Dr. Guy Drijkoningen for his support, and constructive feedback on my thesis.

A special thanks goes to my peers in the IDEA League Applied Geophysics masters program, who have made the past two years incredibly memorable. We have gone on countless adventures and shared some great laughs that I will always remember. Lastly, I cannot thank my parents and family enough. Without them, I would not be at this point today. Thank you!

Delft University of Technology Quentin Hirsch

August 5, 2022

Bibliography

Abma, R. and Kabir, N. (2006). 3D interpolation of irregular data with a pocs algorithm. *Geophysics*, 71(6):E91–E97.

Acuna, Y. P. V. and Sun, Y. (2020). An efficiency-improved genetic algorithm and its application on multimodal functions and a 2D common reflection surface stacking problem. *Geophysical Prospecting*, 68(4):1189–1210.

Bagaini, C. and Spagnolini, U. (1996). 2-D continuation operators and their applications. *Geophysics*, 61(6):1846–1858.

Bakulin, A., Silvestrov, I., Dmitriev, M., Neklyudov, D., Protasov, M., Gadylshin, K., and Dolgov, V. (2020). Nonlinear beamforming for enhancement of 3D prestack land seismic data. *Geophysics*, 85(3):V283–V296.

Buzlukov, V. and Landa, E. (2013). Imaging improvement by prestack signal enhancement. *Geophysical Prospecting*, 61(6-Challenges of Seismic Imaging and Inversion Devoted to Goldin):1150–1158.

Claerbout, J. F. (1985). *Imaging the earth's interior*, volume 1. Blackwell scientific publications Oxford.

Claerbout, J. F. and Abma, R. (1992). *Earth soundings analysis: Processing versus inversion*, volume 6. Blackwell Scientific Publications London.

Deregowski, S. (1986). What is DMO? *First break*, 4(7).

Engl, H. W., Hanke, M., and Neubauer, A. (1996). *Regularization of inverse problems*, volume 375. Springer Science & Business Media.

Fomel, S. and Kazinnik, R. (2013). Non-hyperbolic common reflection surface. *Geophysical Prospecting*, 61(1):21–27.

Gao, J.-J., Chen, X.-H., Li, J.-Y., Liu, G.-C., and Ma, J. (2010). Irregular seismic data reconstruction based on exponential threshold model of POCS method. *Applied Geophysics*, 7(3):229–238.

Gelchinsky, B., Berkovitch, A., and Keydar, S. (1999). Non-hyperbolic common reflection surface. *Applied geophysics*, 42(3-4):229–242.

Geldart, L. P. and Sheriff, R. E. (2004). *Problems in exploration seismology and their solutions*. Society of Exploration Geophysicists.

Hirsch, Q. B. (2022). Image reconstruction optimization: Parameter testing for an alternating projections onto convex sets (APOCS) algorithm: Research module project, rwth aachen.

Hoecht, G., Landa, E., and Baina, R. (2004). Seimic data interpolation using local wavefield characteristics. *In 66th EAGE meeting, Paris, France,*, 31(Extended Abstracts D).

Hoecht, G., Ricarte, P., Bergler, S., and Landa, E. (2009). Operator-oriented CRS interpolation. *Geophysical Prospecting*, 57(6):957–979.

Kabir, M. N. and Verschuur, D. (1995). Restoration of missing offsets by parabolic Radon transform 1. *Geophysical Prospecting*, 43(3):347–368.

Knutsson, H. and Westin, C.-F. (1993). Normalized and differential convolution. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 515–523. IEEE.

Kuehl, H. and Sacchi, M. (2002). Robust AVP estimation using least-squares wave-equation migration. In *SEG technical program expanded abstracts 2002*, pages 281–284. Society of Exploration Geophysicists.

Kurucz, M., Benczúr, A. A., and Csalogány, K. (2007). Methods for large scale SVD with missing values. In *Proceedings of KDD cup and workshop*, volume 12, pages 31–38. Citeseer.

Kutscha, H. (2014). The double focal transformation and its application to data reconstruction: Phd thesis, delft university of technology.

Liu, B. and Sacchi, M. D. (2004). Minimum weighted norm interpolation of seismic records. *Geophysics*, 69(6):1560–1568.

Liu, Y. and Fomel, S. (2011). Seismic data interpolation beyond aliasing using regularized nonstationary autoregression. *Geophysics*, 76(5):V69–V77.

Malcolm, A., Shabelansky, A. H., Fehler, M., and Rodi, W. (2014). Migration-based seismic trace interpolation of sparse converted phase micro-seismic data. *SEG Technical Program Expanded Abstracts 2014*, 2014:3642–3647.

Marks II, R. J. (1997). Alternating projections onto convex sets. *Deconvolution of Images and Spectra*, pages 476–501.

Naghizadeh, M. and Sacchi, M. D. (2009). F-X adaptive seismic-trace interpolation. *Geophysics*, 74(1):V9–V16.

Naghizadeh, M. and Sacchi, M. D. (2010). Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data. *Geophysics*, 75(6):WB189–WB202.

Nemeth, T., Wu, C., and Schuster, G. T. (1999). Least-squares migration of incomplete reflection data. *Geophysics*, 64(1):208–221.

Olsson, C. and Oskarsson, M. (2009). A convex approach to low rank matrix approximation with missing data. In *Scandinavian Conference on Image Analysis*, pages 301–309. Springer.

Oristaglio, M. (2015). SEAM update: The Arid model—seismic exploration in desert terrains. *The Leading Edge*, 34(4):466–468.

Pecholcs, P. I., Al-Saad, R., Al-Sannaa, M., Quigley, J., Bagaini, C., Zarkhidze, A., May, R., Guellili, M., Sinanaj, S., and Membrouk, M. (2012). A broadband full azimuth land seismic case study from Saudi Arabia using a 100,000 channel recording system at 6 terabytes per day: Acquisition and processing lessons learned. In *SEG Technical Program Expanded Abstracts 2012*, pages 1–5. Society of Exploration Geophysicists.

Qin, F., Leger, P. C., Ren, J., Aleksic, V., Rowe, R. W., Zainaldin, S., and Hadab, S. A. (2018). A robust implementation and application of antileakage fourier transform interpolation. *The Leading Edge*, 37(7):538–543.

Regone, C., Fry, M., and Etgen, J. (2015). Dense sources vs. dense receivers in the presence of coherent noise: a land modeling study. In *2015 SEG Annual Meeting*. OnePetro.

Ronen, J. (1987). Wave-equation trace interpolation. *Geophysics*, 52(7):973–984.

Sambridge, M. and Drijkoningen, G. (1992). Genetic algorithms in seismic waveform inversion. *Geophysical Journal International*, 109(2):323–342.

Schwarz, B., Vanelle, C., Gajewski, D., and Kashtan, B. (2014). Curvatures and inhomogeneities: An improved common-reflection-surface approach. *Geophysics*, 79(5):S231–S240.

Spagnolini, U. and Opreni, S. (1996). 3-d shot continuation operator. In *SEG Technical Program Expanded Abstracts 1996*, pages 439–442. Society of Exploration Geophysicists.

Spitz, S. (1991). Seismic trace interpolation in the FX domain. *Geophysics*, 56(6):785–794.

Sun, Y., Silvestrov, I., and Bakulin, A. (2022). Enhancing 3D land seismic data using nonlinear beamforming based on the efficiency-improved genetic algorithm. *IEEE Transactions on Evolutionary Computation*.

Sun, Y. and Verschuur, D. J. (2014). A self-adjustable input genetic algorithm for the near-surface problem in geophysics. *IEEE Transactions on Evolutionary Computation*, 18(3):309–325.

Trickett, S. (2008). F-xy Cadzow noise suppression. In *SEG Technical Program Expanded Abstracts 2008*, pages 2586–2590. Society of Exploration Geophysicists.

Trickett, S. and Burroughs, L. (2009). Prestack rank-reduction-based noise suppression. *CSEG Recorder*, 34(9):24–31.

Trickett, S., Burroughs, L., Milton, A., Walton, L., and Dack, R. (2010). Rank-reduction-based trace interpolation. In *SEG Technical Program Expanded Abstracts 2010*, pages 3829–3833. Society of Exploration Geophysicists.

Wang, B. (2016). An efficient POCS interpolation method in the frequency-space domain. *IEEE Geoscience and Remote Sensing Letters*, 13(9):1384–1387.

Wapenaar, C. P. A. (2014). *Elastic wave field extrapolation: Redatuming of single-and multi-component seismic data*. Elsevier.

Xie, Y. and Gajewski, D. (2017). 5-D interpolation with wave-front attributes. *Geophysical Journal International*, 211(2):897–919.

Yilmaz, Ö. (2001). *Seismic data analysis: Processing, inversion, and interpretation of seismic data*. Society of exploration geophysicists.

Zhang, L., Li, A., Yang, J., Li, S., Yao, Y., Xiao, F., and Huang, Y. (2021). Five-dimensional interpolation in the OVT domain for wide-azimuth seismic data. *Journal of Geophysics and Engineering*, 18(4):529–538.

Zhang, Y.-Q. and Lu, W.-K. (2014). 2D and 3D prestack seismic data regularization using an accelerated sparse time-invariant Radon transform. *Geophysics*, 79(5):V165–V177.

Zwartjes, P. and Gisolf, A. (2007). Fourier reconstruction with sparse inversion. *Geophysical prospecting*, 55(2):199–221.

Zwartjes, P. and Sacchi, M. (2007). Fourier reconstruction of nonuniformly sampled, aliased seismic data. *Geophysics*, 72(1):V21–V32.