UNIVERSITY OF MISKOLC

Faculty of Earth Science and Engineering

Department of Geophysics

INVERSION-BASED FOURIER TRANSFORMATION ALGORITHM USED IN
PROCESSING GEOPHYSICAL DATA

Ph.D. THESIS

By

OMAR AL MARASHLY

Scientific supervisor

Prof. Dr. Mihály Dobróka

MIKOVINY SÁMUEL DOCTORAL SCHOOL OF EARTH SCIENCES

Head of the Doctoral School: Prof. Dr. habil. Péter Szűcs

Miskolc 2023

HUNGARY

ADVISOR'S FOREWORD

for the (PhD) thesis

„INVERSION-BASED FOURIER TRANSFORMATION ALGORITHM USED IN
PROCESSING GEOPHYSICAL DATA"

by

OMAR AL MARASHLY

The topic of the Candidate's thesis - inversion-based geophysical data processing - is in the focus of international research. The new method developments introduced by the Candidate in the thesis belong to the range of modern data processing tools of applied geophysics.

Following the main objectives of the thesis, the Candidate introduced new series expansion-based Fourier- and Hilbert Transformation methods and investigated them on synthetic and in-situ measured datasets in two distinct categories:

-   noise reduction capability and outlier sensitivity of the newly developed inversion-based Fourier transformation using Chebishev polynomials of the first- and second-kind as basis functions.
-   improving the noise reduction capability and robustness of Hilbert Transform by applying the inversion-based Fourier Transformation method using Hermite functions and Chebishev polynomials as basis functions.

The Candidate completed his PhD doctoral research with a high degree of independence and precision and reached scientifically valuable new results. His continuous efforts towards scientific research, his creativity, and the results presented in this thesis prove the scientific knowledge and the suitability of the Candidate for independent research. In my opinion, the Candidate's results are worth to be published in ranked international journals of applied geophysics. I certify that this dissertation contains only valid data and that the presented results are representing the Candidate's own work. I also declare that it is fully adequate in scope and quality required by the Mikoviny Sámuel Doctoral School of Earth Sciences. Based on the above, I support and recommend the public defence of the thesis and the award of the PhD title.

08/07/2023, Miskolc

Dr. Mihály Dobróka

professor emeritus

1. Table of Contents

Chapter 1

1. INTRODUCTION

Geophysics, when combined with data processing, plays an essential role in the exploration of raw materials, especially in the constantly changing oil and gas industry. It faces new challenges from the increased availability of unconventional resources, shifting supply and demand factors, and environmental pressures, as well as legislative and tax changes and new trends in supply and demand. Geophysics methods are considered the most cost-effective way to gather subsurface data, as they provide low-cost results that offer a large amount of information about the ground structure and other properties. Initially, these methods focused on assessing the potential resources of a basin. However, seismic imaging data obtained through two or three-dimensional processing proved very helpful in understanding the long-term effects of large structures on the earth, leading to positive results and a successful demonstration of the cost-effectiveness of acquiring data about subsurface layers without direct measurement. Advances in geoscience continue to support this premise. Our understanding of the ground layers grows as discoveries are made through cost-effective methods. Currently, geophysicists find seismic to be a particularly compelling method because its workflows are more cost-effective than previous ones for exploring and developing new reservoirs. Modern acquisition and processing techniques provide high-quality subsurface illumination for both conventional and unconventional reservoirs at a lower cost than before. Prestack methods like AVO are used in conjunction with anisotropic velocity analysis to predict the fluid properties of a reservoir. Seismic attributes can be used to indicate stress orientation, the integrity of the overlying cap rock, fault distribution, and reservoir quality. Additionally, these attributes can be used to identify statistical relationships between payout rate and quality through cross-correlation. Easy access to 3D seismic data makes it an essential aspect of many projects, and it is used in a variety of ways, including predicting fluid distribution and migration through rock formations and assessing the performance of enhanced recovery efforts.

When understanding the earth's properties through seismic data, it is frequently suggested to perform a Fourier transform on the x, y, and z axes. This allows for the conversion of complex differential equations into more easily manageable algebraic equations. Performing this with the time axis also helps convert temporal frequency differential equations into traditional algebraic

ones. Seismic data analysis relies heavily on the use of the Fourier transform throughout the entire process. A seismic wavefield recorded at a receiver location is called a seismic trace, and its digital form is a time-based series of sinusoids with distinct peaks, frequencies, and phases. The digital data can be transformed into sinusoidal components by performing the Fourier forward transform on the trace, and then an individual Fourier transform is performed in reverse to create the seismic trace from sinusoidal components. Seismic data processing algorithms are often more easily understood when converted into the frequency domain, instead of the traditional time-based method. Frequency filters work well for this process because they use Fourier analysis for their design. These filters are typically multichannel or single-channel and use an operand (such as a trace of seismic activity) and an operator (such as a filter) to process data. To work with digital samples of signals in seismic traces, we need to use the Discrete Fourier transform (DFT) which provides a way to analyze and understand discrete signals in the frequency domain. It is essentially the digital counterpart to Fourier transforms. On the other hand, the inverse Fourier transform is also known as the Inverse Discrete Fourier transform (IDFT), representing the discrete-time version of the inverse transform.

Several factors can affect the accuracy of seismic data, which can be improved using appropriate noise attenuation strategies and implementing signal processing methodologies. The subsequent stages of data analysis incorporate various approaches, including signal processing, statistical analysis, and algebraic calculations. These methods require expertise from both geophysics and other disciplines. Each method typically pertains to a specific target, such as frequency and wavelength filtering, velocity analysis, static correction, deconvolution, and time/depth migration (Yilmaz, 2001). Noise can generate unwanted features, and can be divided into incoherent (random) and coherent noise. Incoherent noise can be shown as temporal and spatially random noise, while coherent noise can be shown as linear noise, repercussion, or multiples. The ground roll may also appear in land data surveying and dominate the reflection energy in recorded data. Coherent noise contains low frequencies and large amplitudes (Yilmaz, 2001). This type of noise can be processed by F-K filtering and inverse velocity stacking, but field results have shown that the noise is still present in the seismic data (Maurya et al., 2019). All of these types of noise can cause significant artefacts that have a significant negative impact on interpretation results, from simple structural attributes to prestack impedance inversion and amplitude variations from seismic azimuth-gathers (AVAz) analysis. These noises are linearly projected into the frequency domain

during Fourier transformation, meaning that traditional FT algorithms are most sensitive to non-Gaussian noise.

Due to technical issues and the current state of the economy, construction projects may be delayed, leading to missing data in seismic records. Many processes, such as reverse time migration, reverse vector tilt, full waveform inversion, amplitude variation with offset, and more, are heavily impacted by missing data. Therefore, before the data can be used to obtain high-quality results, it needs to be reconstructed from initial seismic data. This becomes increasingly important as the preprocessing stages progress. Different methods have been developed to deal with gaps in seismic records. One such method is wave equation-based reconstruction, which uses the physical properties of seismic waves to create a wave field. However, this method can be costly to use in some applications due to computational limitations. The F-X and F-K domains provide a linear, predictable pattern for prediction filters to use, but this requires the seismic signal data to be equally spaced.

Attributes are specific pieces of hidden information contained within a seismic wave. They represent a smaller subset of the overall information in the original wave and are displayed at the same scale, making them easy to see and understand. They can be computed using a variety of methods. However, the relationships between each attribute can often be complicated. These attributes are calculated to remove extraneous information from the data, making trends or patterns visible that were not present in the original data. They are intended to provide insight into reservoirs that contain oil and gas, as well as information on how they moved and were trapped, by mapping out the geological features associated with hydrocarbon deposits. Cosentino et al., (2001) describes these features as structural elements, such as the thickness and shape of reservoirs, faults, and other geological features. Additionally, a well's petrophysical properties, such as permeability and porosity, are critical to its functionality. The problem here appears when noise exists as in other geophysical methods. Turning inverse problem theory that can help reduce the noisy effect, and consists of a collection of methods that can reject noise and handle it as an overdetermined inverse problem (Dobróka et al., 2012).

Auken et al., (2005) presented an inversion blueprint for continuous resistivity as laterally constrained data as a 1D model and inverted to one system with lateral transitions. They regularized the model and generated the estimated model, which was full of sensitivity analysis of

model parameters to apply evaluation of the inversion outputs. The inversion of time-lapse data was used to improve the characterization of the reservoir using full wavefield inversion (FWI), which is capable of determining subsurface changes due to production (Routh et al., 2012). The series expansion-based inversion method was then applied to borehole data in the interpretation process (Szabó, 2015, 2011) and also in processing induced polarization data (Turai, 2011). The 1D Fourier transformation was handled by Szegedi and Dobróka (2014) as a robust inverse problem using the Iteratively Reweighted Least Squares (IRLS) algorithm with Cauchy-Steiner weights (Steiner, 1997), and the results appeared as a significant reduction in noise sensitivity of the continuous Fourier transform. The inversion method gives the ability to estimate the underlying model of physical properties of the rock and fluids to achieve good reservoir characterization (Berteussen and Ursin, 1983). Maurya and Singh, (2020) presented many attributes, such as P-impedance, S-impedance, P-wave, and S-wave velocity, using inversion that depends on creating forward modelling that generates a set of model parameters. Marashly and Dobroka, (2021) showed that the IRLS inversion-based Fourier transform method achieved good results in noise resistance when applied to a synthetic wavelet.

In this thesis, a new inversion-based Fourier transformation (C-IRLS-FT) is introduced using Chebyshev polynomials in discretizing the Fourier spectrum. The procedure is applied to synthetic wavelets and synthetic seismic complex models. The outlier sensitivity rejection for 1D data is assessed using both DFT and C-IRLS-FT methods to demonstrate the ability of Chebyshev polynomials in noise rejection for both Cauchy and Gaussian noise. The technique is then compared with the Legendre Polynomial-based Fourier transform method (L-IRLS-FT) for evaluating its performance in eliminating noise in pre-generated seismic data. In more depth, the inversion-based Fourier transform process (C-IRLS-FT) combined with the Most Frequent Value method (MFV) developed by Steiner can effectively make the Fourier transform more robust.

The robustness of the C-IRLS-FT to outliers and its outstanding noise suppression capability justifies the method being applied in the field of seismic data processing. To make the Hilbert transform more robust, we applied our C-IRLS-FT in its calculation and define a robust analytical signal. As an application example, we calculate the absolute value of the analytical signal that can be produced as an attribute gauge (instantaneous amplitude), also instantaneous phase was tested. The new algorithm is based on dual inversion: the Fourier spectrum of the time signal (channel) is

determined by inversion, and the spectrum obtained by the transformation required for the Hilbert transform is transformed back into the time domain using robust inversion. The latter operation is carried out using the Steiner weights calculated inside of the Iterative Reweighting Least Squares (IRLS) method (robust inverse Fourier transform based on inversion). To discretize the spectrum of the time signal, we use the again Chebyshev polynomials in a series expansion. The expansion coefficients are the unknowns in the inversion. The results show that the procedure has remarkable resistance to outlier noise and noise suppression, an order of magnitude better than that calculated by using the conventional DFT in calculating the Hilbert Transform.

Building on the principles of the one-dimensional Hilbert Transform, the two-dimensional (2D) variant has emerged as a potent instrument in image processing. Particularly, it has made considerable strides in earth science-oriented applications, such as edge detection, noise reduction, and image enhancement. The 2D Hilbert Transform is characterized by a convolution operation with a 2D Hilbert kernel, yielding a 2D analytic signal that combines the original image data with its Hilbert Transform. This operation provides valuable insights into spatial frequency content and phase information. In this research, we examine the conventional 2D Hilbert method and the newly developed 2D C-IRLS-HT and 2D H-IRLS-HT (with Hermite basis functions) methods, focusing on their capacity to detect anomalies within noisy datasets. Our findings reveal that the traditional Discrete Fourier Transform (DFT)-based 2D Hilbert Transform encounters numerous challenges, such as substantial noise levels, less distinct edges, and the presence of outliers, which hinders accurate anomaly detection. However, applying the 2D H-IRLS-HT method to a noisy synthetic dataset showed notable improvements in terms of clarity, quality, and noise reduction. The elimination of outliers further enhanced the interpretability of the data, enabling clearer identification of anomaly boundaries. The superior performance of the 2D H-IRLS-HT and 2D C-IRLS-HT methods underscores their potential as reliable alternatives for anomaly detection, especially in challenging environments where conventional techniques might be less effective.

Continuing our exploration, comprehensive geophysical measurements were performed in the eastern region of area Syria, spanning approximately 128 km^2, by an exploration company. Employing a gravity meter and GPS locating, robust measurements were obtained and corrected considering the rock density of the region. These measurements aimed to extract valuable subsurface geological information. During the initial stages of previous research, we utilized

Surfer software to identify the structure within the dataset. As the study progressed, we incorporated a more advanced method, the 2D C-IRLS-FT, particularly in the low-pass filter process. This method was systematically applied to evaluate the gravity measurement dataset. Our primary objective was to assess the effectiveness of this method before any comprehensive analysis using a two-dimensional low-pass Butterworth filter. The 2D C-IRLS-FT method, when integrated with Fourier Transform, enabled the decomposition of gravity data into its frequency components, facilitating a more detailed interpretation of subsurface structures. As the investigation continued, the MINIMAG device captured magnetic readings every twenty seconds, highlighting the necessity of applying the 2D C-IRLS-FT method for pole reduction and overall data quality enhancement. The pole reduction process using C-IRLS-FT enhanced the coherence of the geological structures within the dataset, leading to a noticeable improvement in data quality. This process, coupled with the elimination of noise and outliers, resulted in a more distinct representation of the structure, facilitating a more reliable interpretation. Our work also included a novel technique that combined the 2D C-IRLS-FT method and k-means clustering for denoising seismic data, specifically within a seismic section that contained significant noise. This method involved transforming seismic data into the frequency domain, identifying significant spectral components through clustering, and filtering out noise components. The results demonstrated the effectiveness of this method in denoising seismic data, with the filtered data preserving significant spectral components while eliminating the noise. In essence, this study underscores the value of employing advanced methods like the 2D C-IRLS-FT and k-means clustering for data analysis and interpretation in geophysical studies. These techniques fostered robust and accurate interpretations, paving the way for identifying key geological characteristics and assessing the underlying spatial patterns and relationships more accurately.

Chapter 2

2. THE METHODS OF GEOPHYSICAL INVERSION

Geophysical methods involve examining various physical fields, such as gravity, magnetic, electromagnetic, and seismic waves, as they pass through the earth. These fields are influenced by the physical characteristics of rocks, and their values can be measured and studied. In geophysical data analysis, various geological models are created, and the predicted geophysical data for these models are compared to the actual observed data. This process, known as the forward problem, involves the numerical modeling of geophysical data based on given model parameters and allows for the prediction of geophysical data for specific geological structures. The ultimate aim of geophysical observations is to identify geological structures from the geophysical data. However, this can be challenging due to the intricate nature of the earth's interior. To overcome this difficulty, we often approximate the actual geology with a simplified model and try to determine the model parameters from the data. This process is known as the inverse problem. The effectiveness of geophysical interpretation relies on our ability to approximate real geological structures with reasonable models and solve inverse problems efficiently (Zhdanov, 2015).

Inversion is the process of mathematically reconstructing an image of the earth's subsurface from measured data. It involves complex mathematics and requires a deep understanding of the physics of the geophysical method being used. Where considered an essential tool for exploring and understanding the earth's subsurface, and has various applications in fields such as oil and gas exploration, mineral exploration, and environmental studies. One of the main challenges in inversion is dealing with the inherent uncertainty in the measured data and the complexity of the subsurface. To overcome this, inversion algorithms must consider this uncertainty and try to find the most likely model that is consistent with the data. This often involves iterative processes and the use of regularization techniques to constrain the solution. In addition to creating a subsurface model, inversion can also be used to estimate the physical properties of the subsurface, such as the density, conductivity, or elastic properties of the earth's materials. This information can be used to understand the region's geology and make predictions about the location of natural resources or the presence of subsurface contaminants (Oldenburg and Li, 2005) (Sen and Stoffa, 2013).

In recent times, optimization methods have gained popularity among scholars due to advancements in theory, resulting in improved geophysical inversion methods and data processing. Geologists can use inversion, which involves combining various techniques that measure different physical responses to the same geological structure, to solve geophysical inverse problems. More specifically, to tackle the inverse problem, experts need to first build an initial model of their case and then determine the model type (such as horizontal stratification, dip, or tectonic structure) and its quantitative properties (such as petrophysical and geometric factors). Before examining a structure geologically, it must be examined computationally, requiring access to hardware and software resources. These resources are necessary for measuring physical principles and designing appropriate measurement setups. In addition, this information is essential for calculating theoretical data, which is similar to measured data but in a digital format, about the effects of complex structures.

The first step in the modeling process is to input data and previous model parameters into the model. Once these values are established, forward modeling can be performed to generate a prediction of the resulting data. After the inversion phase is completed, the calculated data is compared to the actual field data. The preliminary model is gradually improved through an iterative process until the computed data and observed phenomena consistently match. This is done by adjusting the model parameters until they consistently align. To achieve this task, an objective function is calculated that characterizes the difference between the two sets of data. An optimization algorithm is then applied to the iterative process until a stopping criterion is met (Ito and Jin, 2014).

2.1. THE APPLIED LINEARIZATION TECHNIQUES OF GEOPHYSICAL INVERSION

Linear inverse methods are a type of geophysical method that has been used for a long time. They work well when the data can be described as a linear function of the model parameters. However, in some cases, it may be possible to linearize the relationship between the data and the model by applying certain conditions. For example, the Zoeppritz equations, which describe reflection coefficients, are a nonlinear function of certain parameters (Aki and Richards, 2002), but they can be approximated by a linear equation when the angles of reflection are small. Another way to linearize the relationship between the data and the model is to consider perturbations in the data and perturbations in the model and find a linear relationship between them. Linear inversion

methods are relatively fast methods that are used to solve several geophysical problems. Let us assume that the M number of model parameters is introduced as:

$$\vec{m} = \{m_1, m_2, \ldots, m_M\}^T \qquad (2\text{-}1)$$

and the number of the measured data (N) is collected into a data vector $\vec{d}_{meas}$

$$\vec{d}_{meas} = \{d_1^{meas}, d_2^{meas}, \ldots, d_N^{meas}\}^T \qquad (2\text{-}2)$$

The calculated data $\vec{d}_{calc}$ can be presented as the following

$$\vec{d}_{calc} = \{d_1^{calc}, d_2^{calc}, \ldots, d_N^{calc}\}^T \qquad (2\text{-}3)$$

The connection between $\vec{d}_{calc}$ and $\vec{m}$ can be represented like

$$\vec{d}_{calc} = \vec{g}(\vec{m}) \qquad (2\text{-}4)$$

where $\vec{g}$ is the response function. To calculate the overall error between measured and calculated data, we used the deviation vector

$$\vec{e} = \vec{d}_{meas} - \vec{d}_{calc} \qquad (2\text{-}5)$$

which is a nonlinear function of the model parameters. To characterize the misfit between the calculated and measured data, we can introduce a scalar function which is usually minimized in the inversion procedure

$$E = E\left(\vec{d}_{meas} - \vec{g}(\vec{m})\right). \qquad (2\text{-}6)$$

To linearize our problem we calculate the response function at $\vec{m} = \vec{m}_0 + \Delta\vec{m}$, in a close neighborhood of a $\vec{m}_0$ reference model. After expanding Eq(2-4) using Taylor's series around the reference model and ignoring the higher-order terms we can obtain

$$\vec{g}(\vec{m}_0 + \Delta\vec{m}) = \vec{g}(\vec{m}_0) + G\Delta\vec{m} \qquad (2\text{-}7)$$

where G is the Jacobi matrice with the elements

$$G_{kj} = \left|\frac{\partial g_k}{\partial m_j}\right|_{\vec{m}_0} \qquad (2\text{-}8)$$

This gives the deviation vector in Eq(2-5) as

$$\vec{e} = \Delta \vec{d} - \boldsymbol{G}\Delta \vec{m} \tag{2-9}$$

where $\Delta \vec{d} = \vec{d}_{meas} - \vec{g}(\vec{m}_0)$. After this, the misfit function in Eq(2-6) takes the form

$$E = E\left(\Delta \vec{d} - \boldsymbol{G}\Delta \vec{m}\right) \tag{2-10}$$

Depending on the form of the misfit function a variety of different inversion procedures can be defined. The geophysical inverse theory has brought about numerous innovative ideas and concepts regarding solution methods, the characteristics of solutions obtained, and the circumstances under which solutions may be obtained (Backus and Gilbert, 1967; Backus, 1970; Parker, 1977; Jackson, 1979).

2.1.1. THE OVER-DETERMINATION CASE

When we have more data than unknown model parameters and the data contains information on all model parameters, the problem is called overdetermined. In this situation, all of the model parameters are represented in the data and the observed data points can be effectively fitted by the equations in a linear system by using the methods of least squares. The best estimate in this situation can be found using the least squares method, which minimizes the so-called L_2 norm of the derivation vector (Sen and Stoffa, 2013). This means in our linearized case in Eq(2-10) that

$$E = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N}(\Delta \vec{d} - \boldsymbol{G}\Delta \vec{m})_k^2$$

To simplify the notations, in the following, we omit the symbol Δ and write

$$E = \sum_{k=1}^{N}(\vec{d} - \boldsymbol{G}\vec{m})_k^2 \tag{2-11}$$

2.1.1.1 THE GAUSSIAN LEAST SQUARES (LSQ) METHOD

The Gaussian Least Squares method is a significant linear inversion technique that involves minimizing the squared $L2$ norm of the deviation vector. This method was first introduced and formulated by Gauss and Davis in (1809) as a solution to the overdetermined inverse problem. The objective (or misfit) function in Eq(2-11) that needs to be minimized can be written as:

$$E = \sum_{k=1}^{N} d_k^2 + \sum_{i=1}^{M}\sum_{j=1}^{M} m_i m_j \left(\sum_{k=1}^{N} G_{kj} G_{ki}\right) - 2\sum_{j=1}^{M} m_j \left(\sum_{k=1}^{N} G_{kj} d_k\right)$$

When E is minimized:

$$\frac{\partial E}{\partial m_l} = 0 \qquad for \; l = 1,2,3, \dots, M \qquad (2\text{-}12)$$

Zero will be the result of the derivative of the first term, for the second term will be:

$$\frac{\partial}{\partial m_l}\left(\sum_{i=1}^{M}\sum_{j=1}^{M} m_i m_j \sum_{k=1}^{N} G_{Ki} G_{Kj}\right) = 2\sum_{i=1}^{M} m_i \sum_{k=1}^{N} G_{ki} G_{kl} \qquad (2\text{-}13)$$

Where j and i are exchanged. The third term can write like the following:

$$\frac{\partial}{\partial m_l}\sum_{i=1}^{M} m_i \sum_{k=1}^{N} G_{ki} G_{kl} = 2\sum_{k=1}^{N} G_{ki} d_k \qquad (2\text{-}14)$$

Inserting equations 2-13 and 2-14 in 2-12:

$$\sum_{i=1}^{M} m_i \sum_{k=1}^{N} G_{ki} G_{kl} = \sum_{k=1}^{N} G_{ki} d_k \qquad (2\text{-}15)$$

The vector shape of this equation can be formed like this:

$$\boldsymbol{G}^T \boldsymbol{G} \vec{m} = \boldsymbol{G}^T \vec{d} \qquad (2\text{-}16)$$

By multiplying $(\boldsymbol{G}^T \boldsymbol{G})^{-1}$ with Eq (2-16) we can have:

$$\vec{m} = (\boldsymbol{G}^T \boldsymbol{G})^{-1} \boldsymbol{G}^T \vec{d} \qquad (2\text{-}17)$$

The solution can be formed as the following:

$$\vec{m} = \boldsymbol{G}^{-g} \vec{d} \qquad (2\text{-}18)$$

where $\boldsymbol{G}^{-g} = (\boldsymbol{G}^T \boldsymbol{G})^{-1} \boldsymbol{G}^T$. In an overdetermined problem scenario, the model vector can be calculated using $\boldsymbol{G}^{-g}$ concerning the generalized inverse of the Gussin Least Squares method.

2.1.1.2 THE WEIGHTED LEAST SQUARES (WLSQ) METHOD

Noise is always present in geophysical measurements, so it is crucial for the methods used to process these data to be sensitive to it. Geophysicists often encounter regular noise in their data that can obscure their observations, which has been a persistent problem for them to address. This noise can come from various sources and can be caused by inadequate data processing. Removing estimates of noise from records is believed to help reduce both systematic and non-systematic noise in the records, as proposed by Nyman and Gaiser in 1983, and also by Butler and Russell in 1993. Other methods proposed to reduce noise include subtracting an estimate of noise from recorded data (JeffryesB, 2002; Butler and Russell, 2003; Meunier and Bianchi, 2002; Saucier et al., 2005), and assuming that the noise is stationary for the entire duration of the record (Butler and Russell, 2003). There are numerous reasons why the attributes of noise inevitably change over time. This can be mitigated by using filters that employ inversion techniques or patterns that use pattern-based schemes, such as those suggested by Guitton and Symes, (2003). Additionally, effective methods require using filters with estimable models, which can be time-consuming, as well as implementing filters with pattern-based schemes (Haines et al., 2007).

When gathering data with varying degrees of accuracy due to different types of noise, it is important to proportionally weight each datum to achieve a solution that accounts for all variables. This is best accomplished by using a symmetric weight matrix, such as W. Data that is included in the main diagonal of the matrix represents the weights. The weighted scheme improves the inversion's performance by reducing the impact of noisy data with large errors. This is particularly helpful when dealing with outliers.

The distance between data predictions and measurements can be mitigated by incorporating a weighted matrix as the following:

$$E = \sum_{k=1}^{N}\left(d_k - \sum_{i=1}^{M} G_{ki}m_i\right)\sum_{r=1}^{N} W_{kr}\left(d_r - \sum_{j=1}^{M} G_{rj}m_j\right) \tag{2-19}$$

The optimum of the objective function is fulfilled when $\frac{\partial E}{\partial m_l} = 0$

$$\sum_{i=1}^{M} m_i \sum_{k=1}^{N}\sum_{r=1}^{N} W_{kr}G_{ki}G_{rl} - \sum_{k=1}^{N} d_k \sum_{r=1}^{N} W_{kr}G_{rl} = 0 \tag{2-20}$$

The solution will be:

$$\vec{m} = \boldsymbol{G}^{-g}\vec{d} \qquad (2\text{-}21)$$

Where $\boldsymbol{G}^{-g} = (\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{G})^{-1}\boldsymbol{G}\boldsymbol{W}$ is a generalized inverse matrix.

2.1.1.3 THE ITERATIVELY REWEIGHTED LEAST SQUARES (IRLS) METHOD

IRLS allows for finding the most likely estimates for a generalized linear model by using robust regression to mitigate the influence of outliers on data. This is done by solving Lp approximation in the inverse problem through the use of a weighting matrix that is iteratively updated multiple times before completion. This method, as opposed to finding the least square errors, minimizes the effects of data distribution irregularities. For example, IRLS searches for an estimator that minimizes least absolute errors rather than least square errors. IRLS benefits from the ability to work with Gauss-Newton and Levenberg-Marquardt algorithms as well as other non-convex and linear programming methods by taking advantage of the Lp norm function's overall approach:

$$E_P = \sum_{k=1}^{N} \left| d_k - \sum_{j=1}^{M} G_{kj}m_j \right|^{P} \qquad (2\text{-}22)$$

When $\frac{\partial E}{\partial m_l} = 0$ to solve the problem we can form the Eq (2-22) like:

$$\sum_{i=1}^{N} m_i \sum_{k=1}^{N} G_{ki} \sum_{s=1}^{N} W_{ks}G_{sl} = \sum_{s=1}^{N} G_{sl} \sum_{k=1}^{N} W_{ks}d_k \qquad (2\text{-}23)$$

Eq (2-23) can be written in vectorial shape as the following:

$$\boldsymbol{G}^T\boldsymbol{W}\boldsymbol{G}\vec{m} = \boldsymbol{G}^T\boldsymbol{W}\vec{d} \qquad (2\text{-}24)$$

The $\boldsymbol{W}$ symbol represents the matrix weight $\boldsymbol{W}_{kk} = \left| d_k - \sum_{i=1}^{M} G_{ki}m_i \right|^{p-2}$ where m_i are the unknowns. Noisy data negatively influences the solution because its weight is inversely proportional to the differences between calculated and observed data. Consequently, noisy data makes less of a contribution than data that's less noisy. The weighting matrix is chosen as an identity matrix when (p = 2) and the model vector is estimated as a first loop from:

$$\vec{m}^{(1)} = (\boldsymbol{G}^T\boldsymbol{G})^{-1}\boldsymbol{G}^T\vec{d} \qquad (2\text{-}25)$$

In the second loop of iteration, Eq (2-25) is solved iteratively as:

$$\boldsymbol{G}^T\boldsymbol{W}^{(1)}\vec{m}^{(2)} = \boldsymbol{G}^T\boldsymbol{W}^{(1)}\vec{d} \qquad with \ \boldsymbol{W}_{kk} = \left|d_k - \sum_{i=1}^{M}G_{ki}m_i\right|^{p-2} \qquad (2\text{-}26)$$

Eq (2-26) is linear, the solution for it where $\boldsymbol{W}^{(1)}$ is connected to $\vec{m}^{(1)}$ is:

$$\vec{m}^{(2)} = \left(\boldsymbol{G}^T\boldsymbol{W}^{(1)}\boldsymbol{G}\right)^{-1}\boldsymbol{G}^T\boldsymbol{W}^{(1)}\vec{d} \qquad (2\text{-}27)$$

For the i-th loop of iteration to the nonlinear equation, we can type:

$$\boldsymbol{G}^T\boldsymbol{W}^{(i-1)}\vec{m}^{(i)} = \boldsymbol{G}^T\boldsymbol{W}^{(i-1)}\vec{d} \qquad (2\text{-}28)$$

The weighting matrix is updated iteratively using the Least Squares method, which is why the method is referred to as Iteratively Reweighted Least Squares.

2.2. ACCURACY CALCULATION AND ERROR ESTIMATION

Geophysicists must test and evaluate their inversion methods using synthetic datasets before using real field data measurements. This allows them to understand the impact of random noise on their model's parameters and to properly perform the geophysical inversion. They use noise-filled synthetic datasets to solve geophysical problems, and when working with real field data measurements, they may see a correlation between model parameters and associated errors. It is important to accurately measure and account for random noise in the measurements and to quantify estimation errors, which are often assessed through various model acceptance criteria. The relative data distance is typically utilized to measure the discrepancy between the measured and calculated data using the following equation where N represents the number of inverted data:

$$D = \sqrt{\frac{1}{N}\sum_{K=1}^{N}\left(\frac{d_K^{(measured)} - d_K^{(calculated)}}{d_K^{(measured)}}\right)^2} * 100\% \qquad (2\text{-}29)$$

In synthetic inversion experiments, the relative model distance is used to assess the quality of the estimated model using the following:

$$D = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{m_i^{(estimated)} - m_i^{(exact)}}{m_i^{(estimated)}}\right)^2} * 100\% \qquad (2\text{-}30)$$

Where M represents the number of model parameters. The mathematical expression for a covariance matrix, as described by Menke (1984), to evaluate the quality of model parameter estimates is typically represented by the following equation:

$$cov(\vec{m}) = G^{-g}cov(\vec{d})(G^{-g})^T \tag{2-31}$$

Where $cov(\vec{m})$ and $cov(\vec{d})$ are the covariance matrix of the model parameters and the measurement data, respectively, G^{-g} is the Generalized Inverse matrix and it can be represented as:

$$G^{-g} = (G^TWG)^{-1}G^TW \tag{2-32}$$

and W is the weighting matrix (used to account for any known errors or uncertainties in the data). The superscript T denotes the transpose of a matrix. It's important to notice that this is a general expression, and the specific form of the weighting matrix, W, and Jacobian matrix, G, will depend on the specific problem and data at hand. The i-th element of the main diagonal of the model covariance matrix is used to calculate the estimation error of the i-th model parameter:

$$\sigma_i^{(m)} = \sqrt{COV_{ii}^{(m)}} \tag{2-33}$$

Inversion methods often produce correlated parameter estimates. To determine the level of correlation between the model parameters, a correlation matrix is calculated:

$$corr_{ij}^{(m)} = \frac{COV_{ij}^{(m)}}{\sqrt{COV_{ii}^{(m)}COV_{jj}^{(m)}}} \tag{2-34}$$

Model correlation matrices include the main diagonal, which includes elements that always equal 1. The elements on the off-diagonal range from -1 to +1. Taking into consideration correlation coefficients, the best result is when these range from 0 to ±0.5. This tool can be used to determine the level of correlation between parameters in a model. When values for both problems are close to negative or positive 1, a suggested solution will be considered unreliable. The correlation matrix can be formed by one scalar like:

$$S = \sqrt{\frac{1}{M(M-1)}\sum_{i=1}^{M}\sum_{J=1}^{M}\left(corr_{ij}^{(m)} - \delta_{ij}\right)^2} \tag{2-35}$$

Where S represents the mean spread which considers a good way to measure reliability

Chapter 3

3. CHEBYSHEV POLYNOMIAL-BASED ROBUST FOURIER TRANSFORMATION

It is well known that in traditional Fourier transformation (DFT, FFT) the data noise is directly projected from the time domain to the frequency domain. It was proved by Dobróka et al., (2012) that in the framework of a series expansion-based inversion algorithm, the noise sensitivity of the Fourier transformation can be sufficiently reduced. The essential step of the procedure is expanding the complex Fourier spectrum in series utilizing a set of orthogonal basis functions. The expansion coefficients are determined in the framework of an overdetermined inverse problem. This inversion-based Fourier transformation can be robustified by introducing Cauchy-Steiner weights (Steiner, 1997) in an Iteratively Reweighted Least Squares (IRLS) algorithm. As base functions, Dobróka et al., (2012) used Hermite functions, which are eigenfunctions of the Fourier transformation. In discretizing the spectrum Legendre polynomials were used by Nuamah et al., (2021). In this chapter, a new Chebyshev polynomial-based robust Fourier transformation will be introduced.

3.1. CHEBYSHEV POLYNOMIALS AS BASIS FUNCTIONS

The Chebyshev polynomials, named after the Russian mathematician Pafnuty Chebyshev who studied them in the 19th century, are a sequence of orthogonal polynomials with many important properties and applications in various fields such as numerical analysis and signal processing. The Chebyshev polynomials of the first kind fulfill the differential equation:

$$(1 - x^2)y'' - xy' + n^2 y = 0 \tag{3-1}$$

where $|x| \leq 1$ and n represent integer numbers. They can also be generated by using the recurrence equations

$$T_0(x) = 1 \tag{3-2}$$

$$T_1(x) = x \tag{3-3}$$

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \quad, n > 1 \qquad (3\text{-}4)$$

The first five Chebyshev polynomials of the first kind are plotted below (Figure 3-1). It can be seen that these functions fulfill the symmetry conditions

$$T_n(-x) = (-1)^n T_n(x) \qquad (3\text{-}5)$$

and on the interval $-1 \leq x \leq 1$ all of the extrema have values that are either -1 or 1:
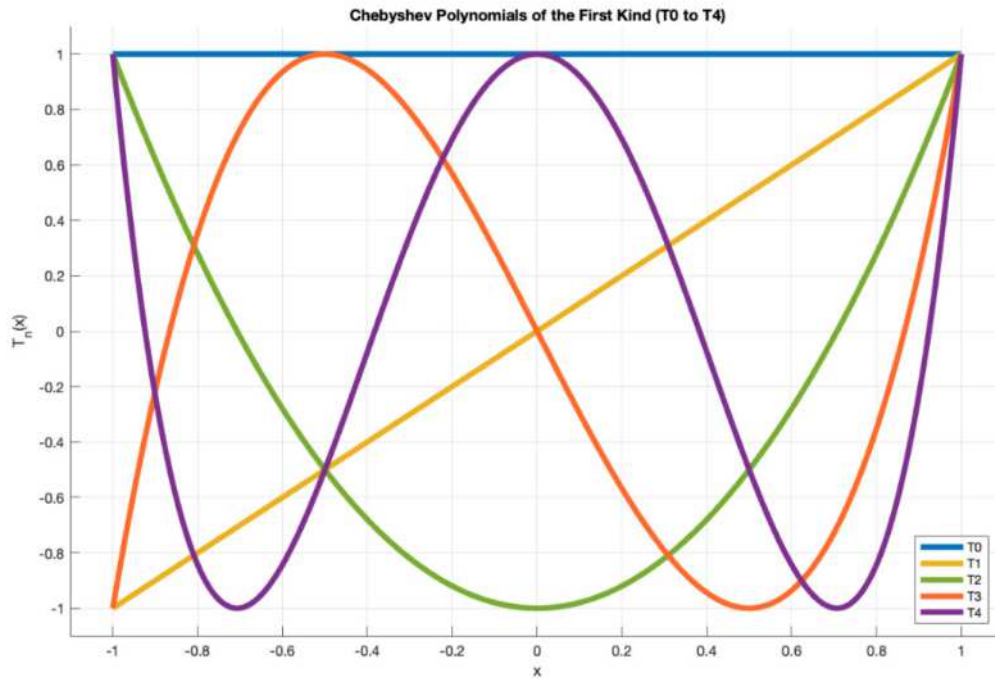
$$T_n(1) = 1, \quad T_n(-1) = (-1)^n \qquad (3\text{-}6)$$

Figure 3-1: Chebyshev Polynomials - First type

In applying these polynomials as basis functions in a series expansion, the extrema features can result in resolution problems at large n values because $T(n, x)$ and $T(n + 2, x)$ are close to each other near x=(+/-) 1. Because of this reason, we will apply also the Chebyshev polynomials of the second kind which fulfill the differential equation

$$(1 - x^2)y'' - 3xy' + n(n + 2)y = 0 \qquad (3\text{-}7)$$

and can be generated by the recurrence formulae

$$U_0(x) = 1 \tag{3-8}$$

$$U_1(x) = 2x \tag{3-9}$$

$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x) \quad , n > 1 \tag{3-10}$$

The first five Chebyshev polynomials of the second kind are plotted below (Figure 3-2).

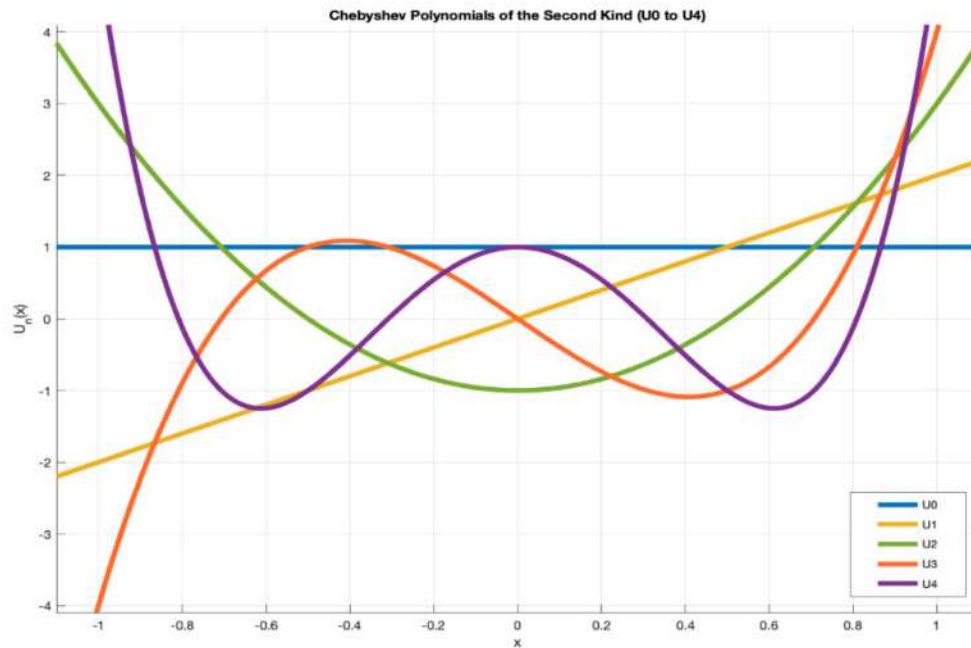*Figure 3-2: Chebyshev Polynomials - Second type*

It can be seen that these functions fulfill the symmetry conditions

$$U_n(-x) = (-1)^n U_n(x) \tag{3-11}$$

and on the interval $-1 \leq x \leq 1$ the extrema have different values

$$U_n(1) = n + 1 \tag{3-12}$$

$$U_n(-1) = (-1)^n(n+1) \tag{3-13}$$

The Chebishev polynomials have important orthogonality properties. The first kind of polynomials considers orthogonal on the interval $-1 \le x \le 1$ with weight function $w(x) = \frac{1}{\sqrt{1-x^2}}$

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 \ if \ n \ne m \\ \pi \ if \ n = m = 0 \\ \dfrac{\pi}{2} \ if \ n \ m \ne 0. \end{cases} \tag{3-14}$$

Also, the Chebyshev polynomials of the second kind can be considered orthogonal on the interval $-1 \le x \le 1$ with weight function $w(x) = \sqrt{1-x^2}$

$$\int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2} \, dx = \begin{cases} 0 \ if \ n \ne m \\ \dfrac{\pi}{2} \ if \ n = m \end{cases} \tag{3-15}$$

3.2. THE C-LSQ-FT AND C-IRLS-FT ALGORITHM IN 1D

As is well known, the traditional Fourier transformation is highly noise sensitive. On the other hand, inverse problem theory encompasses a variety of noise rejection methods. So it is straightforward to expect, that by formulating the Fourier transform as an inverse problem, the noise sensitivity can be appreciably reduced. This idea was followed at the Department of Geophysics (University of Miskolc). Concerning the selection of the basis functions, two kinds of inversion-based Fourier transform algorithms were developed: Hermite function-based (H-LSQ-FT and H-IRLS-FT) method (Dobróka et al., (2012); Szegedi and Dobróka, (2014); Dobróka et al., (2015)) and the Legendre polynomial-based (L-LSQ-FT and L-IRLS-FT) algorithm (Nuamah and Dobroka, (2019) and Nuamah et al., (2021)). As a third possibility in this chapter, we formulate inversion-based Fourier transformation algorithms employing the Chebyshev polynomials for the discretization of the continuous Fourier spectra. A new 1D Fourier transformation algorithm, the Chebyshev Polynomial Least-Squares Fourier Transformation (C-LSQ-FT), is presented together with the robust Chebyshev Polynomial Iteratively Reweighted Least-Squares Fourier Transformation (C-IRLS-FT).

Data conversion from the time domain to the frequency domain can be established using a Fourier transform. For the one-dimensional case, the $D(\omega)$ Fourier transform of the time-dependent $d(t)$ function is defined as

$$D(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d(t)e^{-j\omega t} dt$$

where t denotes the time, ω is the angular frequency and j is the imaginary unit. The inverse Fourier transform ensures a return from the frequency domain to the time domain:

$$d(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D(\omega)e^{j\omega t} d\omega.$$

In the framework of the inversion-based Fourier transformation the $D(\omega)$ frequency spectrum should be discretized using a finite series expansion

$$D(\omega) = \sum_{n=1}^{M} B_n \, \psi_n(\omega) \tag{3-16}$$

Where the parameter B_n is a complex-valued expansion coefficient and ψ_n is a member of an accordingly chosen set of real-valued basis functions. Using the terminology of (discrete) inverse problem theory, the theoretical values of time-domain data in the k-th sampling time t_k (forward problem) can be given by the inverse Fourier transform

$$d(t_k) = d_k^{theor} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} D(\omega)e^{j\omega t_k} d\omega$$

Inserting the expression given in Eq(3-16) one finds that

$$d_k^{theor} = \sum_{n=1}^{M} B_n G_{k,n} \tag{3-17}$$

where the Jacobi matrix is introduced

$$G_{k,n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_n(\omega)e^{j\omega t_k} d\omega$$

as the inverse Fourier transform of the ψ_n basis function:

$$G_{k,n} = \mathcal{F}_k^{-1}\{\psi_n(\omega)\} \tag{3-18}$$

In our present approach, the Chebyshev polynomials serve as the model's basis function for parameterization using $T_n(\omega)$ for the first kind

$$G_{k,n} = \mathcal{F}_k^{-1}\{T_n(\omega)\} \tag{3-19}$$

and $U_n(\omega)$ for the second kind.

To calculate the elements of the Jacoby matrix we can use a standard inverse DFT procedure:

$$G_{k,n} = IDFT_k\{T_n(\omega)\} \tag{3-20}$$

In this case, the deviation vector can be formed like this:

$$e_k = d_k^{meas} - d_k^{theor} = d_k^{meas} - \sum_{N=1}^{M} B_n G_{kn} \tag{3-21}$$

where d_k^{meas} represent the measurement signal. The normal equation of the Gaussian Least Squares method, after using the L$_2$-norm to measure the misfit and minimizing it, can be written as:

$$\vec{B} = (G^T G)^{-1} G^T \vec{d}^{meas} \tag{3-22}$$

In the knowledge of the expansion coefficients the estimated spectrum can be given as the following:

$$D^{est}(\omega) = \sum_{n=1}^{M} B_n T_n(\omega) \tag{3-23}$$

The previous equations rely on the inversion-based Fourier Transformation method, which uses Chebyshev polynomials based on the Least Square Fourier Transformation method (C-LSQ-FT). However, this method only works effectively with data sets that have noise that is distributed in a regular pattern. In cases where the data contains outliers, the procedure has less efficiency in processing them. As Barnett and Lewis (1994) stated, outliers are different from the remaining data, and LSQ will not give acceptable results. Outliers can appear as abnormal, deviant, incongruous, or anomalous (Aggarwal, 2013), leading to serious problems such as high error

variance in statistical power, decreased normality in the data, and corrupting the true relationship between exposure and outcome in model bias (Osborne and Overbay, 2004). To solve this problem, we used the Iterative Reweighted Least Squares method that minimizes the deviation vector via Cauchy-Steiner weights in combination with the Fourier transform using Chebyshev polynomials as basis functions for discretization, creating a robust algorithm, C-IRLS-FT.

From Eq (3-19), which shows the Jacobian matrix derived from the inverse FT, and indicates the calculation of the theoretical value of the signal in Eq (3-17), we can use the IRLS inversion algorithm as described by Dobróka et al. (2012), and write the weighted norm as:

$$E_w = \sum_{k=1}^{N} w_k e_k^2 \qquad (3\text{-}24)$$

where Cauchy-Steiner weights represent the term w_k and can be defined as the following (ε is the scale parameter):

$$w_k = \frac{\varepsilon^2}{\varepsilon^2 + e_k^2} \qquad (3\text{-}25)$$

Scales and Gersztenkorn (1988) shows that the problem of a nonlinear inverse problem caused by a non-quadratic misfit function can be solved by applying the IRLS method. As a first step, the misfit function:

$$E_w^0 = \sum_{K=1}^{N} e_k^2$$

can be minimized in a linear set of normal equations:

$$\vec{B}^0 = (\boldsymbol{G}^T\boldsymbol{G})^{-1}\boldsymbol{G}^T\vec{d}^{meas} \qquad (3\text{-}26)$$

and the deviation error:

$$e_k^0 = d_k^{meas} - \sum_{n=1}^{M} B_n^0 G_{kn} \qquad (3\text{-}27)$$

The weight equation can be written like this:

$$w_k^0 = \frac{\varepsilon^2}{\varepsilon^2 + (e_k^0)^2} \tag{3-28}$$

and the new misfit function takes the form:

$$E_w^1 = \sum_{k=1}^{N} w_k^0 (e_k^1)^2 \tag{3-29}$$

The minimization of Eq (3-29) in a linear set of equations can be solved where the weighting matrix $\boldsymbol{W}^0$ is independent of $\vec{B}^1$:

$$\vec{B}^0 = (\boldsymbol{G}^T \boldsymbol{W}^0 \boldsymbol{G})^{-1} \boldsymbol{G}^T \boldsymbol{W}^0 \vec{d}^{meas} \tag{3-30}$$

And this will be in an iterative loop until matching the stop criterion; thus, it's called the Chebyshev polynomials-based Iteratively-Reweighted Least Square Fourier Transform method (C-IRLS-FT).

3.2.1. NUMERICAL TESTING

As we mentioned earlier, the new algorithms should be tested on synthetic data first to test their reliability and output quality.
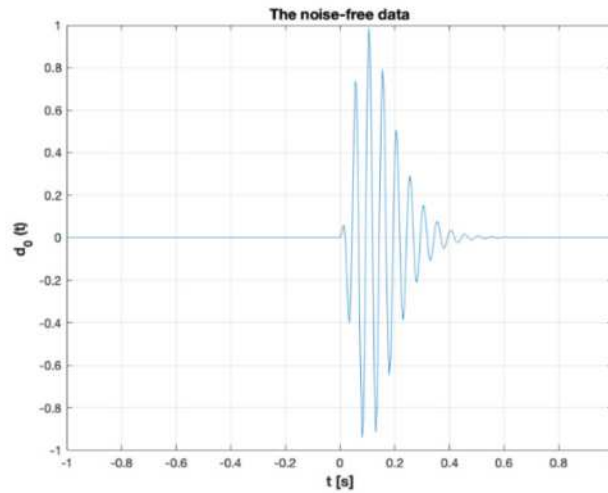
Figure 3-3 : The generated noise-free wave

For that, we create a wave loaded with different types of noise (Gaussian and Cauchy noise) to qualify the noise reduction capability of the introduced algorithms. We start with the following equation to create the data set:

$$d(t_k) = ct_k^n e^{-\lambda t_k}\sin(\omega t_k + \gamma) \qquad (3\text{-}31)$$

and the items n =1, $\lambda = 20$, $\omega = 40\pi$, c = 739, and $\gamma = \pi/4$ are the parameters of the generated wave. The sample rate of the generated waveform is $\Delta t = 0.005\ sec.$ within the interval of [-1,1], as shown in Figure (3-3). The DFT method was applied to generate the real and imaginary parts of the Fourier transform without any noise contamination. In the same step, the C-LSQ-FT and C-IRLS-FT for the first and second kinds of Chebyshev polynomials were also applied, as shown in Figure (3-4).
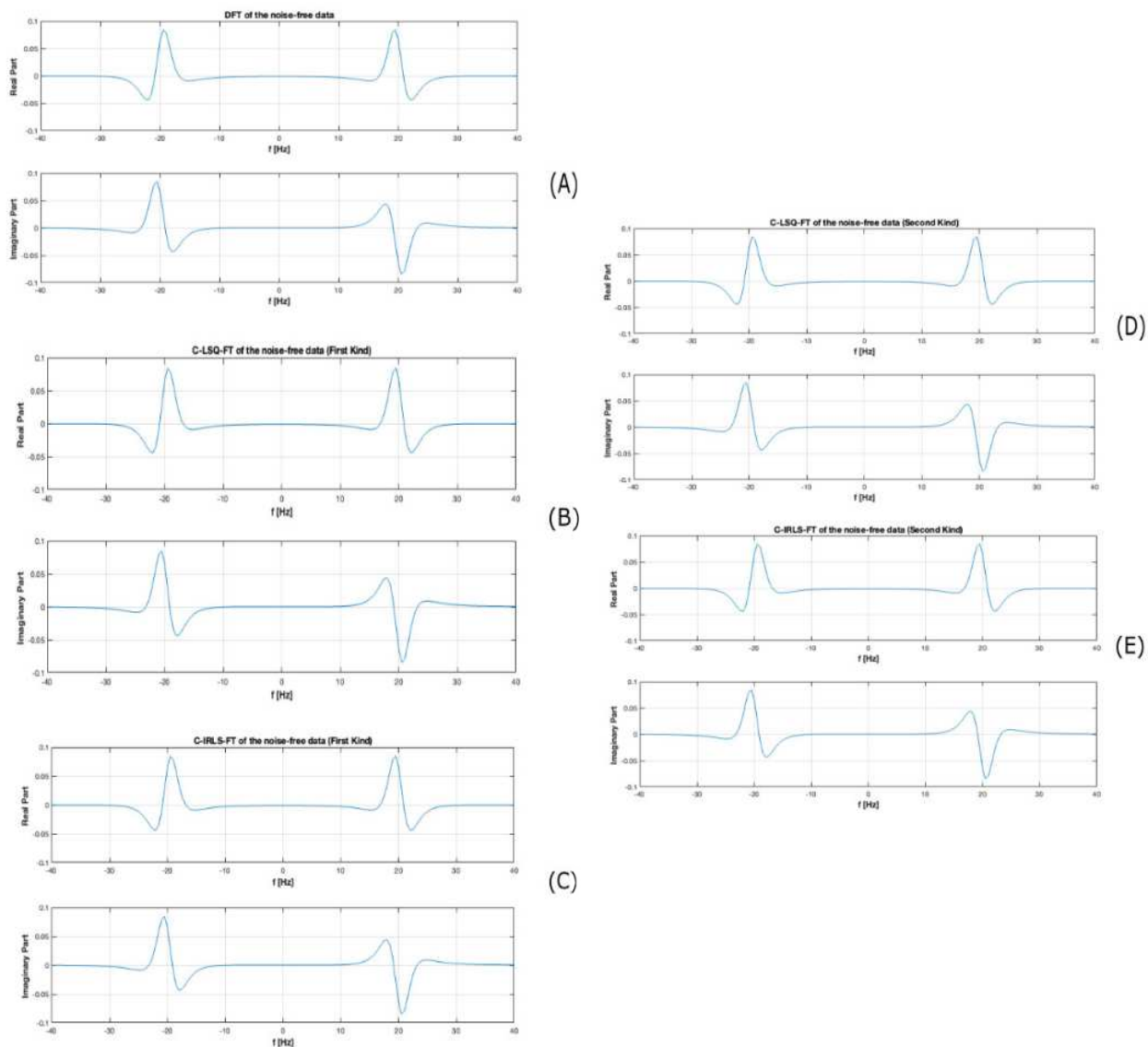
Figure 3-4: The spectrum of the noise-free signal in the frequency domain using (A) DFT (B) C-LSQ-FT FK (C) C-IRLS-FT FK (D) C-LSQ-FT SK (E) C-IRLS-FT SK

The DFT, C-LSQ-FT, and C-IRLS-FT methods produced similar results for the real and imaginary parts of the Fourier-transformed spectrum. Furthermore, in the inversion-based method for high-quality discretization, we use Chebyshev polynomials of the order of M=150. In summary, demonstrated methods are highly suitable for datasets without noise.

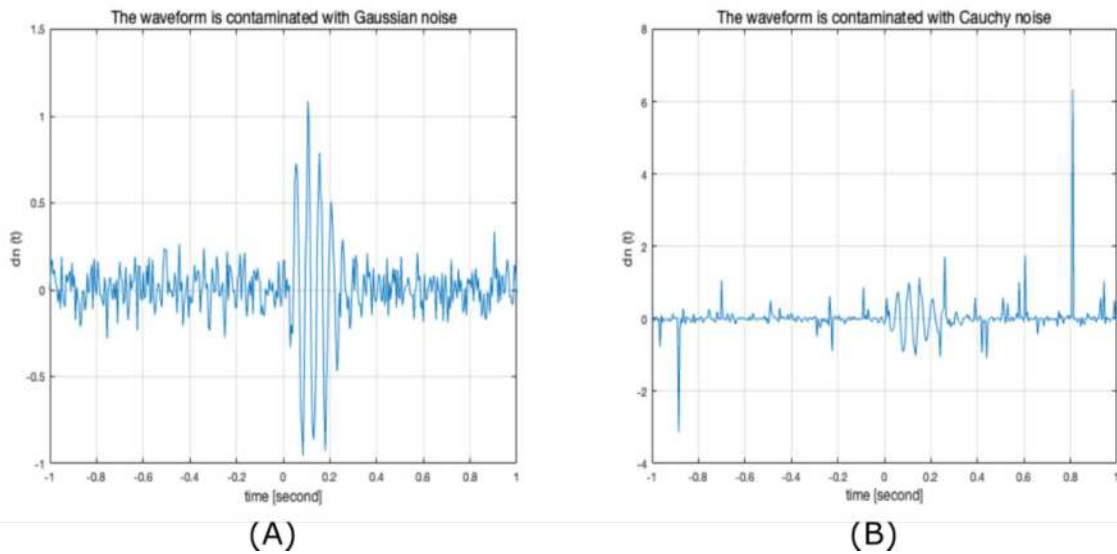Gaussian and Cauchy's noise are respectively loaded onto the previous waveform to test the algorithms separately and see the efficiency, as demonstrated in Figure (3-5).

Figure 3-5: Generated noisy waveform with (A) Gaussian noise (B) Cauchy noise in the time domain

Gaussian noise is generally considered a type of random noise that follows a normal distribution with a zero mean and a specified variance. It is often utilized as a model for external noise in signal processing. It is considered additive, meaning that it does not depend on the signal and can be added to it without altering the signal's distribution. Due to its properties, it is commonly used to simulate the effects of real-world noise in simulations and performance evaluations of signal processing algorithms. It is very important to reduce its impact on the interpretation of the real signal. This includes the development of filters using various transforms such as the Wavelet Transform (Deighan and Watts, 1997), S-Transform (Askari and Siahkoohi, 2008), and Fourier Transform (Dobroka et al., 2012). For testing purposes, the parameters of the applied Gaussian noise are the mean = 0 and $\sigma = 0.01$. As we can see in Figure (3-6), the DFT method was applied to the Gaussian noisy data sets to demonstrate the efficiency of the presented algorithm.

Furthermore, C-LSQ-FT and C-IRLS-Ft were also applied to the same data, as shown in Figure (3-7). To quantify the results, we used data distance:

$$D^{data} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(d_k^{noisy} - d_k^{noise-free}\right)^2} \qquad (3\text{-}32)$$

and the spectral distance

$$D^{spect} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} Re\left(u_i^{noisy} - u_i^{noise-free}\right)^2 + Im\left(u_i^{noisy} - u_i^{noise-free}\right)^2} \qquad (3\text{-}33)$$
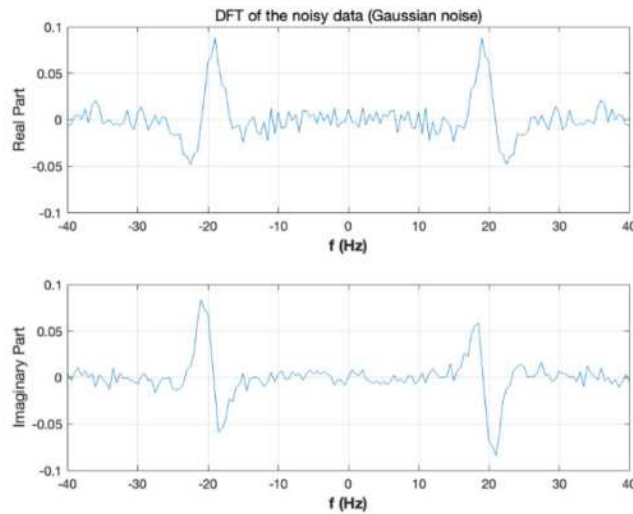
Figure 3-6: Applying DFT on the noisy data set (Gaussian noise) to extracting the real and the imaginary part
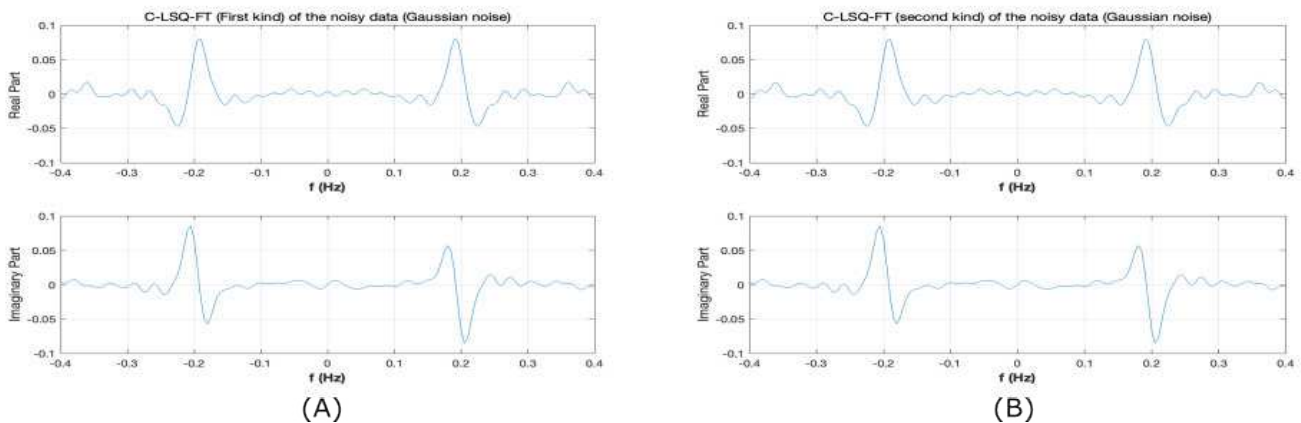
Figure 3-7: The result of noisy data (Gaussian noise) after applying (A) C-LSQ-FT first kind (B) C-LSQ-FT second kind
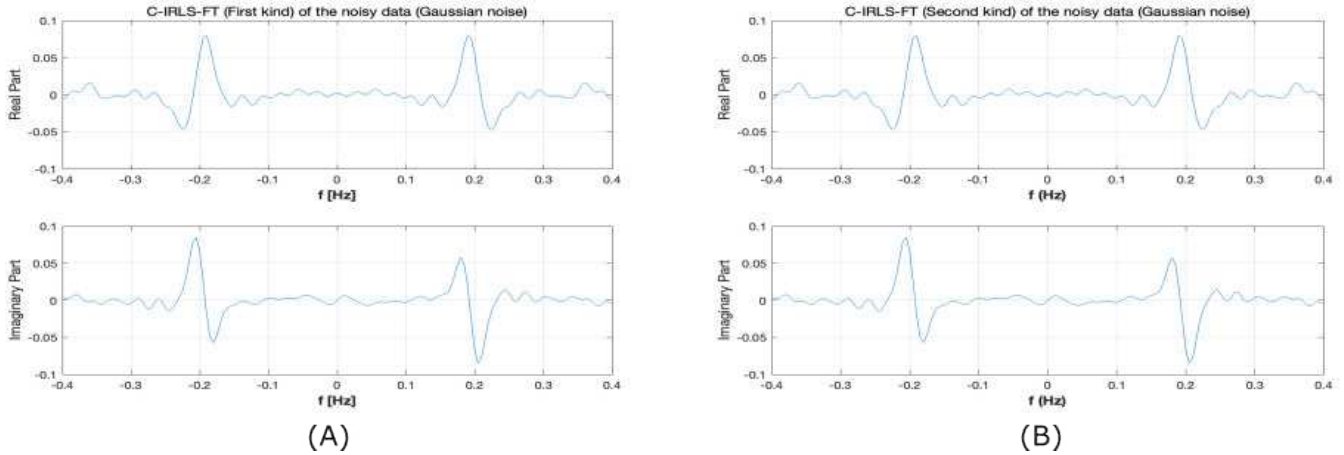
Figure 3-8: The result of noisy data (Gaussian noise) after applying (A) C-IRLS-FT first kind (B) C-IRLS-FT second kind

Figures (3-6), (3-7), and (3-8) show that the present algorithms, C-LSQ-FT and C-IRLS-FT, have similar effective results in reducing Gaussian noise in synthetic data than the traditional DFT.

When conducting a comparison between different methods, it is important to consider the level of error associated with each approach. In the case of methods C-LSQ-FT, C-IRLS-FT, and conventional Discrete Fourier transform method (DFT), notable differences in error levels can be observed. Upon analyzing a Gaussian noisy data set, the data distance between the noisy and noise-free data is 0.1032 (Figures (3-3) and (3-5A)). The conventional DFT displays a spectrum distance of 0.0103 (Figure 3-6), while method C-LSQ-FT display of 0.0079 for the first type of Chebyshev polynomials (Figure 3-7A) and a slight difference of 0.0078 to the second one (Figure 3-7B). And method C-IRLS-FT displays a spectrum distance of 0.0081 for the first type (Figure 3-8A) and 0.0077 for the second one (Figure 3-8B). Note that there is a slight improvement using the second type of Chebyshev polynomials, probably due to the different extrema (Eqs. (3-12), (3-13)).

Processing speed is a critical factor in many computational tasks, and various methods have been developed to enhance it because serial processing involves executing one task at a time, which can be time-consuming and limit overall processing speed. The method C-IRLS-FT consumes slightly more time (between 14-20 seconds) in both types than the C-LSQ-FT (between 8-11second). This is related to the more computing procedure in IRLS.

This data highlights that method C-LSQ-FT and C-IRLS-FT is likely the most accurate approach among the DFT method, which exhibits progressively higher (28-29 %) spectrum distance.
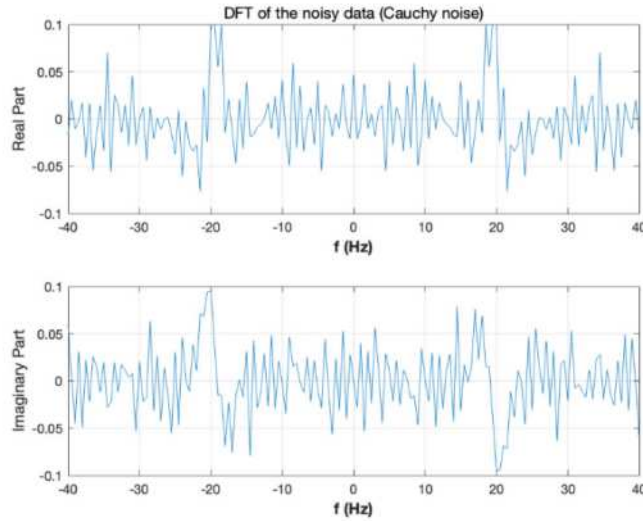
Figure 3-9: Applying DFT on the noisy data set (Cauchy noise) to extracting the real and the imaginary part

In the case of Cauchy noise, Figures (3-9), (3-10), and (3-11) compare the presented algorithms C-LSQ-FT and C-IRLS-FT with the traditional DFT. It reveals a significant reduction in noise and spikes when compared to conventional ones. C-IRLS-FT demonstrates better results (Figure 3-11 (A) and (B)) with both types of Chebyshev polynomials compared to C-LSQ-FT (Figure 3-10 (A) and (B)). The data distance between noisy and noise-free data sets is 0.414. As we compared previous methods in the Gaussian noise data set, we applied the same procedure and observed that the conventional DFT displays a spectrum distance of 0.0416 (Figure 3-9) while method C-LSQ-FT display of 0.0224 for the first kind of Chebyshev polynomials (Figure 3-10A) and slightly different 0.0220 to the second one (Figure 3-10B).
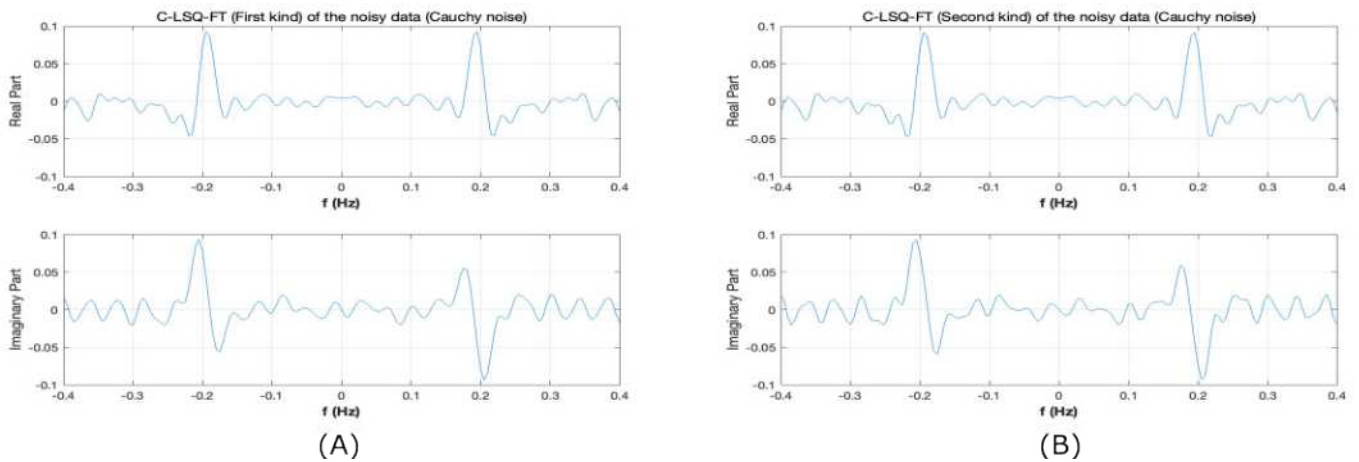

(A)   (B)

Figure 3-10: The result of noisy data (Cauchy noise) after applying (A) C-LSQ-FT first kind
(B) C-LSQ-FT second kind

31

Compared to C-LSQ-FT the method C-IRLS-FT shows an improvement of around 60% in the spectrum distance (0.0143 for the first kind (Figure 3-11A) and 0.0131 for the second one (Figure 3-11B)).

Also, the method C-IRLS-FT consumes slightly more time (between 13-15 seconds) in both types compared to the C-LSQ-FT (between 10-11 seconds). This is related to the more computing procedure in IRLS.
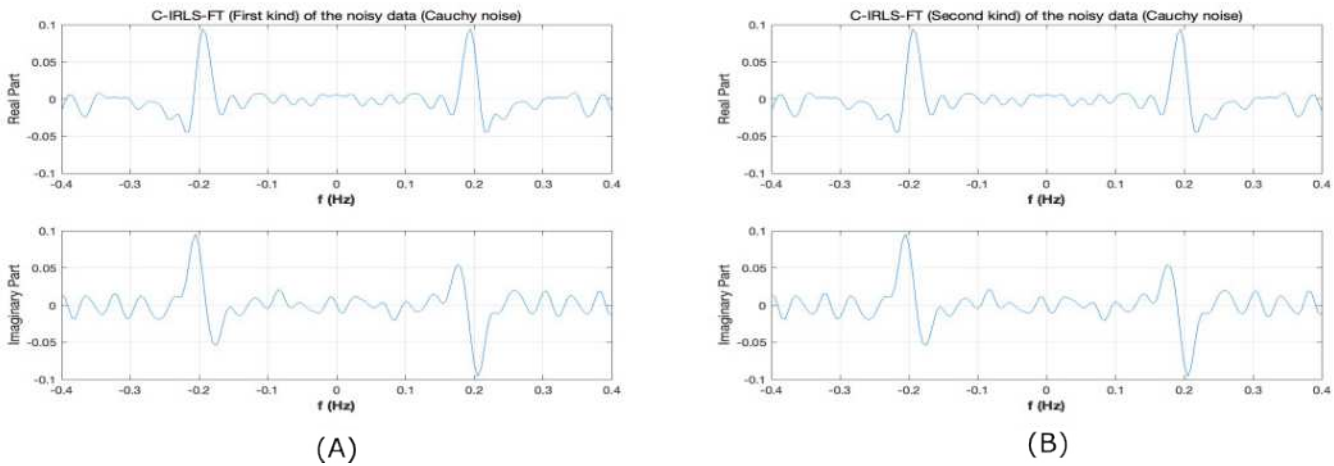

Figure 3-11: The result of noisy data (Cauchy noise) after applying (A) C-IRLS-FT first kind (B) C-IRLS-FT second kind

This greatly highlights the limitations of traditional DFT in effectively eliminating randomly occurring outliers and recursive random noise from a waveform compared with C-LSQ-FT and C-IRLS-FT.

3.2.2. COMPARISON BETWEEN C-IRLS-FT AND AND L-IRLS-FT

Nuamah et al.(2021) introduced two advanced Fourier transformation algorithms the Legendre-Polynomials Least-Squares Fourier Transformation (L-LSQ-FT) and the Legendre-Polynomials Iteratively Reweighted Least-Squares Fourier Transformation (L-IRLS-FT). These algorithms use Legendre polynomials as basis functions for discretizing the Fourier spectrum through series expansion. The algorithms were found to be robust and resistant to sparsely distributed data outliers, particularly when compared to the traditional DFT algorithm in the reduction of magnetic data poles.

It's important, as mentioned before, to compare different methods to determine the right situation in which they are suitable. The same data, without any changes in the waveform or types of noise (Gaussian and Cauchy), was used on the same machine to apply the L-IRLS-FT algorithm and compare it with C-IRLS-FT (Figure 3-12).
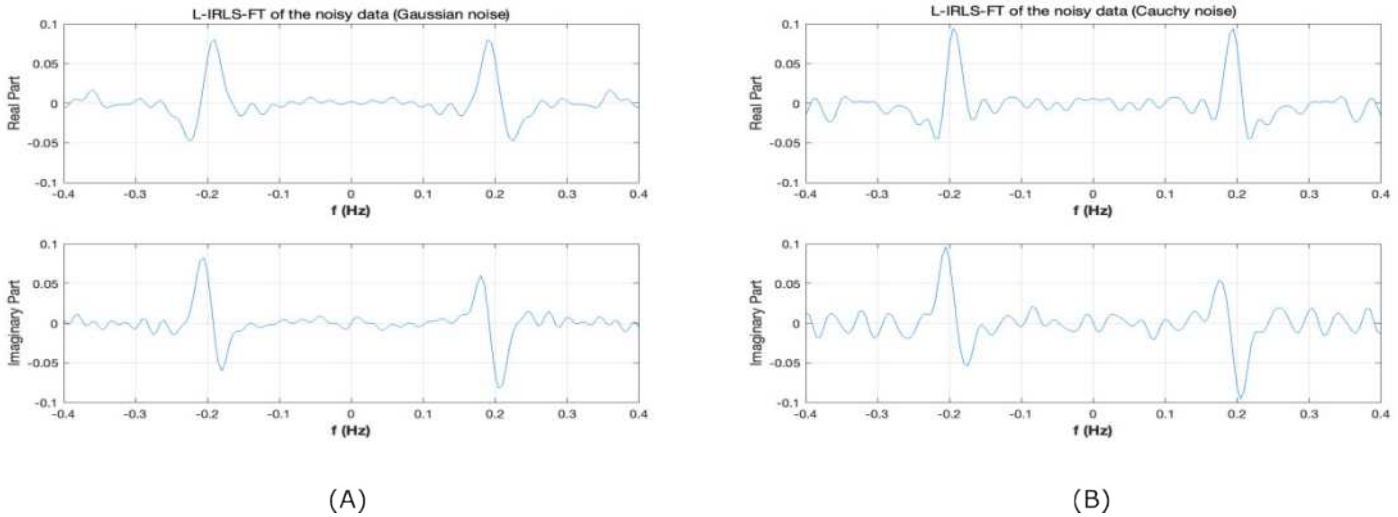

(A)   (B)

Figure 3-12: : The result of L-IRLS-FT algorithm on (A) Gaussian noise (B)Cauchy noise

When comparing the presented algorithm X with Y, an improvement can be observed in terms of lower spectral distance and reduced computation time (from an average of 3 to 5 seconds on the same data set). These factors are very important when selecting a suitable algorithm for the study criteria. The spectral distance of the L-IRLS-FT method is 0.0092, compared to 0.0081 for the first type of the C-IRLS-FT method and 0.0077 for the second type, which is associated with noisy data sets contaminated with Gaussian noise. When Cauchy noise is considered, the spectral distance of the L-IRLS-FT method is 0.0143, compared to 0.0143 for the first type of C-IRLS-FT method and 0.013 for the second type. It can be mentioned that the methods using the Chebishev polynomial of the second kind as a base function give better results compared to methods based on both Legendre- and Chebishev polynomial (of the first kind). The reason is that Legendre polynomials have the same extrema at x=+/- 1 as the Chebishev polynomial (of the first kind).

3.3. The C-IRLS-FT algorithm in 2D

The advancement in signal processing has led to remarkable improvements in the extraction and analysis of vital information from various sources, as we can see before. One-dimensional (1D) noise reduction methods are being extensively utilized for their simplicity and computational efficiency. However, the inherent limitations of 1D techniques have prompted us to explore higher-dimensional approaches, offering improved performance and adaptability. We introduce a new algorithm C-IRLS-FT that enhances the existing 1D noise reduction techniques by extending their applicability to two-dimensional (2D) signals. Many papers discussed noise reduction in 2D geophysical data sets like gravity and magnetic (Dobróka et al., 2017; Nuamah and Dobroka, (2019);Abdelaziz and Dobróka, (2020); Nuamah et al., 2021) using the Hermite function or Legendre polynomials. This innovative using Chebyshev polynomials for the first time offers several advantages over traditional 1D methods, including improved noise reduction capabilities, better handling of complex data structures, and enhanced adaptability to various signals and data types.

Furthermore, the proposed 2D method incorporates advanced signal decomposition techniques, adaptive filtering, and intelligent thresholding mechanisms, collectively contributing to its superior performance. In particular, the method's ability to exploit the spatial and spectral correlations present in 2D signals enables it to effectively suppress noise while preserving the underlying data's essential features and details. This is achieved through the use of inversion-based 2D Fourier transform 2D-IRLS-FT that are designed to automatically identify and eliminate noise components while retaining the informative elements of the signal.

Data conversion from the space domain to the space-frequency domain can be established using a 2D Fourier transform. For the two-dimensional case, the $U(\omega_x, \omega_y)$ Fourier transform of the space-dependent $u(x, y)$ function is defined as

$$U(\omega_x, \omega_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} u(x, y)e^{-j(\omega_x x + \omega_y y)}dxdy$$

where (x, y) denote the space coordinates, (ω_x, ω_y) are the (angular) space frequencies in 2D and j is the imaginary unit. The inverse Fourier transform ensures a return from the space-frequency domain to the space domain:

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} U(\omega_x, \omega_y) e^{j(\omega_x x + \omega_y y)} d\omega_x d\omega_y$$

In the framework of the inversion-based Fourier transformation the $U(\omega_x, \omega_y)$ frequency spectrum should be discretized using a finite series expansion

$$U(\omega_x, \omega_y) = \sum_{n=1}^{N} \sum_{m=1}^{M} B_{n,m} \, \psi_n(\omega_x) \, \psi_m(\omega_y)$$

where the parameters $B_{n,m}$ are complex-valued expansion coefficients, ψ_n and ψ_m are members of an accordingly chosen set of real-valued basis functions. Using the terminology of (discrete) inverse problem theory, the theoretical values of spatial domain data in the (x_k, x_l) sampling point (forward problem) can be given by the inverse Fourier transform

$$u(x_k, x_l) = u_{k,l}^{theor} = \sum_{n=1}^{N} \sum_{m=1}^{M} B_{n,m} \, G_{k,l}^{n,m} \tag{3-34}$$

where the Jacobi matrix is introduced as

$$G_{k,l}^{n,m} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_n(\omega_x) e^{j\omega_x x_k} d\omega_x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi_m(\omega_y) e^{j\omega_y y_l} d\omega_y$$

The Jacobian matrix is the inverse Fourier transform of the basis function ψ_n and ψ_m.

In our investigations the Chebyshev polynomials serve as the model's basis function for parameterization using Eq(3-10) for the first kind and Eq(3-15) for the second one:

$$G_{k,l}^{n,m} = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} T_n(\omega_x) e^{j\omega_x x_k} d\omega_x \cdot \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} T_m(\omega_y) e^{j\omega_y y_l} d\omega_y \tag{3-35}$$

Or in another form:

$$G_{k,l}^{n,m} = \mathcal{F}_k^{-1}\{T_n(\omega_x)\} \cdot \mathcal{F}_l^{-1}\{T_m(\omega_y)\} \tag{3-36}$$

The reason for creating a new Fourier Transformation method revolves around creating a 2D inversion-based version of Eq (3-35). To accomplish this, we will use a standard 2D inverse DFT process:

$$G_{k,l}^{n,m} = IDFT_k\{T_n(\omega_x)\} \cdot IDFT_l\{T_m(\omega_y)\} \tag{3-37}$$

We can create the theoretical data at the sampling points (x_k, y_l):

$$d_{k,l}^{theor} = \sum_{n=1}^{N} \sum_{m=1}^{M} B_{n,m} G_{k,l}^{n,m} \tag{3-38}$$

The programming of the algorithm is more simple after using the transformation of the indices

$i = n + (m-1)N, \quad s = k + (l-1)K$. With these notations, the total number of the unknown expansion coefficient is $I = N + (M-1)N = NM$ and that of the measurement data is $S = K + (L-1)K = KL$. The theoretical data can be calculated as

$$d_s^{theor} = \sum_{i=1}^{M} B_i G_{si}$$

In this case, the deviation vector takes the form:

$$e_s = d_s^{meas} - d_s^{theor} = d_s^{meas} - \sum_{i=1}^{M} B_i G_{si} \tag{3-39}$$

The normal equation of the Gaussian Least Squares method, after using the L$_2$-norm to measure the misfit of the function and minimizing it, can be written as:

$$\vec{B} = (G^T G)^{-1} G^T \vec{d}^{me} \tag{3-40}$$

where the estimated spectrum can be given as the following:

$$D^{est}(\omega_x, \omega_y) = \sum_{n=1}^{N} \sum_{m=1}^{M} B_{n,m} T_n(\omega_x) T_m(\omega_y) \tag{3-41}$$

The previous equations rely on the inversion-based Fourier Transformation method, which uses Chebyshev polynomials based on the Least Square Fourier Transformation method (C-LSQ-FT) as we saw in the 1D case. Also, this method only works effectively with data sets that have noise that is distributed in a regular pattern. In cases where the data contains outliers, the procedure has less efficiency in processing them. To solve this problem, we used the Iteratively Reweighted Least Squares method that minimizes the deviation vector via Cauchy-Steiner weights in combination

with the Fourier transform using Chebyshev polynomials as basis functions for discretization, creating a robust algorithm, C-IRLS-FT.

As in the 1D case from Eq (3-34), which shows the Jacobian matrix derived from the inverse FT, and indicates the calculation of the theoretical data, we can use the IRLS inversion algorithm as described by Dobróka et al., 2012, and write the weighted norm as:

$$E_w = \sum_{s=1}^{S} w_s e_s^2 \tag{3-42}$$

where Cauchy-Steiner weights represent with term w_s and can be defined as the following:

$$w_s = \frac{\varepsilon^2}{\varepsilon^2 + e_s^2} \tag{3-43}$$

Scales and Gersztenkorn (1988) shows that the problem of a nonlinear inverse problem caused by a non-quadratic misfit function can be solved by applying IRLS. As a first step, the misfit function:

$$E_w^0 = \sum_{s=1}^{S} e_s^2$$

can be minimized in a linear set of normal equations:

$$\vec{B}^0 = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \vec{d}^{me} \tag{3-44}$$

and the deviation error:

$$e_s^0 = d_s^{meas} - \sum_{i=1}^{I} B_i^0 G_{si} \tag{3-45}$$

The weight equation can be written like this:

$$w_s^0 = \frac{\varepsilon^2}{\varepsilon^2 + (e_s^0)^2} \tag{3-46}$$

and the misfit function:

$$E_w^1 = \sum_{s=1}^{S} w_s^0 (e_s^1)^2 \qquad (3\text{-}47)$$

The minimization of Eq (3-29) in a linear set of equations can be solved where the weighting matrix $\boldsymbol{W}^0$ is independent of $\vec{B}^1$:

$$\vec{B}^1 = (\boldsymbol{G}^T \boldsymbol{W}^0 \boldsymbol{G})^{-1} \boldsymbol{G}^T \boldsymbol{W}^0 \vec{d}^{meas} \qquad (3\text{-}48)$$

And this will be in an iterative loop until matching the criteria; thus, it's called the Chebyshev polynomials-based IRLS Fourier Transform method (2D C-IRLS-FT).

3.3.1. Numerical testing for 2D C-LSQ-FT and 2D C-IRLS-FT method

To test the proposed method, we create a 2D noise-free data set consisting of a rectangular region with the (normalized) dimensions [-1,1] units for directions x and y. In addition, this data set contains an anomaly in the middle with dimensions [-0.2,0.2] units also in directions x and y. The homogeneous background and the anomaly is characterized by the physical value d_{back}, and d_{anom}, respectively, where $d_{anom}=2*d_{back}$. For numerical tests we introduce the dimensionless function
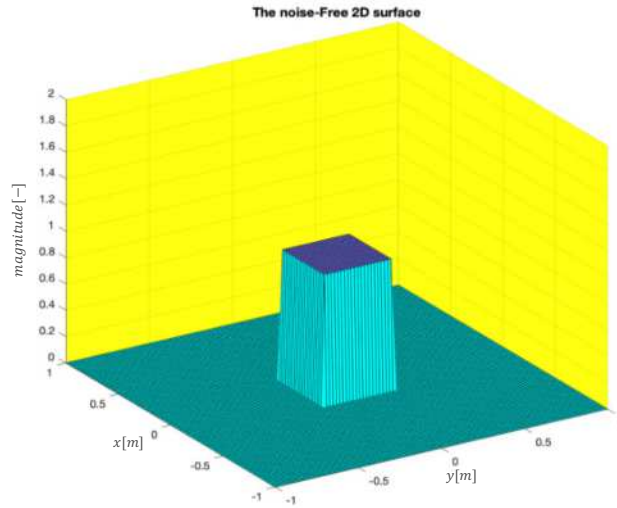

Figure 3-13: The Noise-Free data set

$\text{magnitude}(x,y)=(d(x,y)-d_{back})/d_{back}$

giving the values 0 and 1, in the background and the anomaly, respectively (Figure 3-13). For the sampling intervals, we set it to equal 0.02 units for both directions, creating 101*101 data points.

We started with noise-free data set in comparison to the previously mentioned method. The 2D Fourier spectrum of the data set was computed using 2D DFT, 2D C-LSQ-FT and 2D C-IRLS-FT algorithms.

Figure (3-14) demonstrates the amplitude spectrum using 2D DFT, 2D C-LSQ-FT, and 2D C-IRLS-FT methods. All approaches give similar results without any significant difference, which indicates the effectiveness in processing noise-free datasets. The 2D C-LSQ-FT and 2D C-IRLS-FT used Chebyshev polynomials with M=35 order.
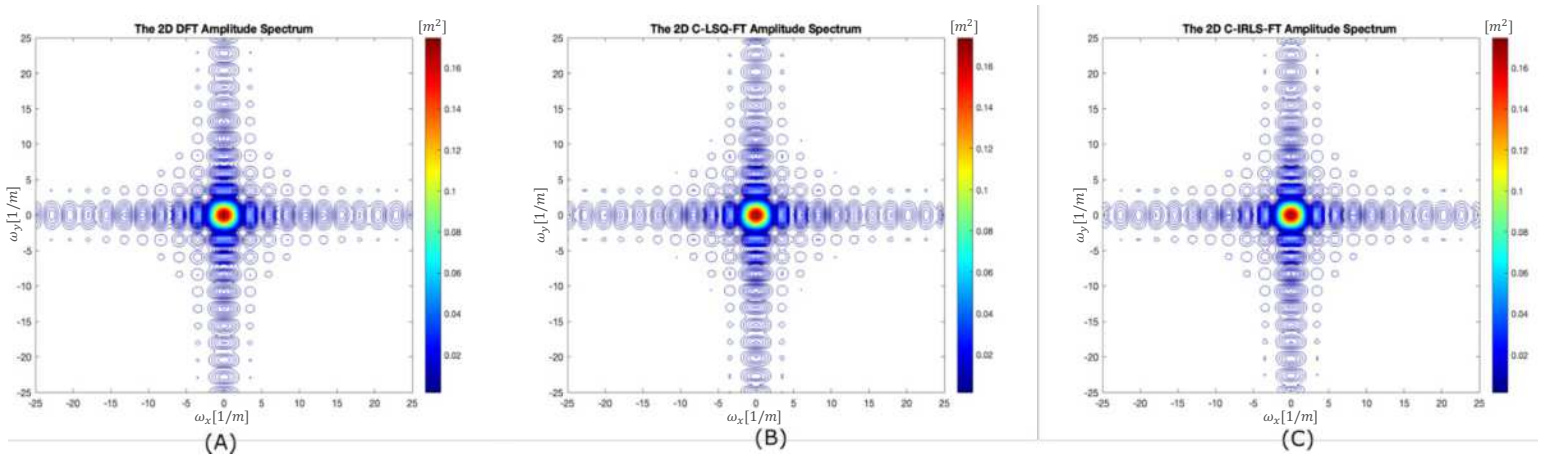

Figure 3-14: The 2D Amplitude Spectrum for noie-free data set (A) Using 2D DFT method (B) Using 2D C-LSQ-FT (C) Using 2D C-IRLS-FT

The efficacy of the three methods was confirmed when applied to the noise-free surface. To test real-world scenarios, the surface was contaminated with Gaussian and Cauchy noise as two separate scenarios, producing much rougher areas- as depicted in Figure (3-15) and Figure (3-17).

To accurately quantify the results, we propose the Root Mean Square (RMS) distance as a measure between the data sets (a) in the space domain:

$$D^{data} = \sqrt{\frac{1}{N}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left[u^{noisy}(x_i, y_j) - u^{noisefree}(x_i, y_j)\right]^2} \qquad (3\text{-}49)$$

where $N_x \; and \; N_y$ are the number of data sets on x and y axis, and (b) in (space) frequency domain:

$$D^{Spectrum} = \sqrt{\begin{aligned} &\frac{1}{N}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(Re\left[U^{noisy}(\omega_{xi},\omega_{yj})\right] - Re[U^{noisefree}(\omega_{xi},\omega_{yj})]\right)^2 \\ &+ \frac{1}{N}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(Im\left[U^{noisy}(\omega_{xi},\omega_{yj})\right] - Im[U^{noisefree}(\omega_{xi},\omega_{yj})]\right)^2 \end{aligned}} \qquad (3\text{-}50)$$

In the Gaussian noise scenario, shown in Figure (3-16), the data distance between the noise-free and noisy data sets is 0.0501. On the other hand, the model distance between the 2D DFT spectrum of the noisy and noise-free data sets is 0.0017. Moving to the introduced methods, the data distance using the 2D C-LSQ-FT method is 0.0176, and the model distance is 7.3347e-04. The 2D C-IRLS-FT showed similar results with a data distance equal to 0.0178; the model distance is 7.3272e-04. From above, we can summarize that the 2D C-LSQ-FT and 2D C-IRLS-FT demonstrate high noise reduction to random noise compared to the 2D DFT method. Also, both methods have similar outputs for amplitude spectrums. This result refers to the power of using Chebyshev polynomials as an alternative basis function in the inversion method.
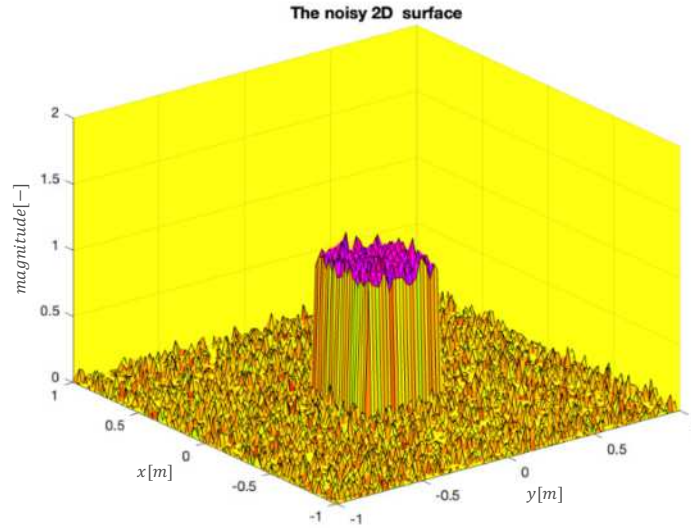

Figure 3-15: The Noisy data set contaminated with Gaussian noise
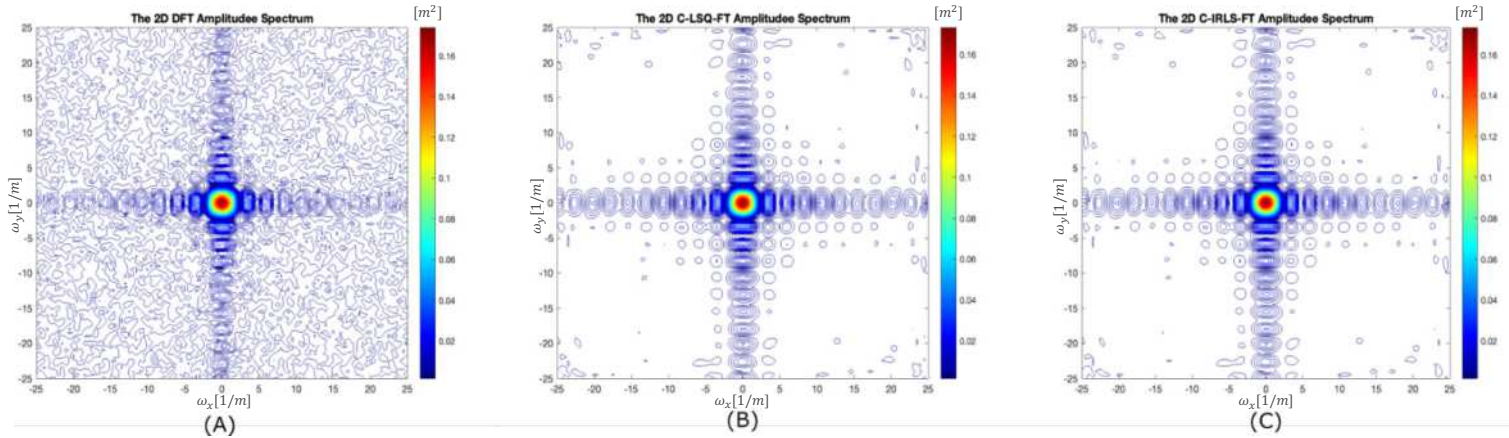
Figure 3-16: The 2D Amplitude Spectrum for noisy data set (Gaussian noise) (A) Using 2D DFT method (B) Using 2D C-LSQ-FT (C) Using 2D C-IRLS-FT

In the Cauchy noise scenario as shown in Figure (3-17), the addition of random Cauchy noise generated outliers with sharp spikes.
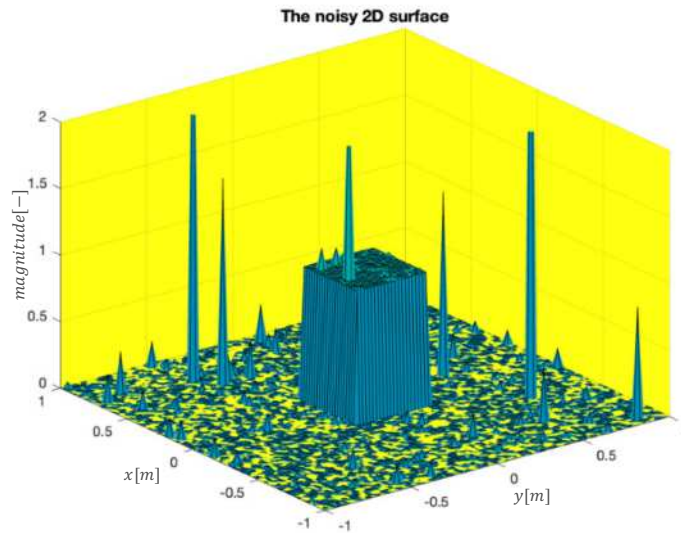

Figure 3-17: The Noisy data set contaminated with Cauchy noise

Figures (3-18) showcase the noise reduction capabilities of the 2D DFT, 2D C-LSQ-FT, and 2D C-IRLS-FT methods, through processing Cauchy's noisy data set. The output spectrums demonstrate that the 2D C-IRLS-FT method suppresses Cauchy noise considerably more effectively than the traditional 2D DFT method and 2D C-LSQ-FT. The 2D DFT method proves inadequate in eliminating the introduced noise, as evidenced by the spread of data noise in its output Fourier spectrums.

41

The data distance between the noise-free and noisy data sets is 0.0958. On the other hand, the model distance between the 2D DFT spectrum of the noisy and noise-free data sets is 0.0034. Moving to the introduced methods, the data distance using the 2D C-LSQ-FT method is 0.0253, and the model distance is 0.0009. The 2D C-IRLS-FT showed similar results with a data distance equal to 0.0161; the model distance is 6.6927e-04.

Upon examination of the data, it became evident that the 2D L-LSQ-FT and 2D C-IRLS-FT techniques offered superior noise reduction capabilities relative to the traditional 2D DFT approach. Nonetheless, it is essential to establish a more resilient and effective method for filtering out random noise and outliers, given the susceptibility of the Least Squares and 2D DFT methods. Thus, it is strongly advised to consider utilizing the 2D C-IRLS-FT method.
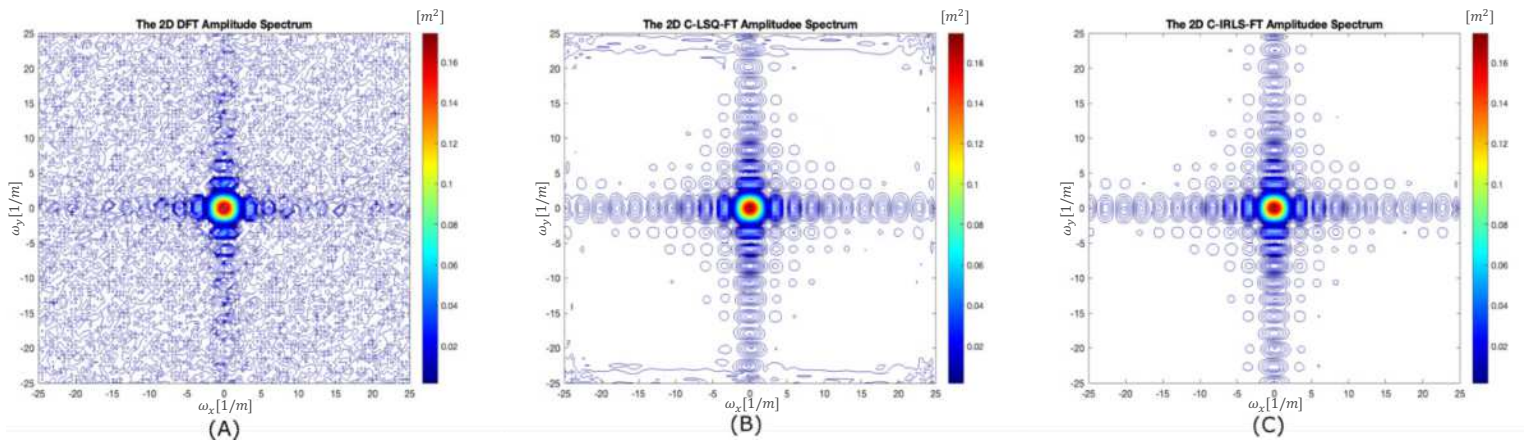

Figure 3-18: The 2D Amplitude Spectrum for noisy data set (Cauchy noise) (A) Using 2D DFT method (B) Using 2D C-LSQ-FT (C) Using 2D C-IRLS-FT

Thesis one:

I have developed an improved inversion-based Fourier transformation method with Chebyshev polynomials, as basis functions (1D C-IRLS-FT) using one-dimensional synthetic datasets. Using a time-domain 1D synthetic wavelet (contaminated with Gaussian- and Cauchy noises) a comparison is made between the proposed inversion-based technique and the conventional Discrete Fourier Transformation (DFT) method. The results indicate that the inversion-based Fourier transformation (1D C-IRLS-FT) method demonstrates the robustness and a significant

ability to mitigate the influence of outliers compared to the outcomes achieved using the conventional DFT method.

Thesis Two:

Owing to the effectiveness of the IRLS inversion technique in processing 1D datasets, the inversion-based Fourier transformation algorithm is further developed for the processing of 2D datasets with the use of the Chebyshev polynomials as a basis function (2D C-IRLS-FT method). To assess the accuracy, stability and outlier sensitivity, a 2D synthetic dataset (contaminated with random Gaussian and Cauchy noise) was ised. Upon data examination, it becomes clear that the 2D C-IRLS-FT technique demonstrates superior noise reduction capabilities in comparison to the traditional 2D DFT approach.

Chapter 4

4. HILBERT TRANSFORM USING C-IRLS-FT ALGORITHM

4.1. HILBERT TRANSFORM USING THE METHOD OF THE MOST FREQUENT VALUES

Attributes play an important role in seismic data processing and interpretation. They can be used to extract certain information, such as physical or geometric parameters, that may not be obtainable otherwise. Since the seminal paper published by Taner et al. in 1979, the field has expanded and undergone significant development. Today, we can discuss a wide range of attributes, which can be divided into physical and geometric categories, and classify procedures based on factors such as whether they are applied before or after processing, interpreted using one or more seismic channels, etc. When developing processing methods, it's important to manage and improve the signal-to-noise ratio as much as possible. Fourier transforms often play an important role in creating attributes mentioned in Péter Vass's Ph.D. thesis (Vass, 2010). Szegedi and Dobroka (2012) proposed a robust Fourier transform procedure (IRLS-FT) based on inversion. Our previous work showed that this method effectively suppresses outliers and can significantly improve the signal-to-noise ratio by up to an order of magnitude. In this chapter, we propose applying this procedure to computing attributes. First, we show how to create the Hilbert transform as part of a robust/resistant inversion framework, which plays a crucial role in defining the complex channel. It is considered a powerful tool in signal processing that allows for analyzing the phase and frequency components of a signal. First proposed by David Hilbert in 1917, the transform is used in a variety of fields, including telecommunications, control systems, and seismology.

4.2. THE ANALYTICAL SIGNAL

The starting point for calculating basic attribute stations is the creation of an analytical signal, also known as the complex channel. The concept of the analytical signal was introduced by Hungarian physicist Dénes Gabor, who won a Nobel Prize, in 1946. His goal was to use the powerful mathematical tools of quantum mechanics in signal processing, using square-integrable complex functions as elements of Hilbert space. To achieve this, he introduced the analytical signal.

$$s(t) = u(t) + ju_H(t) \qquad (4\text{-}1)$$

where

$$u_H(t) = \frac{1}{\pi} \int\limits_{-\infty}^{\infty} u(\tau) \frac{d\tau}{\tau - t} \qquad (4\text{-}2)$$

is the Hilbert -transform of the time signal. According to Eq. (4-2), the Hilbert transform is generated as a convolution of the time signal $u(t)$ with function $-\frac{1}{\pi t}$.

In the frequency domain, this relation can be written as:

$$\mathcal{F}\{u_H(t)\} = \mathcal{F}\{u(t)\}\mathcal{F}\left\{-\frac{1}{\pi t}\right\} \qquad (4\text{-}3)$$

where $\mathcal{F}$ denotes the Fourier transform. As $\mathcal{F}\left\{-\frac{1}{\pi t}\right\} = -j\,\mathrm{sign}(\omega)$, introducing the notation

$$U(\omega) = \mathcal{F}\{u(t)\} \qquad (4\text{-}4)$$

one can write

$$\mathcal{F}\{u_H(t)\} = -j\,\mathrm{sign}(\omega)U(\omega) = U_H(\omega) \qquad (4\text{-}5)$$

giving the Hilbert transform as

$$u_H(t) = \mathcal{F}^{-1}\{U_H(\omega)\} \qquad (4\text{-}6)$$

It can be seen that the calculation of the Hilbert transform requires the use of the Fourier transform and its inverse. The traditional DFT and IDFT algorithms are known to be sensitive to non-Gaussian noise in data sets. Dobroka et al. (2017) introduced an inversion-based Fourier transform method (IRLS-FT) to define a robust procedure, as we mentioned earlier in Chapter 3. To take advantage of its noise-rejection capacity, IRLS-FT can be used in calculating $U(\omega)$ as:

$$U(\omega) = \mathcal{F}_{IRLS}\{u(t)\} \qquad (4\text{-}7)$$

with IDFT in Eq. (4-6). A full inversion-based method can be produced when the $\mathcal{F}^{-1}_{IRLS}$ procedure is also used in Eq. (4-6).

4.3. The robust generation of Hilbert transform

To produce the Hilbert transform, we need to know the signal's U(ω) spectrum. To improve the signal-to-noise ratio, the Fourier transform is performed by the IRLS-FT method, and then the spectrum is multiplied by the function -j sign(ω). We then return to the time domain by an inverse Fourier transform. This can be done by applying IDFT and also by using an inversion-based inverse Fourier transform. The latter procedure can be performed by robust inversion defined here. The starting point is the expression of the Fourier transform

$$U(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t)e^{-j\omega t}dt \,. \tag{4-8}$$

The direct problem is given by the formula of the Fourier transform, where the time function u(t) is discretized in the form of a series expansion

$$u(t) = \sum_{n=1}^{M} B_n \Psi_n(t) \tag{4-9}$$

where B_n denotes the complex expansion coefficients, $\Psi_n(\omega)$ indicates the n-th known basis function, and M is the total number of the basis functions. After substitution, the following formula is obtained for the k-th sampling element of the spectrum

$$U_k(\omega_k) = \sum_{n=1}^{M} B_n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n(t)e^{-j\omega_k t}dt = \sum_{n=1}^{M} B_n G_{kn} \tag{4-10}$$

Where

$$G_{kn} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_n(t)e^{-j\omega_k t}dt = \mathcal{F}_k\{\Psi_n(t)\} \tag{4-11}$$

is the Jacobi matrix, the elements of which can be thought of as Fourier transforms of the basis function system. The calculation of the complex integral in the formula can be avoided by choosing the basic functions of series expansion from the eigenfunctions of the Fourier transform (Hermite

functions). In this chapter, we will use Chebyshev and Legendre polynomials as basis functions because, in the case of inverse problems, it is advised to use complete and orthogonal function systems to reduce the number of unknown parameters and improve the stability of the inversion procedure.

The Chebyshev polynomials were discussed early in Chapter 3. About the Legendre polynomials, we can demonstrate the orthogonality using the following equation that Nuamah et al. (2021) used in Fourier transformation for reduction to the pole of magnetic data (McCarthy et al., 1993):

$$\int_{-1}^{1} P_m(x)\, P_n(x)\, dx = 0 \quad if\ n \neq m \tag{4-12}$$

The Legendre differential equation is:

$$(1 - x^2)\frac{d^2 y}{dx^2} - \frac{2x\, dy}{dx} + n(n+1)y = 0 \tag{4-13}$$

where $n > 0, |x| < 1, or\ equivalently$

The solution can be written as a:

$$Y = AP_n(x) + B\mathbb{Q}_n(x) \tag{4-14}$$

The $P_n(x)\ and\ \mathbb{Q}_n(x)$ are Legendre functions of the first and second kinds. That means we can write for n order:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n \tag{4-15}$$

which can be written also in the recursive form:

$$(n+1)P'_{n+1}(x) = (2n+1)xP_n(x) - nP'_{n-1}(x) \tag{4-16}$$

4.4. USING HILBERT TRANSFORM TO GENERATE SEISMIC ATTRIBUTES

Many years ago, people commonly used seismic information to characterize emerging reservoirs. Seismic qualities help to find structural and stratigraphic information in three dimensions. Also, they can detect flaws and faults, as well as discernible geological lines, with characteristics of seismic data including amplitude, phase, and frequency. Getting instant and precise values for these criteria is crucial for finding hidden submerged features (Barnes, 2000).

Over the last three decades, many people have attempted to group various seismic attributes into specific categories to understand and apply them better. A 2002 version of this classification contains 153 lines, consisting of 69 text lines, a 34-line table, nine lines discussing coherence, 25 lines on hydrocarbon indicators, 16 lines on complex trace analysis, and six figures. In the study by Taner et al. (1994), attributes are classified as either physical or geometric. The former pertains to visualizing the continuity, azimuth, and dip of data collected through geometrical measurements. As for geometric attributes, they change the perceived visibility of data gathered through those measurements.

Norman Neidell originally proposed the Hilbert transform as a mean to achieve complex-trace analysis. Koehler and Taner worked with Neidell and developed a mathematical framework for attributes. The real part of the seismic signal's amplitude is represented by the magnitude of the complex analytical signal. The imaginary part is computed by taking a Hilbert transform of the analytical signal. Next, envelope, phase, and square root calculations are performed. These processes involve calculating the sum of the squares of the real and imaginary components. Three important attributes - the envelope, phase, and frequency - were established. The first one is the instantaneous envelope (reflection strength). It is sensitive to lithology, porosity, hydrocarbons, and thin-bed tuning. The second one is the instantaneous phase, which is useful for tracking reflector continuity, detecting unconformities, faults, and lateral changes in stratigraphy. The third one is the instantaneous frequency. It is useful in identifying abnormal attenuation and thin-bed tuning (Chopra and Marfurt, 2005).

4.5. NUMERICAL TESTS

The Ricker wavelet (Ricker, 1953) has been used to illustrate the response of a new method, as shown in Figure (4-1) in blue. This wavelet was generated at 10 Hz and is localized at 0.1 seconds, using a sampling interval of 5 milliseconds in the range of [-1, 1] seconds (Figure 4-1). Two types of noise were applied to the previous clean data set. As an example, gaussian noise with a standard deviation of 0.0025 was added to the first data set to simulate random noise from devices. On the other hand, Cauchy noise with a scale parameter of $\varepsilon=0.04$ was applied to the second data set to simulate outliers in measured data (Figure 4-2).

4.5.1. CHEBYSHEV POLYNOMIALS AS A BASIS FUNCTION

The first attribute that can be produced is the reflection strength (Instantaneous amplitude), which is considered an absolute signal value. We can represent it with the following equation:

$$A(t) = \sqrt{u(t)^2 + u_H(t)^2} \qquad (4\text{-}17)$$

The noise sensitivity of this attribute can be illustrated in Figure (4-3) using the previous wavelet. The reflection strength calculated by the conventional process for noise-free inputs contaminated Gaussian and Cauchy noise at the same time.
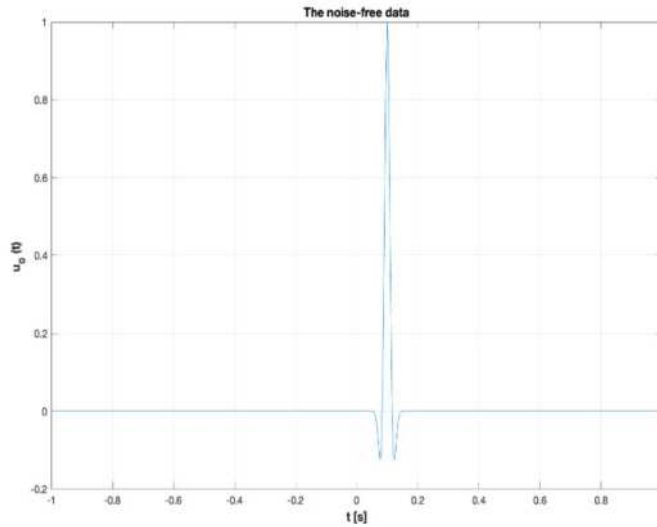
Figure 4-1:The data set without noise
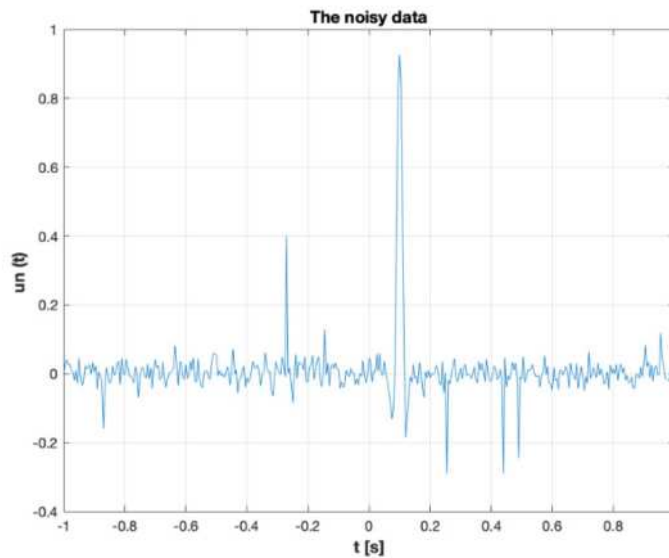
Figure 4-2: The data set with Cauchy and Gaussian noise
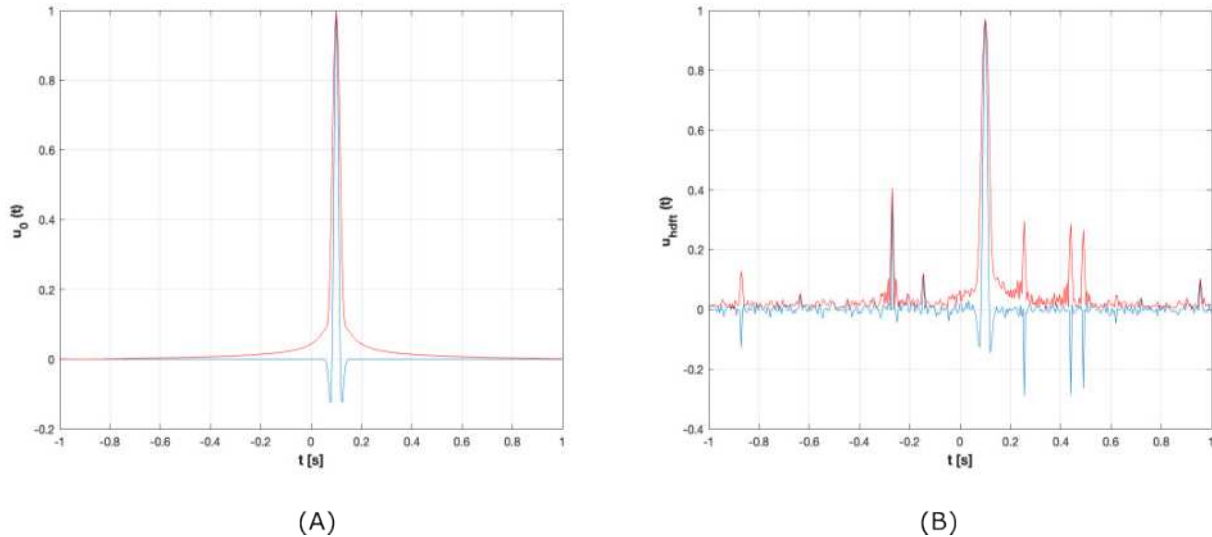
Figure 4-3: The Ricker wavelet (blue) and reflection strength (red) where (A) Noise-free data set and (B) Noisy data set. The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.

The second attribute, which is the instantaneous phase, can be given by:

$$\emptyset(t) = arc\tan|u_H(t)/u(t)| \qquad \text{(4-18)}$$

where u(t) is the signal (Ricker wavelet) and u_H (t) is the Hilbert transform which those related to the envelope E(t) and the phase $\emptyset(t)$ by the following equations:

$$u(t) = E(t)\cos(\Phi(t)) \qquad \text{(4-19)}$$

$$u_H(t) = E(t)\sin(\Phi(t)) \qquad \text{(4-20)}$$

The instantaneous phase is measured in degrees (-180, 180). It is independent of amplitude and shows continuity and discontinuity of events. It shows bedding very well. Phase along horizon should not change in principle. Changes can arise if there is a picking problem or if the layer changes laterally due to "sinkholes" or other phenomena. The noise sensitivity of this attribute can be illustrated in Figure (4-4) using the previous wavelet. The reflection strength calculated by the conventional process for noise-free inputs contaminated Gaussian and Cauchy noise at the same time.
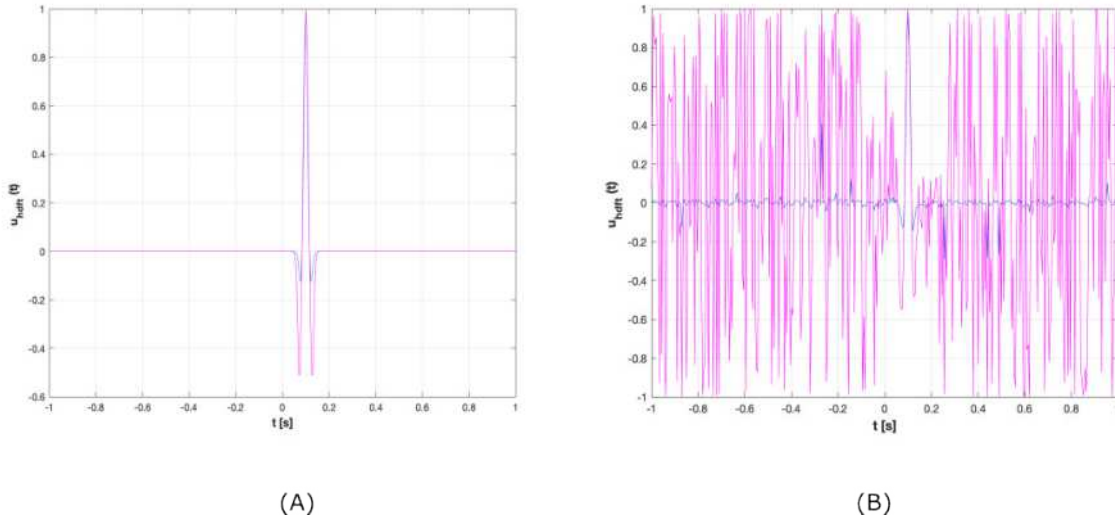
Figure 4-4: The Ricker wavelet (blue) and instantaneous phase (magenta) where (A) Noise-free data set and (B) Noisy data set. The discrete Fourier transform method (DFT) and its inverse (IDFT) were used.

The newly introduced algorithm was applied to the same data set, as shown in Figure (4-5). A new method has been developed that effectively reduces noise in a given noisy data set on the first seismic attribute (Reflection strength) Figure (4-5B) compared to the old one Figure (4-5A). The previous method relied on traditional DFT and signal processing techniques that were ineffective in removing noise from the signal.
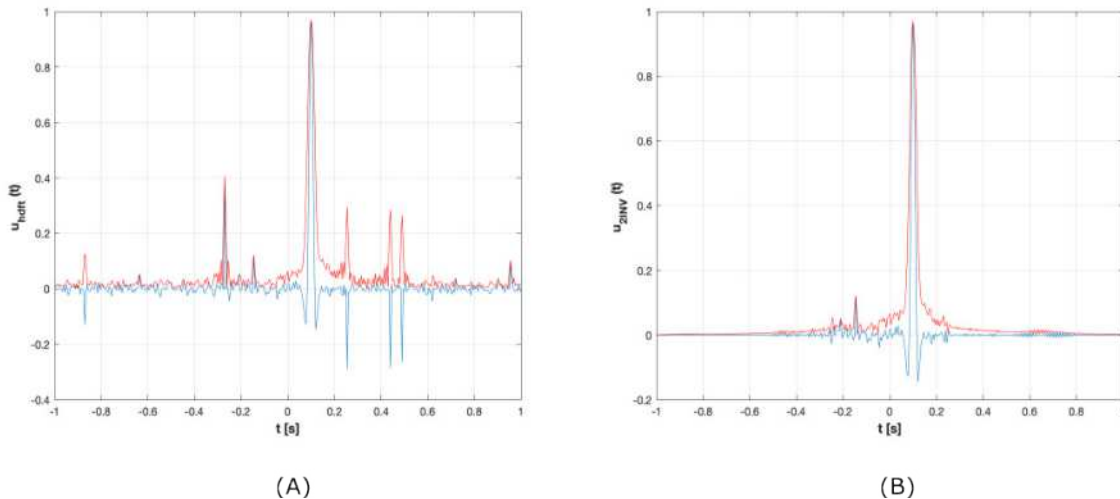
Figure 4-5: The Ricker wavelet (blue) and reflection strength (red) where (A) DFT and IDFT used on noisy data set and (B) The newly introduced method was used on the noisy data set.

Also, the second attribute (Instantaneous phase) was tested on the same noisy data set, as shown in Figure (4-6). This attribute is very sensitive to noise compared to the previous one. Therefore,

we can conclude that the new method effectively reduces noise in a given noisy data set, as seen in Figure (4-6B), in contrast to the transitional process in Figure (4-6A).

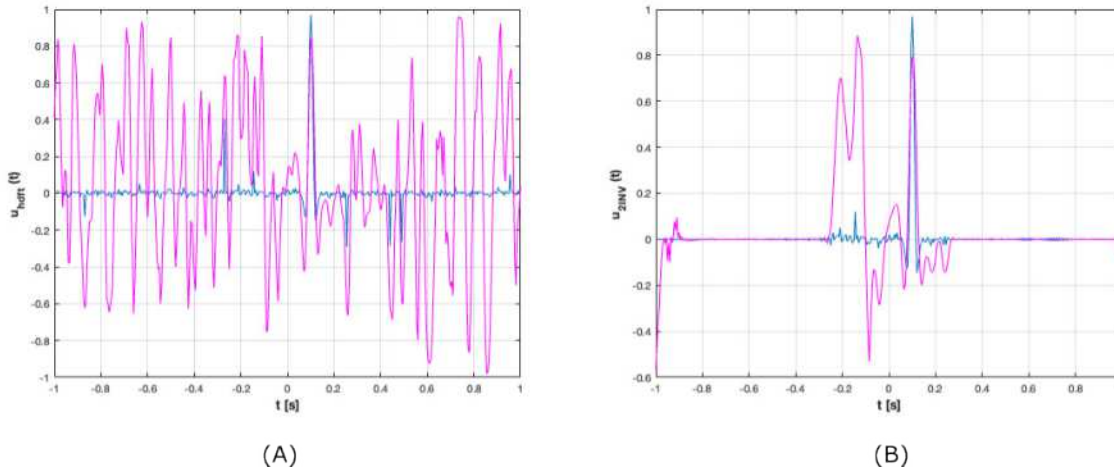(A)   (B)

Figure 4-6: The Ricker wavelet (blue) and instantaneous phase (magenta) where (A) Noisy data set and (B) The newly introduced method was used on the noisy data set.

One common way to measure the accuracy of an introduced algorithm is by comparing its output to the original data set (Noise-free data set) by measuring the data distance as we performed in Chapter 3. In the context of our study, when comparing two methods or approaches in data analysis, it is important to consider the data distance between them, as we discussed earlier. Data distance refers to the degree of difference or similarity between the data generated by each method. Understanding the data distance between two methods is important in determining which method is more appropriate for a particular research question or problem. On the other hand, if the data distance between the two methods is large, it may suggest that they are measuring different aspects of the same phenomenon but with different accuracy. Therefore, a thorough understanding of the data distance between the two methods is essential for making informed decisions about data analysis and interpretation.

The newly developed method called C-IRLS-HT is more effective in eliminating noise than the traditional Hilbert transform method in two seismic attributes, as evidenced by the lower data distance values obtained through its use. In the first attribute (Reflection Strength), when Method C-IRLS-HT was applied to a dataset with high levels of generated noise, it was found to reduce the error with D = 0.009, resulting in more accurate and reliable data. Conversely, when the traditional Hilbert transform was applied to the same dataset, the error reduction was demonstrated

with D = 0.045. Similarly, in the second attribute (Instantaneous Phase), where both methods were applied to a dataset with the same noise, Method C-IRLS-HT was again found to be more effective in reducing the error despite the high sensitivity of the attribute, resulting in data distance D = 0.160 compared to a 0.435 with traditional Hilbert transform. These findings suggest that the C-IRLS-HT is a more robust and reliable method for data analysis in situations where noise reduction is very important.

4.5.2. LEGENDRE POLYNOMIALS AS A BASIS FUNCTION

Another introduced algorithm that uses Legendre polynomials was applied to the same data set, as shown in Figure (4-7). Also, this method effectively reduces noise in a given noisy data set on the first seismic attribute (Reflection strength) Figure (4-7B) compared to the conventional method Figure (4-7A). The conventional method relied on traditional DFT and signal processing techniques that were ineffective in removing noise from the signal.

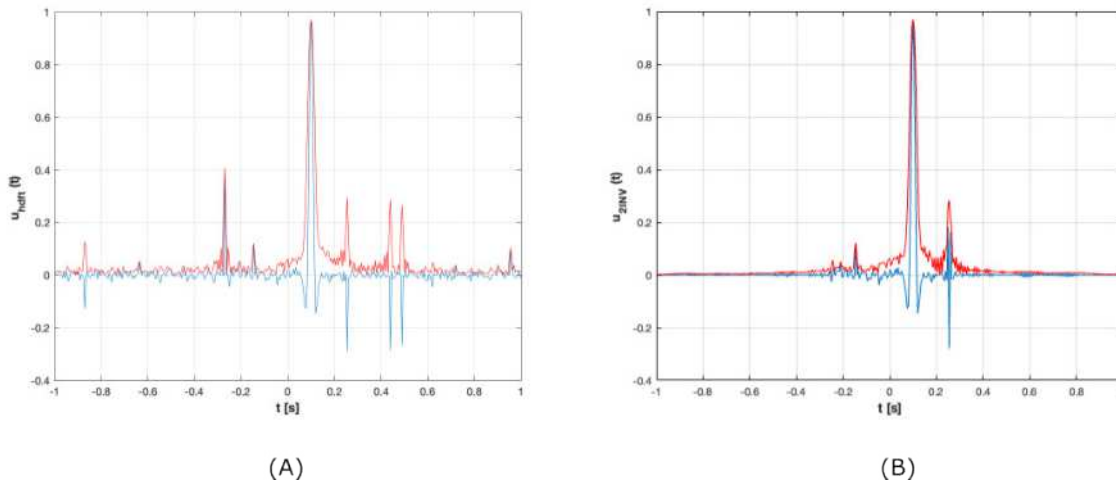(A)   (B)

Figure 4-7: The Ricker wavelet (blue) and reflection strength (red) where (A) Noisy data set and (B) The second introduced method was used on the noisy data set (Using Legendre Polynomials).

Also, the second attribute (Instantaneous phase) was tested on the same noisy data set, as shown in Figure (4-8). Like the previous case, this attribute is more noise-sensitive than the reflection strength. Therefore, we can conclude that also this method effectively reduces noise in a given noisy data set, as seen in Figure (4-8B), in contrast to the conventional method in Figure (4-8A).
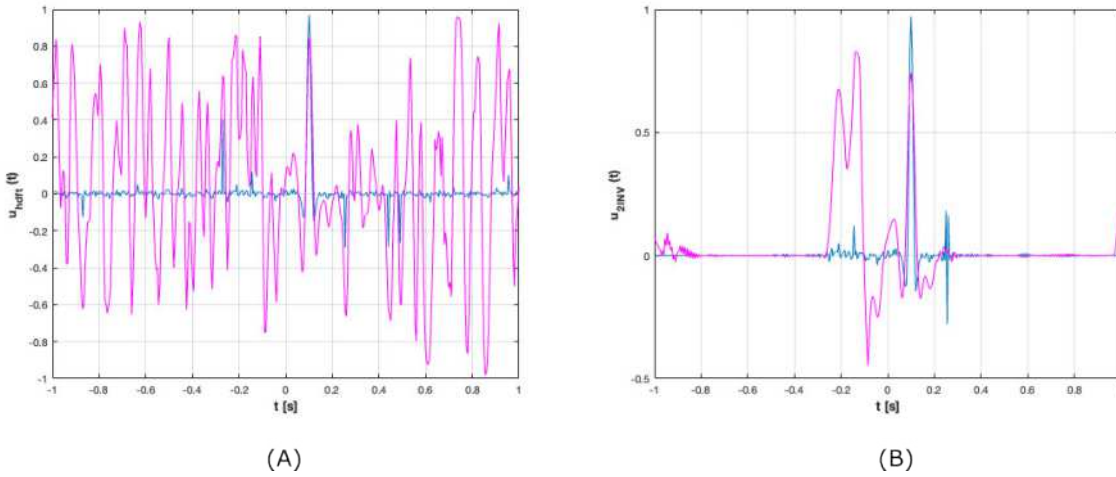
Figure 4-8: The Ricker wavelet (blue) and instantaneous phase (magenta) where (A) Noisy data set and (B) The second introduced method was used on the noisy data set (Using Legendre Polynomials).

Also, Method L-IRLS-HT has been evaluated for its ability to eliminate noise from data using the data distance analytic method. In the first attribute (Reflection Strength), where a dataset with high levels of generated noise was analyzed. Method L-IRLS-HT was found to reduce the error with data distance $D = 0.020$, while Method traditional Hilbert transform method as we saw before only reduced the error by $D = 0.045$. Similarly, the second attribute (Instantaneous Phase) was analyzed with the same noisy data set. The L-IRLS-HT reduced the error by $D = 0.180$, while the traditional Hilbert transform shows that $D = 0.435$ as before without any change. These results also suggest that Method L-IRLS-HT as C-IRLS-HT may be more effective than the traditional Hilbert transform in eliminating noise from data. The comparison between the newly developed C-IRLS-HT and L-IRLS-HT methods shows the superiority of C-IRLS-HT. However, again, the effectiveness of these methods may depend on the specific characteristics of the dataset.

4.5.3. THE COMPARISON BETWEEN LEGENDRE and CHEBYCHV POLYNOMIALS AS BASIS FUNCTIONS FOR HILBERT TRANSFORM

There are various methods available for solving numerical problems, each with its own unique set of advantages and disadvantages. In this section, we will compare two presented methods, Hilbert

transforms using Chebyshev and Legendre polynomials with IRLS, based on their noise reduction ability, speed, iteration, and the number of unknowns.

Regarding noise reduction, the Hilbert transform using the Chebyshev polynomials method (C-IRLS-HT) has the edge over Hilbert transform using Legendre polynomials (L-IRLS-HT). This difference regards C-IRLS-HT utilizing a better filtering technique due to the Chebyshev polynomials of the second kind that helps to eliminate any unwanted noise in the system, making it particularly useful in applications where the data is subject to interference or distortion compared to L-IRLS-HT.
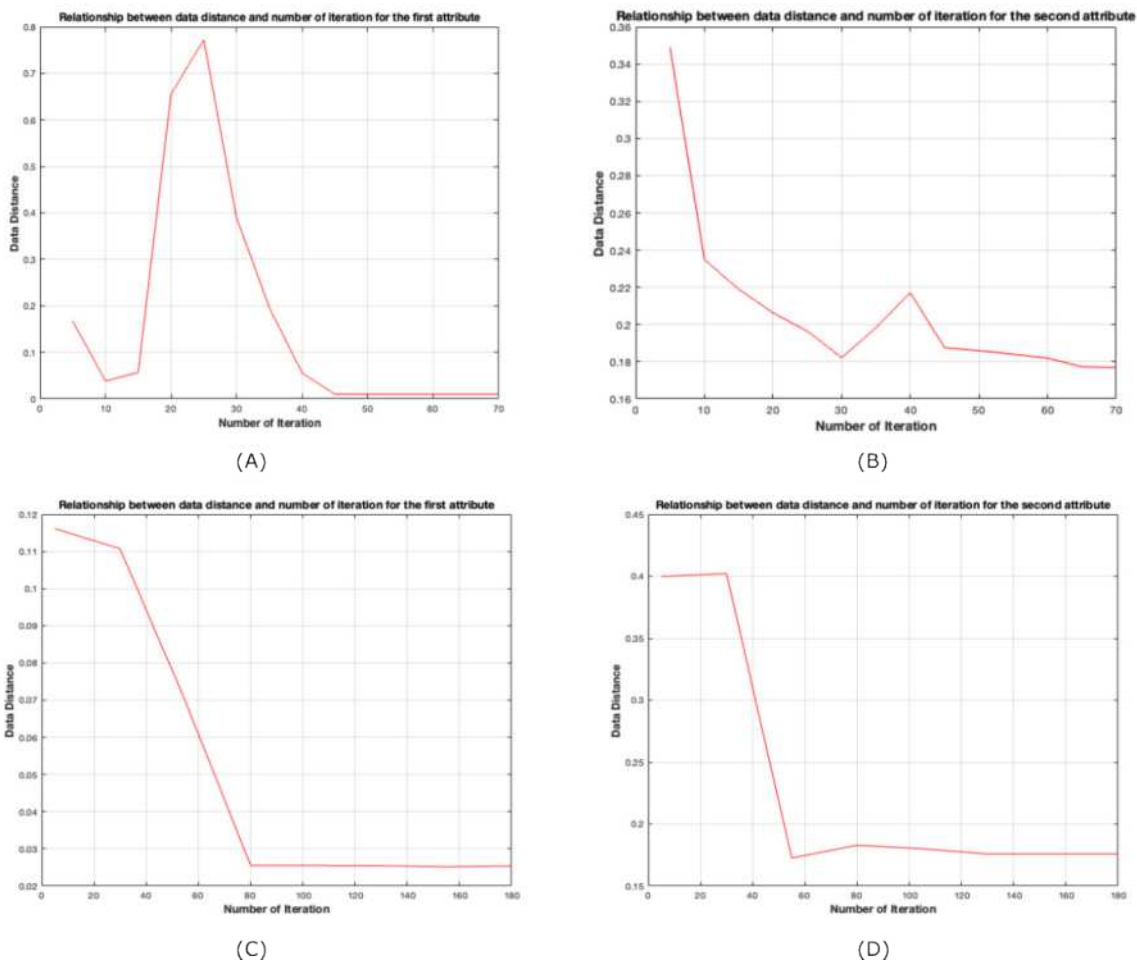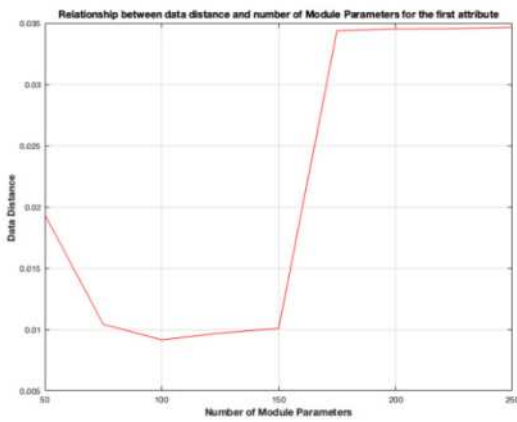

(A)

(B)

(C)

(D)

Figure 4-9: Comparison between the number of iterations and the data distance between the noise-free attribute and noisy attribute A-Using C-IRLS-HT for reflection coefficient B- Using C-IRLS-HT for instantaneous phase C- Using L-IRLS-HT for reflection coefficient D- Using L-IRLS-HT for instantaneous phase

When it comes to iteration, both methods are capable of handling a significant number of iterations. However, method L-IRLS-HT tends to require more iterations (100 iterations) Figure (4-9) C&D

to converge on a solution, mainly when dealing with more complex problems. Method C-IRLS-HT, on the other hand, is typically more efficient and can often converge on a solution with fewer iterations (50 in numerical tests) Figure (4-9) A&B.

The number of unknowns (M)that each method can handle is an important consideration. Both methods should have suitable values that do not distort the signal and, at the same time, eliminate the noise. Figure (4-10) demonstrates that a value of 150 is the most convenient value for both methods without distorting the data. If we choose a value less than 100, it may corrupt the signal, and if we choose a value higher than 150, it may also introduce noise.
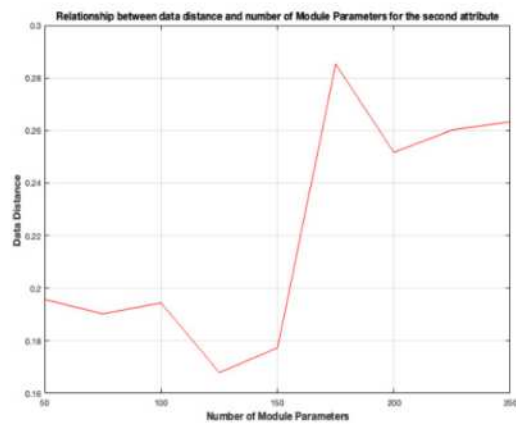

(A)
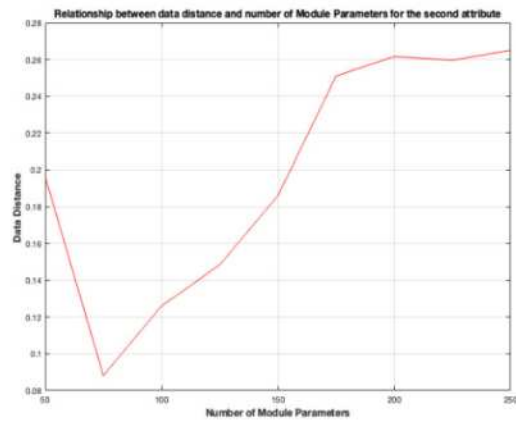
(B)
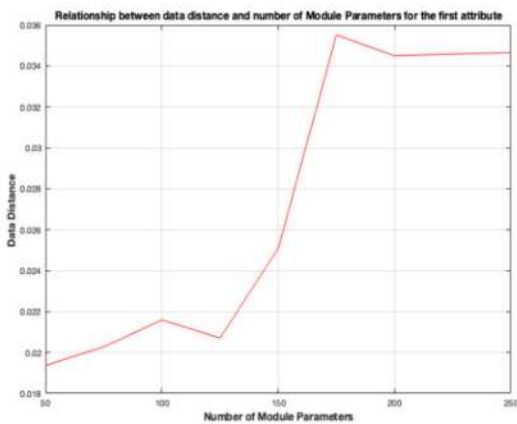
(C)

(D)

Figure 4-10: Comparison between the Model Parameters and the data distance between the noise-free attribute and noisy attribute A-Using C-IRLS-HT for reflection coefficient B- Using C-IRLS-HT for instantaneous phase C- Using L-IRLS-HT for reflection coefficient D- Using L-IRLS-HT for instantaneous phase

Finally, in terms of speed, method C-IRLS-HT has a clear advantage over method L-IRLS-HT. This is because method C-IRLS-HT requires fewer iteration steps. However, in modern computing systems, the difference in time can be negligible if we use powerful CPUs or GPUs. In summary, both methods C-IRLS-HT and L-IRLS-HT have the ability to reduce noise better than conventional methods, taking into consideration the sensitivity of the second attribute that plays a significant role in judging the results.

Thesis Three:

We further explored the application of the inversion-based Fourier transformation procedure to computing seismic attributes. We have demonstrated that the newly developed method, referred to as (1D) C-IRLS-HT, outperforms the traditional Hilbert transform method in reducing noise for two seismic attributes. Specifically, when applied to process noisy Reflection Coefficient (first attribute) and the Instantaneous Phase (second attribute) data, the Method C-IRLS-HT demonstrated a significant noise reduction, resulting in more accurate and reliable data compared to the traditional Hilbert transform method. These findings underscore the effectiveness of the C-IRLS-HT procedure as a more robust and reliable method for data analysis in situations where noise reduction is essential.

Chapter 5

5. 2D HILBERT TRANSFORM WITH IRLS-FT

5.1. 2D HILBERT TRANSFORM

The two-dimensional (2D) Hilbert Transform, a natural extension of its one-dimensional counterpart, has proven to be a powerful and versatile tool in the domain of image processing. By adopting the Hilbert Transform to handle 2D signals, researchers have been able to efficiently analyze and process images to extract spatial information and identify features that may be obscure to traditional methods. This chapter will focus solely on the 2D Hilbert Transform, shedding light on its theoretical foundations, and contemporary developments within the realm of earth science-related image processing (edge detection) from the point of view of noise reduction and image enhancement.

We begin by elaborating on the mathematical formulation of the 2D Hilbert Transform, highlighting its derivation and inherent properties. The 2D Hilbert Transform can be understood as a convolution operation involving the image and a 2D Hilbert kernel. The kernel itself is frequently derived from the product of two 1D Hilbert kernels, with one acting along the rows and the other along the columns of the image. This convolution can be extended to a 2D analytic signal that encapsulates both the original image data and its Hilbert Transform, providing valuable insights into the image's spatial frequency content and phase information.

When applied to a square wave signal, as Figure (5-1) shows, the Hilbert transform accentuates the abrupt changes in amplitude. This observation implies to use the 2D Hilbert transform for edge detection of 2D objects. The edge transitions can be determined by examining also the instantaneous amplitude and phase of the analytic signal. The instantaneous amplitude is given by the magnitude of the analytic signal, while the instantaneous phase is the arctangent of the imaginary part divided by the real part, as we saw before.

In the case of a square wave $u_0(t)$, the $u_H(t)$ Hilbert transform is shown in Fig 1a. and the instantaneous amplitude $u_A(t)$ is depicted in Fig. 1b. Both operations exhibit peaks at the edge transitions. By analyzing these peaks and changes, the precise location of the edge transitions in

the square wave signal can be determined. This information is precious for applications that require mostly the identification of boundaries or changes in one-dimensional signals.
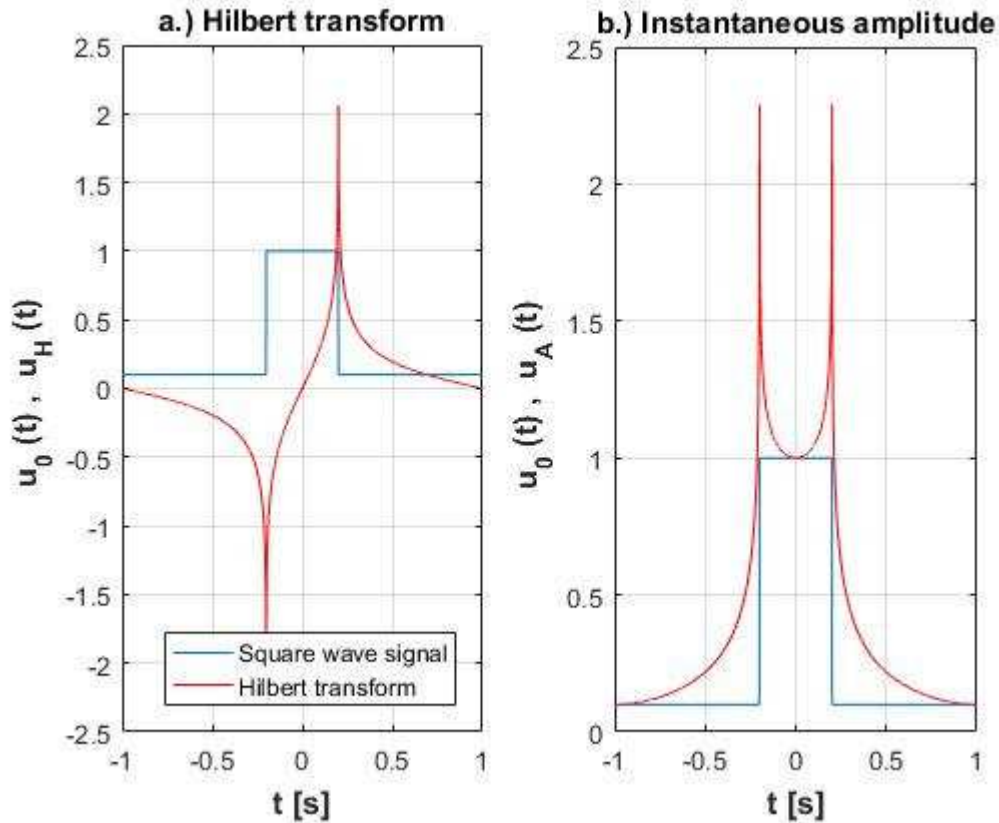

Figure 5-1: a.) The square wave signal and its Hilbert transform, b.) The square wave signal (blue) and Instantaneous amplitude (red)

Edge detection is an important operation in image processing. The most frequently used tools for it are the Prewitt and Sobel operators performing difference calculations along the x- or y-directions of the image. As is well known, differentiation is highly sensitive to noises. To attenuate the effect the difference is multiply calculated (due to the size of the convolution mask) and the arithmetic mean (Prewitt operator) or the binomial mean (Sobel operator) of the differences is applied for the central pixel of the image section covered by the mask. As is shown above, the calculation of the Hilbert transform needs integration, which is less sensitive to the noises. So it can be expected to find less noise-sensitive edge detection results by using Hilbert transform (Pei and Ding, 2003).

One of the most compelling applications of the 2D Hilbert Transform lies in its ability to perform noise reduction in addition to edge feature extraction. By analyzing the local phase information, the transform can accurately discern edges, corners, and other salient features in an image, often outperforming traditional edge detection techniques. Furthermore, the 2D Hilbert Transform enables the decomposition of an image into its constituent amplitude and phase components, which can be invaluable for tasks such as image enhancement, denoising, and compression (Havlicek et al., 2002).

In addition to edge detection, the 2D Hilbert Transform has found utility in texture analysis, where it has been employed to characterize and discriminate between different textures within an image. By examining the local frequency content and phase relationships, the 2D Hilbert Transform can effectively isolate and identify various textural patterns, providing valuable information for tasks such as segmentation, classification, and feature extraction (Guanlei et al., 2018).

Moreover, the 2D Hilbert Transform has been utilized in image reconstruction, particularly in the context of medical imaging and synthetic aperture radar (SAR) systems. By exploiting the phase information provided by the transform, it is possible to reconstruct high-resolution images from limited or incomplete data, leading to improved diagnostics and decision-making. However, despite its current limitations, the 2D Hilbert Transform remains a cornerstone in image processing, with its unique capabilities continuing to inspire novel applications and advancements in the field (Moon et al., 1988).

The main aim that will focus on in this chapter further improve the efficacy of the Hilbert transform in edge detection by reducing its outlier sensitivity. In addition, this approach will allow us to enhance the processing and interpretation phases for more precise output and accurate information. This chapter will further discuss various computational approaches for implementing the 2D Hilbert transformation with our newly developed inversion-based C-IRLS-FT robust Fourier transformation method.

5.2. THE CONVENTIONAL 2D HILBERT TRANSFORM

The one-dimensional (1D) Hilbert Transform (HT) and Analytic Signal have played a significant role in signal processing since their introduction by Gabor (1946). The analytic signal is obtained

by eliminating negative frequency components from a signal's Fourier Transform, as we saw in Chapter 4, resulting in the following:

$$s(t) = u(t) + ju_H(t) \tag{5-1}$$

where $u_H(t)$ denotes the signal's Hilbert Transform, acquired through convolution:

$$u_H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} u(\tau) \frac{d\tau}{\tau - t} \tag{5-2}$$

The Analytic Signal allows for the calculation of instantaneous amplitude and phase as the modulus and angle of the complex function Eq (5-1). Numerous applications have been explored, including time-frequency signal analysis (Cohen, 1995). However, extending HT and AS concepts to two-dimensional (2D) cases and their application to images has been challenging due to the non-uniqueness of the multidimensional HT, leading to various approaches.

Sommer et al. (1997) emphasized that three properties of 1D-AS must be satisfied also in 2D cases:

- the first one is the AS spectrum is causal: $S_A(f) = 0 \; \forall f < 0$ (the function or property $S_A(f)$ equals zero for all negative frequencies.)
- the second one is the mention that the original signal can be recovered from its associated Analytic Signal, with the real part of the AS equaling the original signal
- the last one indicates that the envelope of an actual signal is acquired as its associated AS's magnitude, representing the instantaneous amplitude.

To satisfy these conditions in 2D cases, several approaches have been proposed. Simultaneously, interesting applications for 2D-HT and the corresponding AS have been developed or identified, including corner detection (Kohlmann, 1996), AM-FM image models (Havlicek et al., 1998), phase congruency calculations (Kovesi and others, 1999), and edge detection (Pei and Ding, 2003), among others. In the subsequent discussion, $\mathbf{x} = (x,y)$ represents the spatial domain, and $\mathbf{u} = (u,v)$ is the 2D Fourier-transformed domain. The term s(x) denotes a real image, and S(u) is its 2D Fourier transform. The double asterisk symbolizes a 2D convolution, and sgn(·) represents the signum function, defined in 2D and oriented in the x direction as:

$$sgn(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases}$$

A signum oriented in the y direction has a notation that bears some resemblance. While the idea of an analytic signal does not apply to discrete signals as the ones analyzed in this study, it is still commonly referred to as such in the relevant literature.

The 2D Hilbert Transform (2D-HT) is an extension of the 1D HT and is defined for a two-dimensional function s(x, y) as follows ($s_H(x)$ represent the 2D-HT in time domain (Lorenzo-Ginori, 2007)):

$$s_H(x) = s(x) ** \frac{1}{\pi^2 xy} \tag{5-3}$$

For the frequency domain, the 2D Hilbert transform is:

$$S_H(u) = H(u)S(u) \tag{5-4}$$

where

$$H(u, v) = -sgn(u)sgn(v)$$

is the 2D mask of the Hilbert transform. The 2D Analytic Signal from Eq(5-1):

$$S_A(u) = [1 - j \, sgn(u)sgn(v)]S(u). \tag{5-5}$$

By multiplying the transformed image with a suitable mask, Equations (5-4) and (5-5) yield the 2D-Hilbert Transform or the 2D-Analytic Signal in the frequency domain. The Hilbert transform can be calculated by using the inverse Fourier transformation:

$$s_H(x) = \mathcal{F}^{-1}\{S_H(u)\} \tag{5-6}$$

or in another form

$$s_H(x) = \mathcal{F}^{-1}\{H(u)F\{s(x)\}\} \tag{5-7}$$

In calculating the discrete 2D Hilbert transform the 2D DFT (or 2D FFT) and its inverse, the 2D IDFT (or IFFT) is used resulting in the well-known outlier sensitivity.

5.3. THE ROBUST 2D HILBERT TRANSFORMS C-IRLS-HT AND H-IRLS-HT

Due to using DFT and IDFT, the conventional algorithm of the 2D Hilbert transform given in Equation (5-7) is sensitive to non-Gaussian noises, especially outliers. To make the 2D Hilbert transform more robust we apply our newly developed 2D C-IRLS-FT method and its inverse, the

C-IRLS-FT^{-1}. The inversion-based 1D inverse Fourier transformation was introduced in Chapter 4, its straightforward extension to the 2D case is performed and implemented in our Matlab software. So the robustified 2D Hilbert transform is introduced as

$$s_H(\boldsymbol{x}) = C\_IRLS\_FT^{-1}\{H(\boldsymbol{u})C\_IRLS\_FT\{s(\boldsymbol{x})\}\} \qquad (5\text{-}8)$$

and will be called 2D C-IRLS-HT.

In this section, we introduce also the robust 2D Hilbert transform method based on a previously developed H-IRLS-FT procedure (Dobróka et al., 2012), in which the spectrum is discretized by using the Hermite functions. The new robustified 2D Hilbert transform is introduced as

$$s_H(\boldsymbol{x}) = H\_IRLS\_FT^{-1}\{H(\boldsymbol{u})H\_IRLS\_FT\{s(\boldsymbol{x})\}\}$$

and will be called 2D H-IRLS-HT. In the following, the noise- and outlier sensitivity of the conventional and the robust Hilbert transform methods will be analyzed using synthetic 2D models.

5.4. NUMERICAL INVESTIGATION

In this study, we propose the development of a two-dimensional synthetic model with the size [-1,1] in both x- and y direction (Figure 5-2). The test area is divided into 101*101 cells with a 20*20 cells anomaly in the middle (having a unit value in the zero-valued background). The model's intricate design emulates a very noisy scenario, enabling a comprehensive understanding of the interaction between noisy and real data within the model. This meticulous evaluation will not only provide insights into the method's potential limitations but also elucidate opportunities for optimization.

Ultimately, our synthetic model serves as a potential case for noise reduction approaches using the 2D Hilbert transform with the inversion-based IRLS-FT method (2D C-IRLS-HT, 2D H-IRLS-HT) and refining our understanding of complex phenomena in a controlled academic setting.
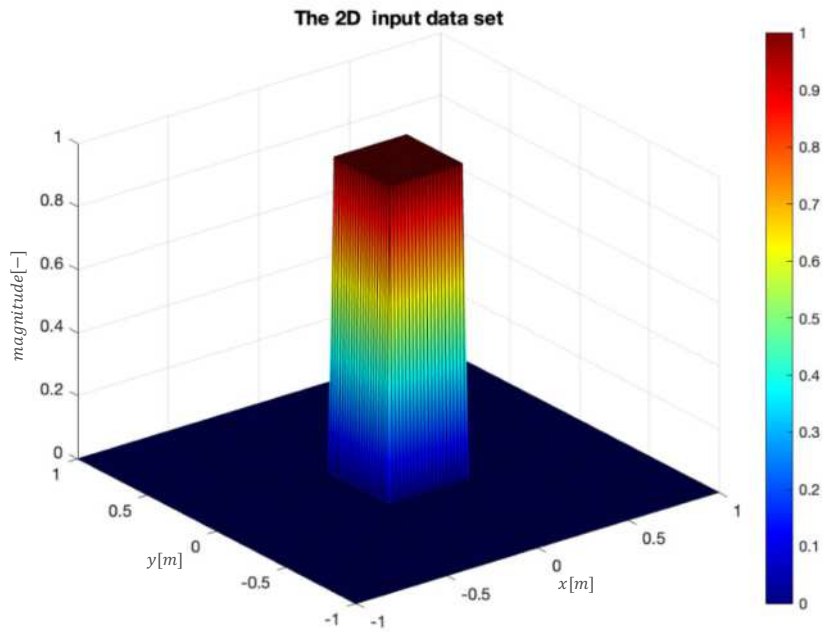
Figure 5-2: The Generated 2D input noise free data set

In the initial phase of our investigation, the conventional 2D Hilbert Transform method using DFT was employed on synthetic data devoid of noise to demonstrate its capacity to delineate the boundaries of anomalies within the dataset, as Figure (5-3) shows. The absence of noise in this controlled environment facilitated the unambiguous identification of the anomaly borders to
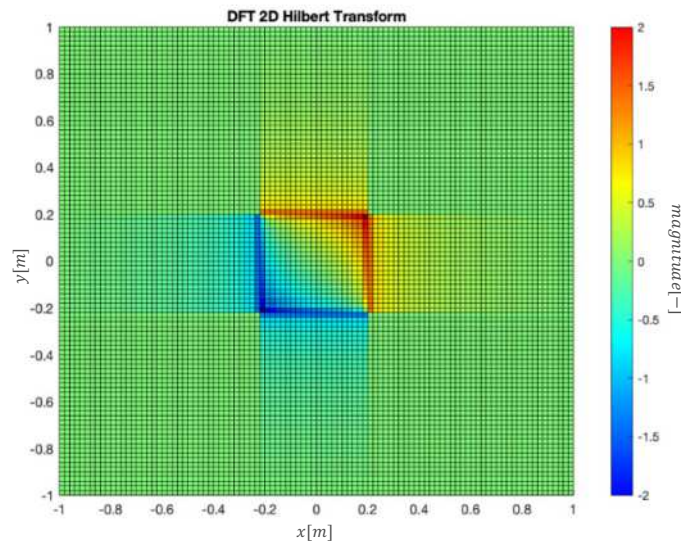

Figure 5-3: The generated conventional 2D Hilbert Transform on
noise free synthetic data

compare it with the noisy one. By using a noise-free dataset, we aimed to establish a baseline performance metric to ascertain the effectiveness of the 2D Hilbert transforms using the IRLS method. This step was important in elucidating the underlying principles and potential limitations of the technique, which would subsequently inform the analysis of more complex, noise-affected data. By applying the 2D DFT, we aimed to accentuate the importance of frequency domain analysis in conjunction with the 2D Hilbert transform for a comprehensive understanding of the data. Consequently, the incorporation of 2D DFT analysis bolsters the robustness of the introduced method in comparison and contributes significantly to the ongoing development of advanced techniques for anomaly detection. We implemented the introduced 2D C-IRLS-HT, and the 2D H-IRLS-HT methods in a Matlab environment and tested them. Using noise-free dataset the 2D IRLS-HT methods give the result shown in Figure (5-4). As it can be seen, the two results are similar with slightly better resolution of the H-IRLS-HT procedure. Because of the fact that the Hermite functions are eigen functions of the Fourier transform, the H-IRLS-HT method is more economic (the computation time is much smaller). So, in our later investigations we use the Hermite function-based method with the note, that the 2D C-IRLS-HT procedure gives similar results.
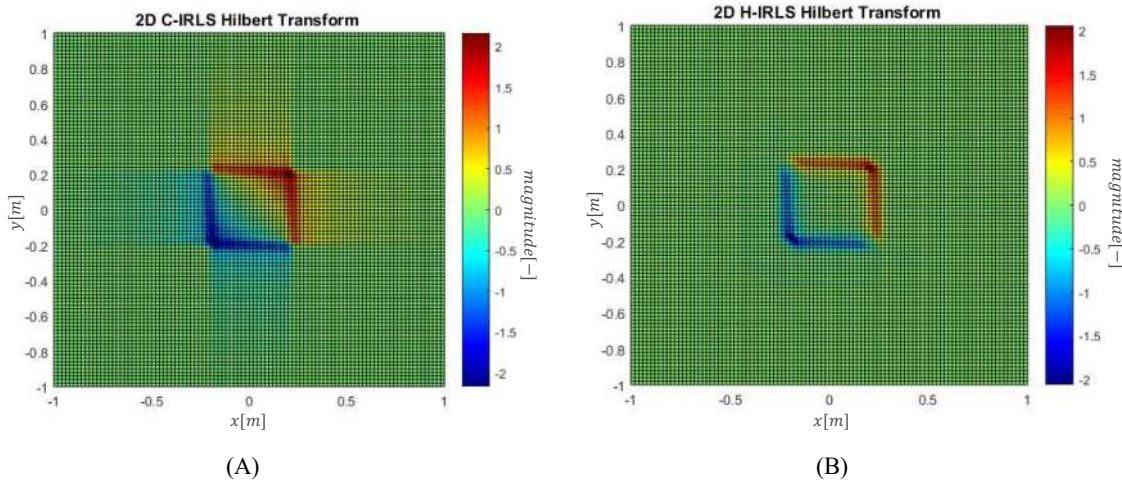

(A)   (B)

Figure 5-4. The 2D IRLS Hilbert transforms A) C-IRLS-HT in 2D view B) H-IRLS-HT in 2D view

As a first step, we show the response of 2D H-IRLS-HT when a 1D Hilbert kernel is applied to the 2D model. Specifically focusing on the X-axis (with the kernel $H(u, v) = jsgn(u)$) the result shown in Figure (5-5) can be found, which are the change of the edges along the X- axes. Similarly,

using Hilbert transform with the kernel $H(u,v) = jsgn(v)$ the change of the edges along the Y-axes are detected (Figure 5-6). In a subsequent step, we used the complete Hilbert mask $H(u,v) = -sgn(u)sgn(v)$ in our 2D H-IRLS-HT method to find the full set of edges. The result is shown in Figure (5-7), which is in complete agreement with the conventional (DFT-based) view shown in Figure (5-3).
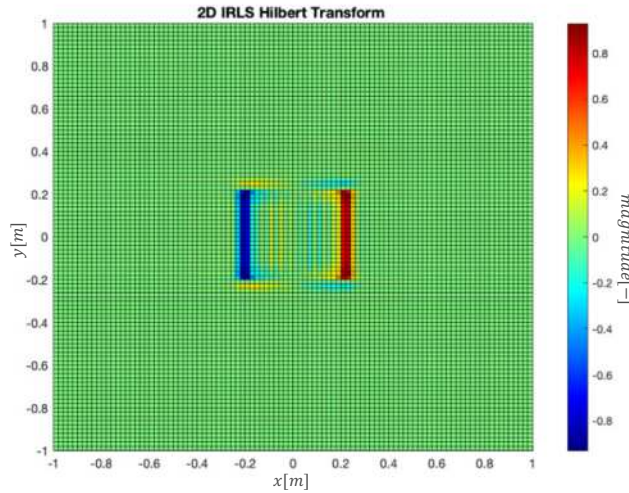

Figure 5-5: The new 2D Hilbert Transform H-IRLS-HT method acting along the X-axis
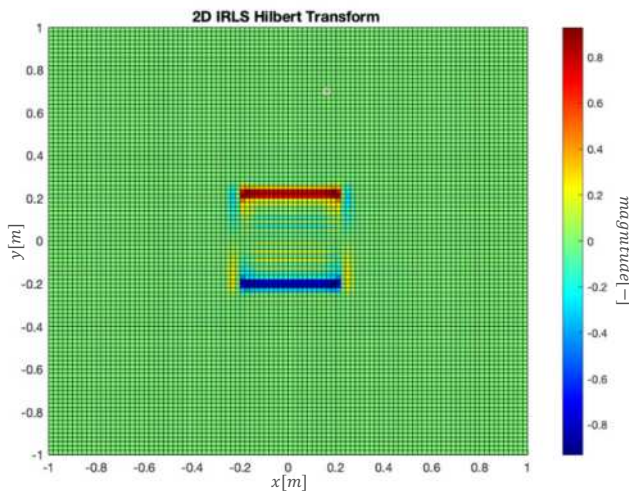

Figure 5-6: The new 2D Hilbert Transform H-IRLS-HT method acting along the Y-axis
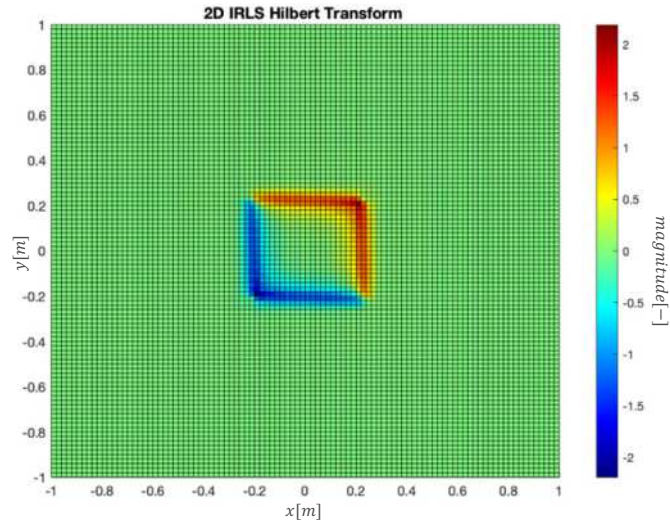
Figure 5-7: The new 2D Hilbert Transform H-IRLS-HT method acting on X and Y axis

To further evaluate the efficacy of the conventional and introduced methods under more realistic conditions, we incorporated noise into the synthetic dataset, simulating challenges similarly encountered in Figure (5-8). The introduction of noise aimed to test the resilience of both techniques in the presence of potentially interfering signals, which often obscure critical features and hinder accurate anomaly detection. By applying the conventional 2D Hilbert method and the 2D H-IRLS-HT to the noisy dataset (with outliers), we sought to assess their respective
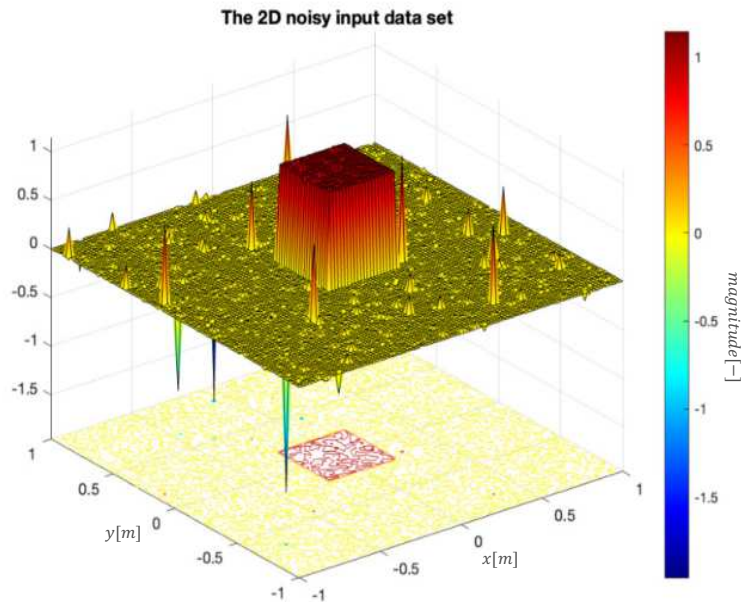

Figure 5-8: The generated synthetic 2D noisy input data set

67

performances and identify the most suitable approach for robust anomaly detection. This comparative analysis enabled us to draw meaningful conclusions regarding the strengths and weaknesses in the field of anomaly detection.

Upon applying the traditional DFT-based 2D Hilbert transform to the noisy synthetic dataset, as shown in Figure (5-9), our analysis revealed several challenges and limitations inherent to this method. The output exhibited a significant level of noise, which adversely affected the clarity and sharpness of the identified anomaly borders. As a result, the edges were less distinct, leading to an impaired ability to accurately detect and characterize anomalies in more complex data sets. Additionally, the presence of numerous outliers within the data further exacerbated the difficulties encountered in the analysis process. These findings underscore the importance of refining existing techniques and exploring alternative approaches to mitigate the impact of noise and outliers, thereby enhancing the reliability and robustness of anomaly detection methods in complex, real-world environments. In contrast, when we applied the newly introduced 2D H-IRLS-HT method to the noisy synthetic dataset, our results exhibited a marked improvement in both the clarity and the quality of the output, as shown in Figure (5-10). The noise present in the dataset was effectively mitigated, revealing a more accurate representation of the underlying spatial patterns. Notably, the absence of outliers in the analysis considerably enhanced the interpretability of the data, allowing for the clear identification and delineation of anomaly boundaries. The improved edge sharpness facilitated a more precise characterization of the anomalies, ultimately yielding more accurate results. This superior performance of the 2D H-IRLS-HT method highlights its potential as a robust
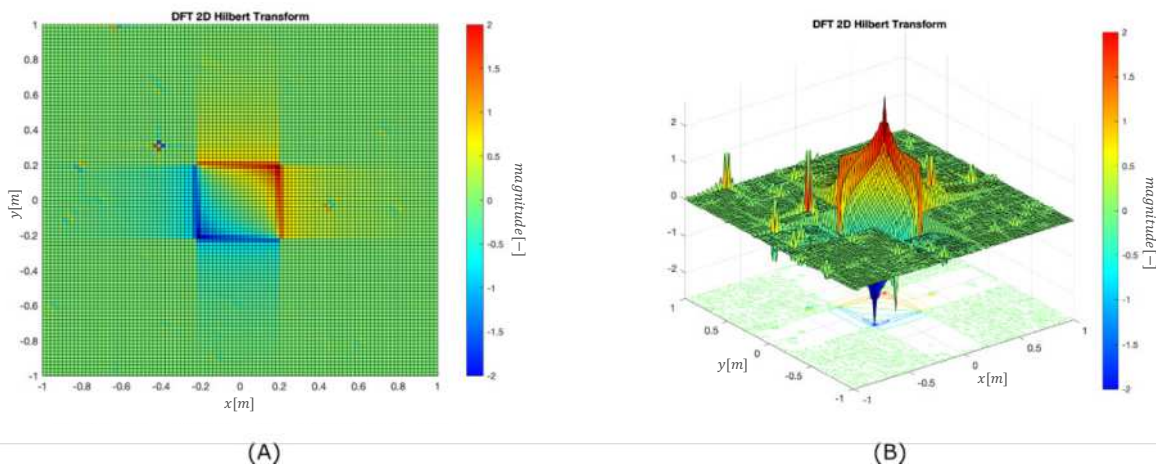

(A)   (B)

Figure 5-9: The 2D DFT-based Hilbert transform of the noisy dataset A) 2D view B) 3D view

and reliable alternative for anomaly detection, particularly in challenging environments where noise and outliers may compromise the efficacy of conventional techniques.
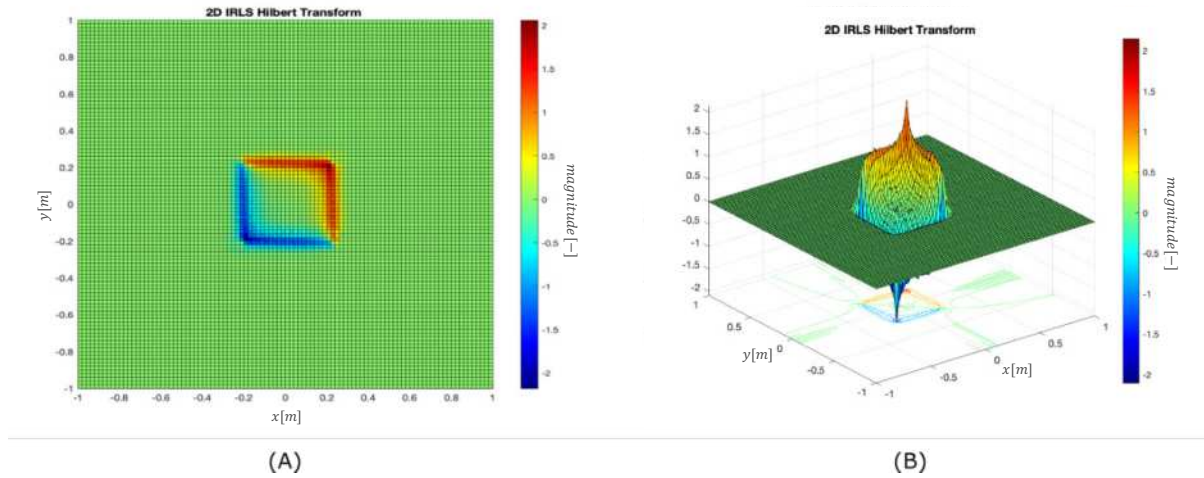

Figure 5-10: The 2D H-IRLS-HT of noisy dataset A) 2D view B) 3D view

In a comprehensive comparative analysis of 2D data distance and model distance methods, it is important to scrutinize their capacity to counteract noise interference for improved edge detection, facilitating more accurate identification of anomalies. These two distinct methodologies differ in their approach to handling data, as we saw in Chapter 3.

Data distance methods primarily involve calculating the dissimilarity between data points by employing various distance metrics. These metrics are instrumental in quantifying the degree of separation between individual points and thus aid in clustering, classification, or outlier detection tasks. Also, we used the proposed Root Mean Square (RMS) distance as a measure between the measured and calculated data in the space domain. The data distance between the noise-free and noisy data sets is 0.4091; meanwhile, the 2D H-DFT-HT method showed the data distance equal to 0.5197. On the other hand, the 2D H-IRLS-HT showed the data distance equal to 0.1074. Ultimately, a thorough evaluation of data distance method should emphasize their noise reduction efficacy, edge detection capabilities, and suitability for the specific application. The introduced method will enable more accurate identification of anomalies, even in the presence of considerable noise.

Furthermore, we aimed to evaluate the effectiveness of our proposed method on a more complex model, which contains the letters H and T (as abbreviation of Hilbert Transform). To simulate the challenge, we intentionally contaminated the model with noise in the form of outliers. The dataset was generated by constructing geometric representations of the H and T characters, followed by the addition of random noise with varying degrees of intensity, as shown in Figure (5-11). This allowed us to assess the robustness of the proposed algorithm in the presence of both structured and unstructured noise. Through iterative testing and refinement, we were able to observe the performance of our approach in identifying and mitigating the effects of outliers on the model's ability to discern the H and T characters accurately. The calculated IRLS-FT spectrum is demonstrated in Figure (5-12).

We employed a defined geometric representation for both the H and T letters, centred in the middle of the model, to ensure a controlled environment for evaluating our proposed method. The T letter was constructed by assigning a binary value of 1 to the model in the following dimensions: the horizontal bar of the T was created using [38:44, 20:44], while the vertical bar was represented by [41:63, 29:35]. Similarly, for the H letter, we assigned a value of 1 to the model in three distinct dimensions: the left vertical bar was created using [38:64, 56:60], the right vertical bar with [38:64, 76:80], and the connecting horizontal bar was represented by [47:51, 61:75].
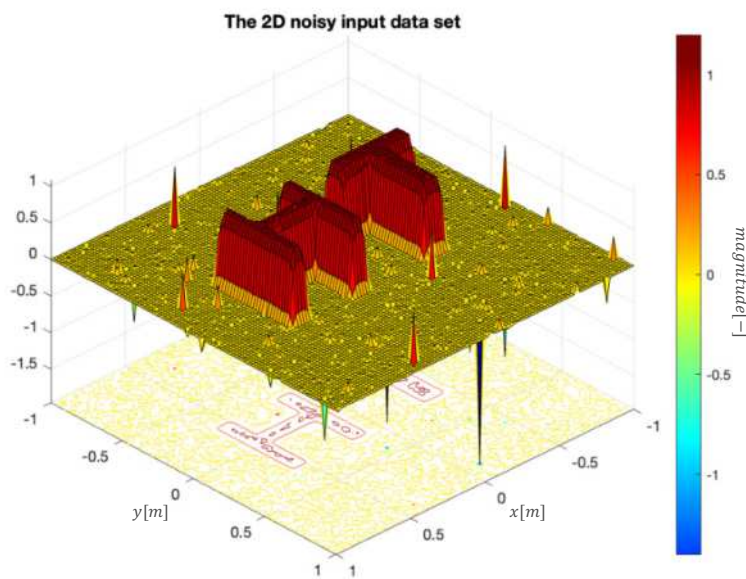

Figure 5-11: The generated synthetic 2D noisy input data set using H and T letters

These predefined positions allowed us to generate a dataset with clear, identifiable patterns of the H and T letters, which facilitated the evaluation of our method's performance in detecting and mitigating the impact of outliers in the presence of structured data.

The Hilbert Transform method using DFT in Figure (5-12) and IRLS-FT with Hermit function in Figure (5-13) were applied to this model to compare the noise rejection capability.
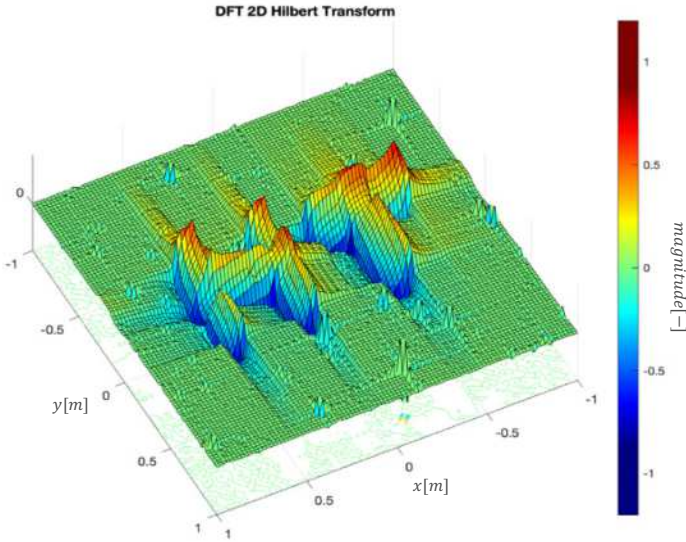

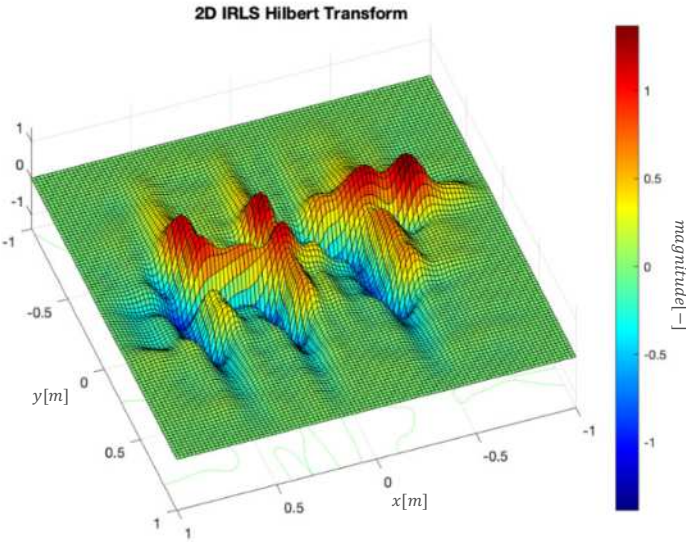*Figure 5-12: The Hilbert Transform method using DFT*

Figure 5-13: The Hilbert Transform method using IRLS-FT

The results provide valuable insights into the effectiveness of the introduced method and suggest potential avenues for future research and optimization in the context of noise and outlier reduction in pattern recognition and edge detection.

Thesis Four:

We introduced the 2D C-IRLS-HT and 2D H-IRLS-HT Hilbert transform methods and sought to evaluate the efficacy of the conventional (DFT-based) 2D Hilbert transformation method and the newly developed procedures in detecting anomalies within a noisy dataset. The traditional DFT-based 2D Hilbert transform encountered numerous challenges, including substantial noise levels, less distinct edges, and the presence of outliers, all of which impeded accurate anomaly detection and characterization in complex datasets. Our investigations show that compared to conventional methods, when the 2D H-IRLS-HT was applied to the noisy synthetic dataset, the results exhibited marked improvements in clarity, quality, and noise reduction, which led to a more accurate representation of the underlying spatial patterns. The superior performance of the 2D H-IRLS-HT and 2D C-IRLS-HT methods underscores their potential as a robust and reliable alternative for anomaly detection, especially in challenging environments where noise and outliers might compromise the effectiveness of conventional techniques.

Chapter 6

6. FIELD DATA

6.1. CASE STUDY

The study area is located in the eastern part of Syria. The total area of the studied area is about 150 km^2 (Figure 6-1).
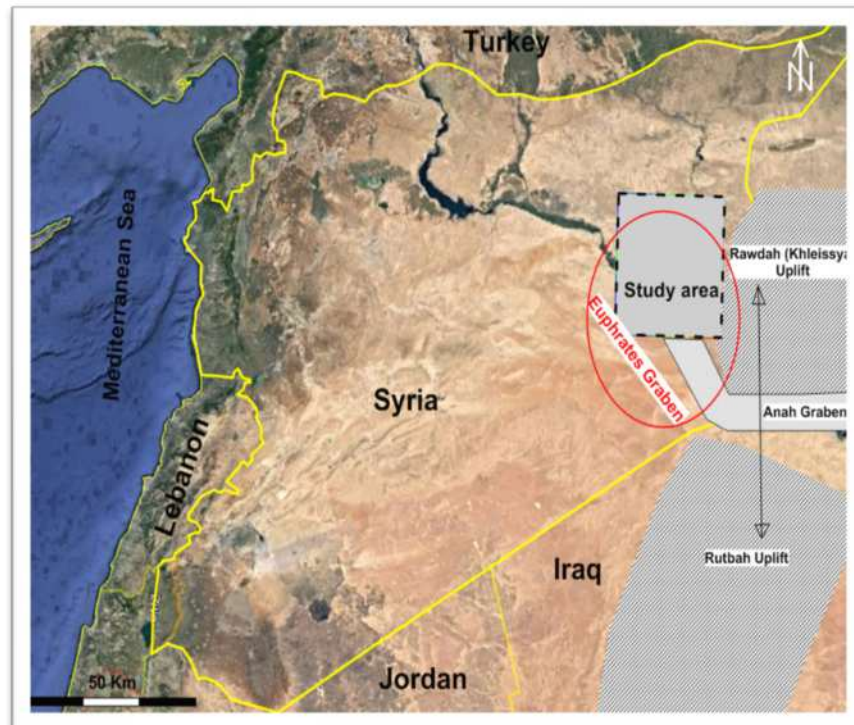

Figure 6-1: Syria map showing the study area situation in East of Syria, and displaying locations of Rutbah and Rawdah Uplifts.

6.2. THE GEOLOGICAL BACKGROUND OF THE MEASUREMENT AREA

6.2.1. THE TECTONIC BACKGROUND OF THE EUPHRATES GRABEN

Syria is located in the northern part of the Arabian Plate, south and west of the main edge where the Arabian and Eurasian plates meet. The Arabian Plate moves in a northeast-north direction and subducts under the Eurasian Plate through the Thrust Fault, forming the Zagros fold belt and the Bitlis Suture (Brew et al., 2001). The Euphrates Basin contains a series of normal faults (resulting from tensile forces) that created a series of uplifted structures that are conducive to trap formation.

As a result, the Euphrates Basin, which is 60-120 km wide, forms an incomplete rift system (Fault-Bounded Failed Rift) that trends northwest-southeast within the Intracratonic (a stable region of the earth's crust composed of an ancient crystalline basement that includes the old Proterozoic shields) located in the northern part of the Arabian Plate, separated from the uplift of the Rataba to the south and the uplift of the Rawdah to the northeast (Figure 6-2).

The Euphrates Basin formed as a result of severe fracturing and vast subsidence during the Upper Cretaceous. The initial phase of the Euphrates Basin rift began with the spread of Toronian - Coniacian discordance due to an uplift that affected the Pre-rift region. Evidence of this rift includes the occurrence of Redbed rocks resulting from the alteration of volcanic tuff rocks that flowed due to early fractures. The Goudia formation below the Redbed was also subjected to weathering and erosion. The fracturing resulting from tensile forces, in addition to the formation of the Euphrates Basin, led to the formation of NW-SE-trending normal faults and the relative separation of the Rawdah and Rataba uplifts. However, these fractures that occurred in the earth's crust did not reach the stage of continental separation.
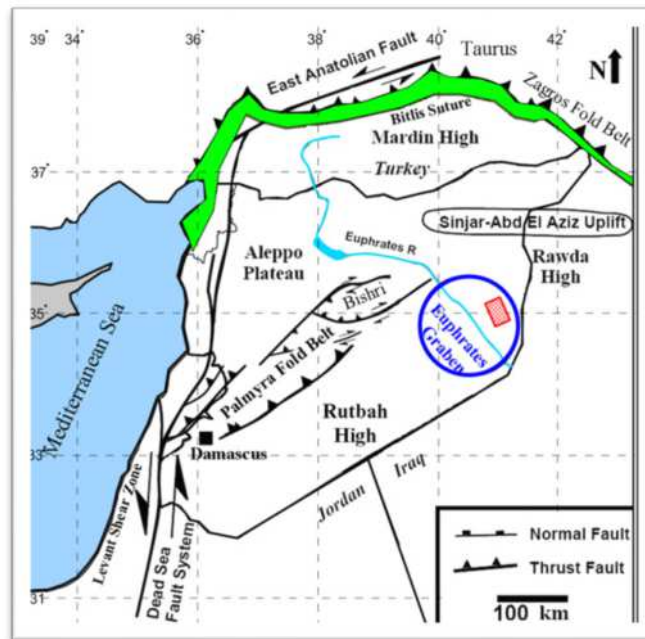
Figure 6-2: A map shows the location of some of the major tectonic units, including the Al-Furat Depression, as well as the Zagros fold belt and the Bitlis suture (Alsouki et al. 2015).

The rifting phase continued during the Santonian period of the formation of the Ramah Group, but with less intensity. Another rifting phase followed it during the Campanian and early Maestrichtian times of the formation of the Ark and Lower Shiranish formations. During this phase, volcanic

activity ceased, and there was a widespread subsidence in the Euphrates Basin, especially in its central part. Seismic data indicates that the rifts and fractures on the wings of the Euphrates Basin, and its southern region, generally stopped at an early stage compared to the central part of the basin, where the growth of faults towards the upper formations of the Lower Shiranish continued. Seismic and well data also show that the upper Ark and Shiranish formations were deposited horizontally (Onlap), with a change in thickness, during the period of significant initial rifting in the tilted fault blocks.

This confirms that the deposition of the lower Ark and Shiranish formations coincided with the later tectonic activity of the initial rifts, where they were deposited with large thicknesses. The activity and growth of the faults began to decline gradually during the deposition of the upper part of the Shiranish formation. This stage coincided with the increasing effect of the subsequent wide subsidence due to the pressures applied to the brittle part of the earth's crust because of the increased thickness and weight of the Upper Cretaceous sediments. It also coincided with a period of thermal subsidence resulting from the decline and cessation of the upper part of the mantle's twisting.

The local changes in the subsidence rate during the deposition of the upper and lower Shiranish formations are shown through the different patterns of their orientation (onlap, offlap) overlying the formations below belonging to the Lower Cretaceous. The onlap pattern of the Shiranish formations' orientation indicates that the uplift or tilting of the fault blocks occurred in a later stage after the deposition of the Upper Cretaceous formations, while the offlap pattern indicates that the uplift occurred in an earlier stage before the deposition of the Shiranish formations, which in turn accelerated and increased the uplift and tilting of the fault blocks.

The third epoch (Tertiary) was characterized, particularly during the Paleogene, by a wide thermal subsidence resulting from thermal equilibrium restoration in the lithosphere after the rifting stage (Litak et al., 1998). During this period, a wide depositional subsidence occurred in the Euphrates area, leading to fragmentation and shattering along active fault lines and blocks in the Euphrates Basin. However, the most significant and dominant event in Syria, both generally and partially on the western outskirts of the Euphrates Basin, occurred during the Miocene, when the Crustal Shortening system controlled Syria's structure as a result of compressive forces caused by the continental collision of the Arabian Plate with the Eurasian Plate, particularly the northern borders

of Syria, near the Anatolian region of Turkey. The compressive deformation led to a secondary inversion (resulting in the uplift and folding of northeast-southwest-trending faults) in the Tadmur and Sinjar sedimentary basins with an SW-NE trend. This event only affected the western outskirts of the Euphrates Basin adjacent to the Tadmur Basin. At the same time, this deformation was utterly absent from the central and southern parts of the Euphrates Basin. Several theories explain the mechanism that caused the Euphrates Rift, as some researchers, such as (Alsdorf et al., 1995), have assumed that the continental collision between the Arabian and Eurasian plates led to the creation of intense vertical forces on the fault line, contributing to the formation of the Euphrates Rift system.

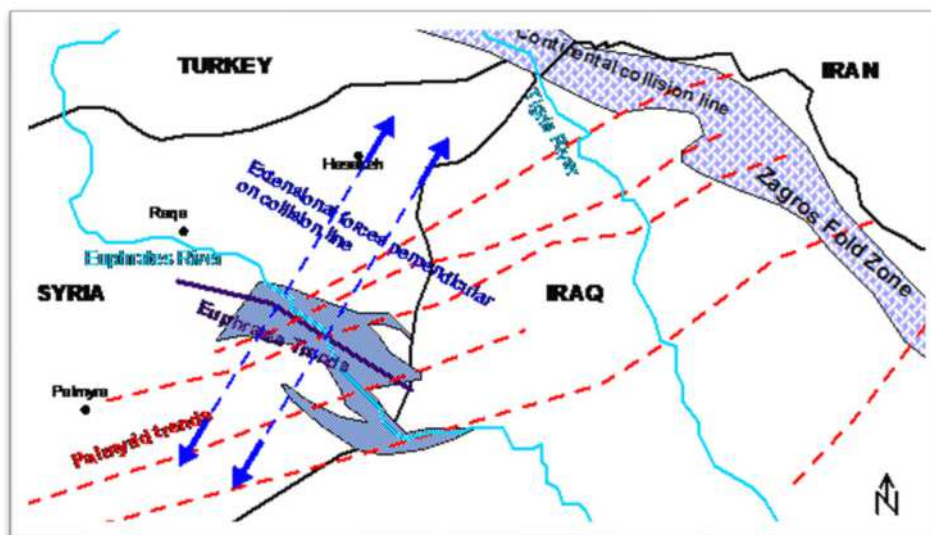However, the early cracking of the Euphratesbefore the mentioned continental collision justifies

Figure 6-3: Illustrates the assumed tensional forces responsible for the formation of the fracture system in the Euphrates (Alsouki et al. 2015).

the rejection of this hypothesis. Others, such as Lovelock, have suggested the Slab-pull theory as a possible mechanism for the formation of passive rifting in the Euphrates. Another hypothesis assumes that the gradually increasing load of volcanic formations led to the formation of strong forces that caused the Euphrates rifting, which ceased with the Arabian plate's continental collision during the Maestrichtian period. The Euphrates rifting system can be divided into several structural units based on the degree of deformation and geometric shape, including the deep central graben unit, a set of peripheral graben units surrounding the central Euphrates graben, stable and relatively elevated platforms, as well as flat, elevated terraces or high plateau areas surrounding the periphery (Figure 6-3) (Alsouki and Taifour, 2015).

The Central Euphartes extends northwest-southeast and forms the deepest part of the Euphrates Rift Basin, where the depth increases towards the northwest, away from the humid and alluvial ridges located in the south and southeast. During the Upper Cretaceous (rift) period, intense tectonic forces were concentrated in the Central Euphrates Basin, resulting in the formation of inverted blocks separated by large NW-SE faults that do not trend towards the axis of the central basin but instead are inclined away from it. Small peripheral grabens, such as the Madaba, Burghuth, Jido, and Ward grabens (Figure 6-4), branch off from the Central Euphrates to the south and southeast. To the north and northeast, there are the Sijan, Tayyani, and Northern grabens, which are characterized by their narrowness and great depth its also surrounded by very large
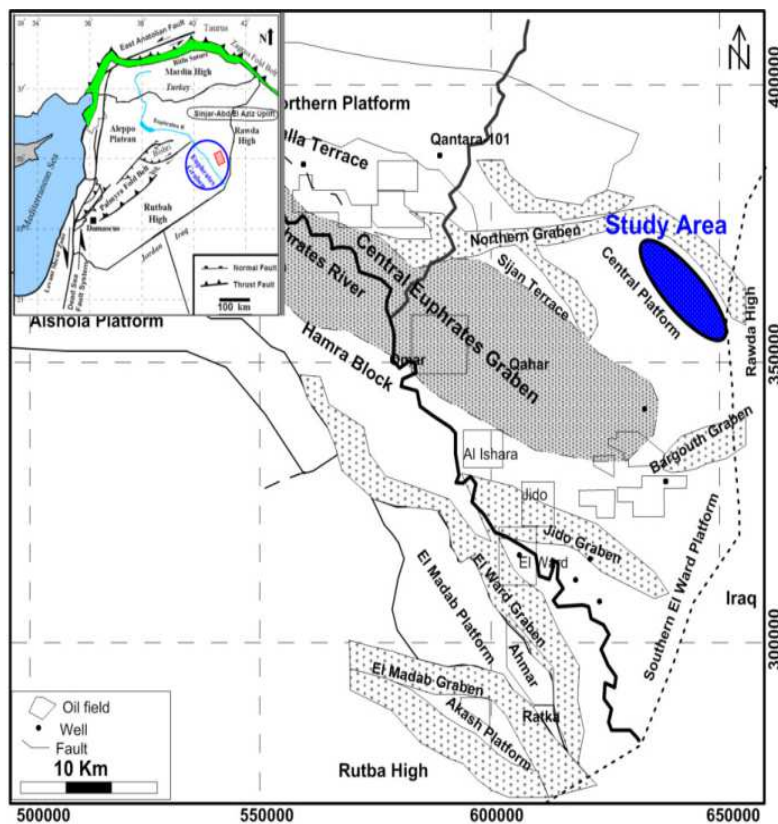

Figure 6-4: Tectonic map of the Al-Furat Basin during the Late Cretaceous base period (Alsouki et al. 2015).

faults. Many of these grabens are filled with continental sediments representing the Dier Abu Saif Formation, which resulted from the weathering of intrusive volcanic rocks during the rift stage. Stable, relatively elevated platforms separate these grabens in the rift basin, such as the Central Platform in the northeast of the central basin and the Sijan block to the north, as well as the southern platform, including the Ward, Madaba, Akash, and Aqaba blocks to the south and southeast.

Most of these platforms are inclined towards the central axis of the rift system, with the inclination of some platforms potentially starting before the early stages of rifting, but most occurring during and after the rift stage, i.e., during the broad subsidence phase of the Upper Shiranish/Mastrichtian and Paleogene times.

6.2.2. THE LOCAL STRUCTURAL ELEMENTS OF THE EUPHRATES GRABEN

The Euphrates Graben, situated in the northern part of the Arabic plate, is considered a rift fault system formed due to intensive intracontinental plate deformation. This Graben with Anah Graben located inside Iraqi territory has evolved within the northern Arabic plate and divided the Rutbah Zone "oriented the N-S" to two stable areas, the Rutbah Uplift to the south and the Rawdah (Khleissya) Uplift to the northeast in Late Miocene-present age (Al-Heety et al., 2017). Both stable areas comprise significant basement uplifts. This rift fault system with general trending E-W to NW-SE could be categorized into several groups of structural elements based on the intensity and shape of the structural deformations: central graben, marginal grabens, and marginal platforms.

The central graben with NW-SE orientation is considered the largest structural element and constitutes the deepest zone of the rift Euphrates Basin. It is governed by a complicated system of overturned fault blocks like Omer Block, the largest oilfield in the Euphrates Graben and Syria. This rift basin also contains several rising platforms or terraces like Sijan Terrace and Madabe Platforms. The Northern platform extends towards the west side of the Rawdah Uplift, the top of which lies on the outlines with Iraqi territory.

The Rawdah Uplift appears to be relatively moderate compared to the Euphrates structure slopes as indicated by seismic profiles. Marginal grabens exist on the sides of the Euphrates Graben, to the south, southeast, and northeastwards of the main graben, like the Madabe Graben, Jido Graben, Elward Graben, the Anah Graben, the Northern Graben, and Sijan Graben (Figure 6-5). These grabens contain vast quantities of rocks that are composed of sedimentary succession and early syn-rift volcanic rocks.
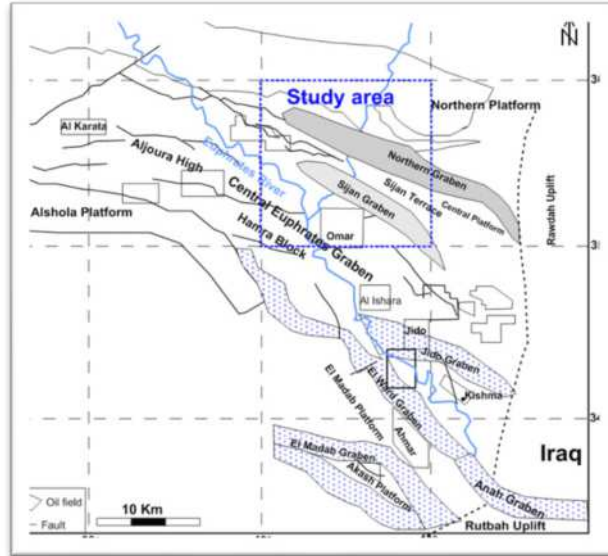
Figure 6-5: The structural map of the Euphrates Graben based on seismic data displaying the local structural elements, particularly Northern and Sijan Grabens besides Omar and Sijan blocks located on the study Area (Alsouki et al. 2015).

6.3. GRAVITY DATA

The exploration company carried out gravity and magnetic geophysical measurements within an area of approximately 128 km^2. The Scintrex CG gravity meter, which is the latest device from Scintrex with a standard deviation of 0.5 mGal, was used along with a GPS locator. Gravity measurements were taken and then interpreted using a set of basic gravity data corrections (terrain, Bouguer, free air, and other corrections) while considering the rock density of the study area, which was determined to be 1.9 g/cm^3.

In order to utilize the gravitational measurements to obtain important subsurface geological information through interpretation processes, it was necessary to carry out some necessary processing operations. In the previous stages of my research in this area, we utilized Surfer software to identify the structure within the dataset. However, as discussed in the later chapters, we have now introduced a more advanced method 2D C-IRLS-FR to effectively use it in low-pass filter process. This innovative approach will contribute to a more robust interpretation of the data and facilitate the development of more reliable conclusions.

In the subsequent phase of the research, we employed MATLAB as the primary tool for data processing and analysis. The first step in this procedure involved importing the dataset into the MATLAB environment as shown in Figure (6-6). This allowed us to leverage the powerful computational capabilities of the software to conduct a comprehensive examination of the data. By utilizing the extensive suite of functions and libraries that we developed in this thesis, we were able to effectively perform a series of operations, such as filtering, and visualization, to better understand the underlying patterns and trends within the data.
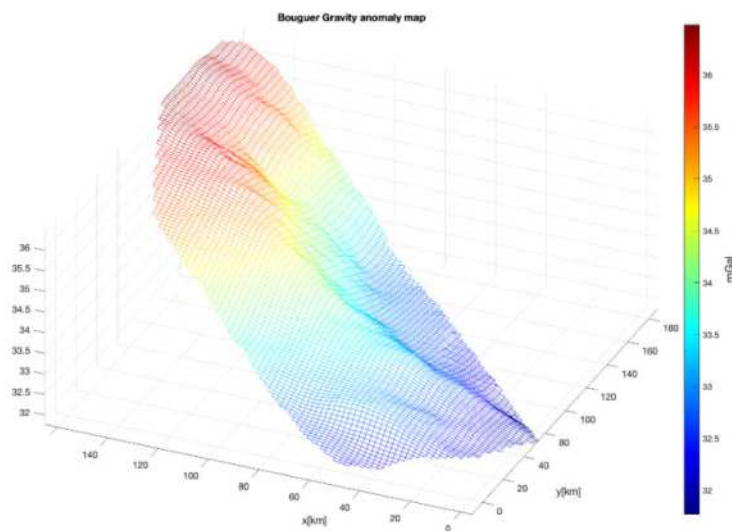

Figure 6-6: The gravity measurements in 3D view

In our investigation, we employed 2D C-IRLS-FT method to systematically evaluate the gravity measurement dataset, with the primary objective of testing the method itself and examining its subsequent effects on the data prior to conducting any in-depth analysis. This rigorous approach facilitated a comparative assessment, enabling us to establish a comprehensive understanding of both the method's efficacy and its impact on the unaltered gravity measurements. Consequently, this preliminary examination served as a robust foundation for further investigation, ensuring the reliability and validity of our findings in the context of gravity measurement research. Figure (6-7) shows the gravity measurements before and after applying the introduced method.
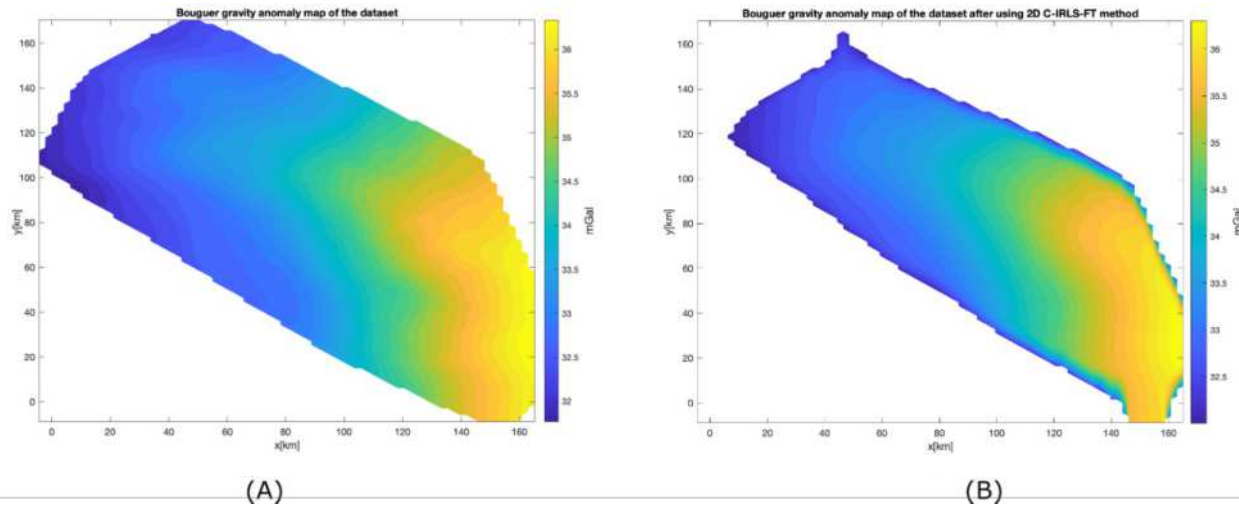
Figure 6-7: The gravity measurements A) Befor B) After apply the 2D C-IRLS-FT method

The minimum values of gravity anomalies are attributed to the gradual subsidence of the basement surface, reaching its lowest point in the Sejan area located southwest of the study area. The low-range areas of the apparent gravity anomalies, colored in blue, are filled with large thicknesses of sediments from the second, third, and fourth periods, which are characterized by their low density.

The Bouguer anomaly values in the study area change from 31 mGal in the western part to more than 36 mGal in the eastern part of the study area, where gravity values increase from the north and northwest to the south and southeast. The high values of gravity anomalies in the southeastern part are attributed to the rise of basement rocks extending in a north-south elongated range. This range represents the western edge of the Rawda uplift, which increases in elevation towards the east in the direction of the Iraq lands.

We demonstrate the application of a two-dimensional low-pass Butterworth filter on gravity data. After converting the data to double precision, a 2D C-IRLS-FT is applied. A grid for the low-pass Butterworth filter is created, with its size defined by the dimensions of the data. The filter parameters, such as the normalized cutoff frequency and filter order, are set to 0.1 and 2, respectively. The low-pass Butterworth filter, which is a function of the distances from the center point of the grid, is then applied to the gravity data in the frequency domain (Butterworth, 1930). The filtered data is then shifted back, and the inverse procedure is applied to convert it back to the spatial domain as a part of the 2D C-IRLS-FT method. Finally, low-pass filtered gravity data are visualized, as shown in Figure (6-8). The Fourier Transform, using 2D C-IRLS-FT, a key

81

component in this procedure, allows for the decomposition of the gravity data into its frequency components, enabling the application of the low pass filter in the frequency domain.
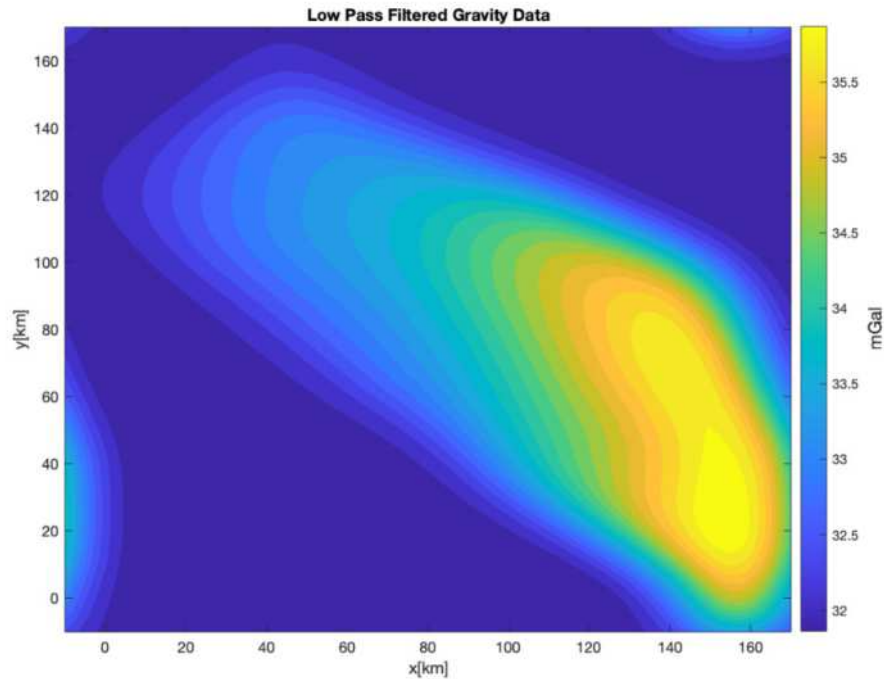
Figure 6-8: Low pass filter for the gravity data using 2D-IRLS-FT method

By providing a more reliable and coherent picture of the subsurface, the method enabled a more efficient and accurate assessment of the geological features and their spatial relationships.

In summary, the utilization of the 2D C-IRLS-FT method for calculating the low pass filter as an example of an application proved to be a valuable addition to the analytical toolkit. The results obtained through this approach facilitated a more comprehensive interpretation of the subsurface structures, contributing to an understanding of the geological context of the study area.

6.4. MAGNETIC DATA

As for the magnetic measurements, the MINIMAG MMPOS-1 device was used, which can be used as a programmable independent magnetic station, taking one reading every twenty seconds, with an operating range between 20-100,000 nT, and an absolute measurement accuracy about ±10 nT due to problems in the device during measurements.

We imported magnetic data into the MATLAB environment to complement our analysis. Incorporating this supplementary dataset provided a broader perspective on the subsurface structures and allowed for a more comprehensive understanding of the underlying geological features, as shown in Figure (6-9).
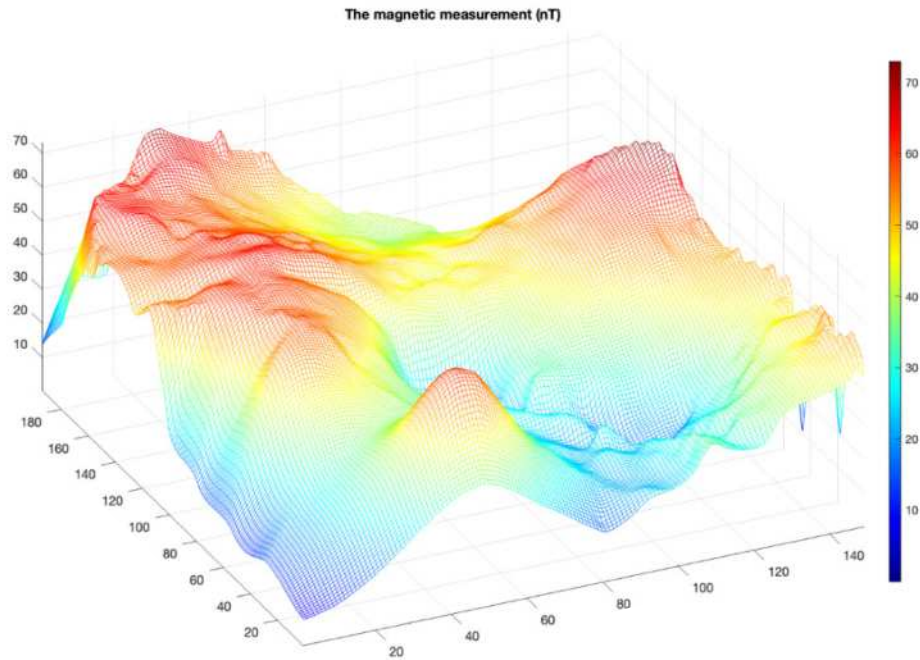
Figure 6-9: The magnetic measurement in 3D view

The magnetic intensity map indicates a consistent increase in magnetic intensity values from the southwestern region toward the northeastern region. The intensity increases from a value of 10 nT in the southern part to a value of 76 nT in the northern and northeastern parts. This can be attributed to the gradual rise of basement rocks in the eastward and northeastward directions within the study area, which corresponds to the uplift of basement rocks towards Rutba. Conversely, the basement subsides towards the south in the direction of depression Sejan and Bargouth, where the thickness of the sedimentary cover reaches its maximum, resulting in a decrease in magnetic field values.

To gain further insight into the properties of the field data, we proceeded to calculate the frequency spectrum of the data set (Figure 6-10).
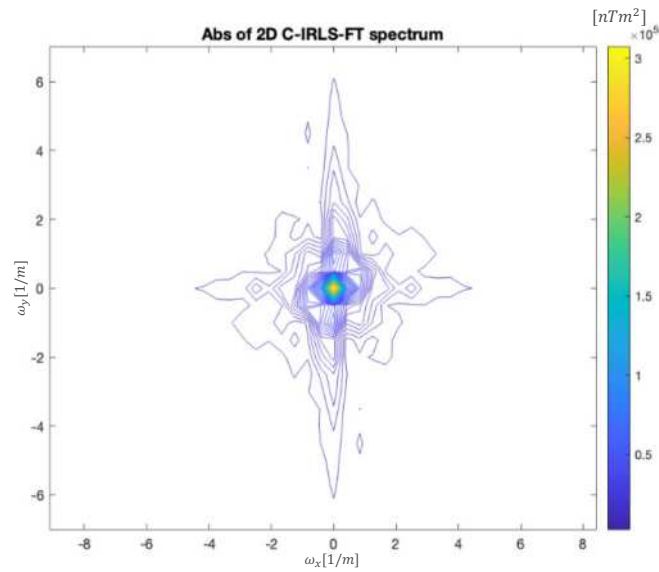

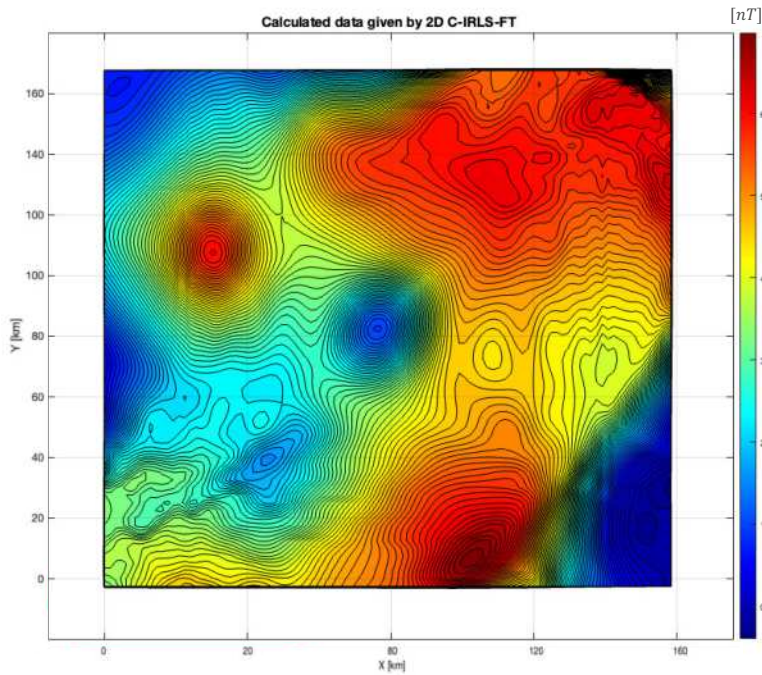*Figure 6-10: The spectrum of the magnetic data using 2D C-IRLS-FT method*

Figure 6-11: The calculated data using C-IRLS-FT method

In light of the findings from the spectral analysis, we proceeded to apply the C-IRLS-FT method (Figure 6-10) to the selected data segment with the aim of applying the pole reduction, reducing

noise, and improving data quality. The implementation of this technique resulted in a significant reduction in noise levels, thereby enhancing the overall clarity and interpretability of the dataset (Figure 6-11).

The pole reduction process, as shown in Figure (6-12), which is a standard approach used in the analysis of magnetic anomalies. This method involves transforming the measured magnetic data to a common reference, typically the magnetic pole, in order to minimize the effects of varying magnetic inclinations (45-50°) and declinations (60° positive). By applying the pole-to-pole reduction, we were able to effectively minimize any potential distortions within the dataset, ultimately enabling a more accurate assessment of the underlying geological structures and properties.
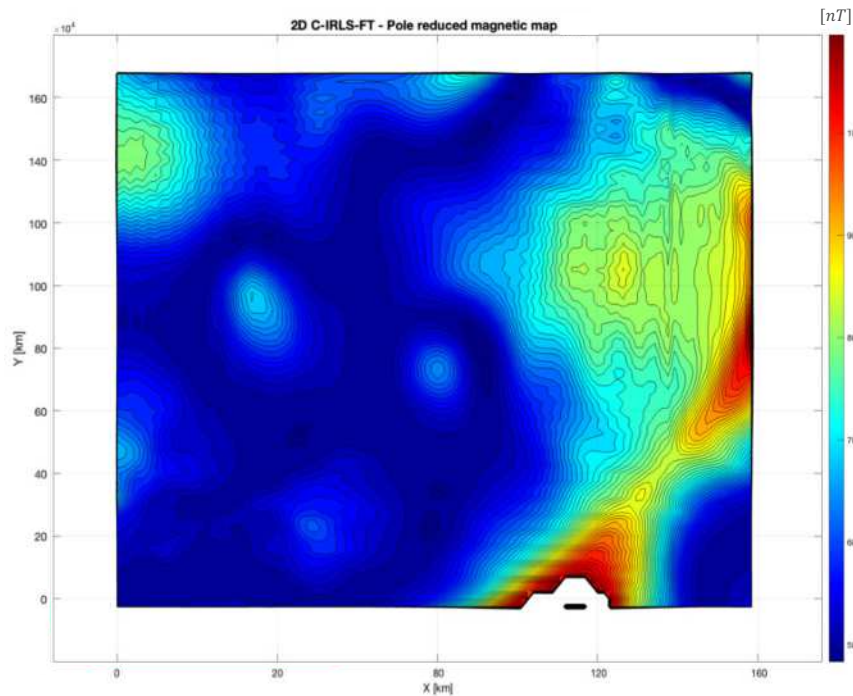
Figure 6-12: Pole reduced magnetic map using 2D C-IRLS-FT method

Following the application of the pole-to-pole reduction technique, the clarity and co herence of the geological structures within the dataset were significantly enhanced. The removal of outliers and noise from the data led to a more distinct representation of the structure, allowing for a more straightforward and reliable interpretation. Consequently, this improved data quality facilitated the identification of key geological characteristics and enabled a more accurate assessment of the underlying spatial patterns and relationships. In summary, the combination of the presented noise

reduction methods and the pole-to-pole reduction technique played an essential role in refining the dataset and ensuring that the subsequent interpretation was both accurate and robust.

6.5. SEISMIC DATA

The data used in this area include a three-dimensional seismic survey (seismic cube) covering an area of approximately $150 \, \text{km}^2$. The boundaries are shown in Figure (6-13). The seismic lines were surveyed in a northeast-southwest direction. The data was processed using the pre-stack time migration method to improve the quality and image of the seismic data, which achieved good resolution throughout the geological section from the Triassic to the top. However, this feature weakens in deeper layers due to energy attenuation and high-frequency absorption. The quality of the seismic data also decreases in the time levels close to the surface due to energy absorption by the salt rocks.

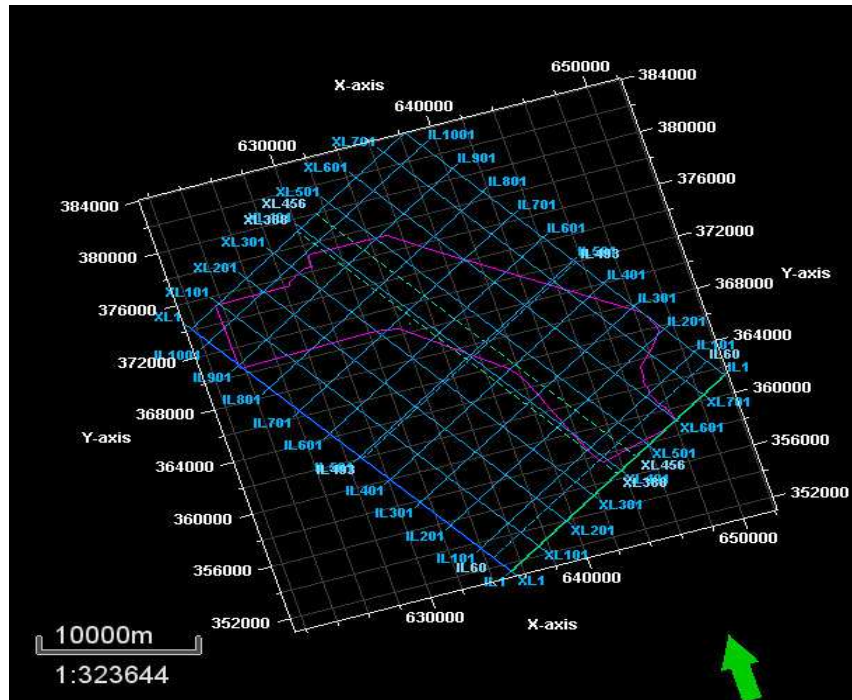*Figure 6-13: The shape of the study area*

We will present a method for denoising seismic data using the 2D C-IRLS-FT and k-means clustering technique that will be applied to a smaller region within the seismic section that contains a significant amount of noise. The method involves transforming the seismic data into the frequency domain, identifying the significant spectral components through clustering, and

subsequently filtering out the noise components. The seismic data is first transformed into the frequency domain in the process of the 2D C-IRLS-FT method. The amplitude spectrum is computed as the absolute value of the transformed seismic data and reshaped into a one-dimensional array to serve as input features for the k-means clustering algorithm. In order to provide a comparative analysis of the denoising method, we also compute the amplitude spectrum of the seismic data without applying the proposed technique, as Figure (6-14) showed. This will serve as a baseline for evaluating the effectiveness of the 2D C-IRLS-FT and k-means clustering method in enhancing the amplitude spectrum and reducing the noise components.

K-means clustering (MacQueen, 1967) is a partition-based clustering technique that aims to minimize the sum of squared distances between each data point and the centroid of the cluster it belongs to. The algorithm is initialized with 'k' centroids, and each data point is assigned to the nearest centroid. The centroids are then updated as the mean of all the data points assigned to them. This process is iterated until convergence or a predetermined number of iterations is reached.

In the introduced method, k-means clustering is applied to the feature set to separate the significant spectral components (signal) from the noise that remains from the first process of the 2D C-IRLS-FT method, with the number of clusters, k, set to 2, representing the signal and noise components.
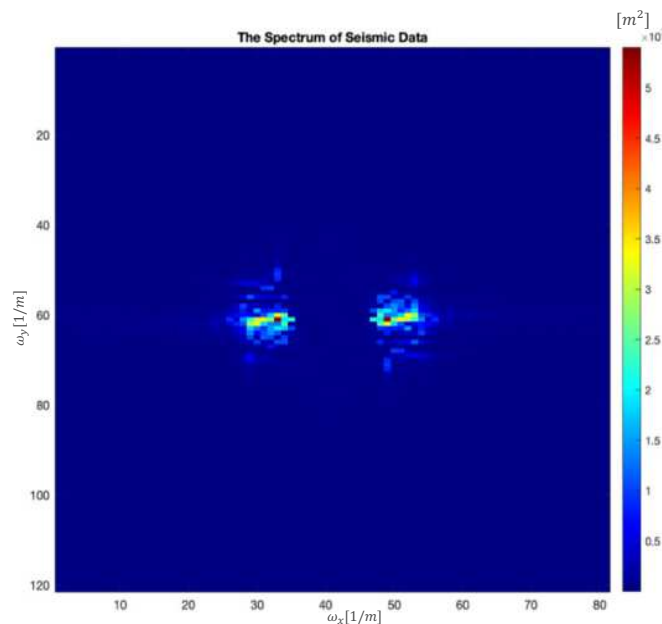
Figure 6-14: The spectrum of the noisy seismic data

The cluster with the highest mean amplitude is identified as the signal cluster. Using the clustering results, the seismic data in the frequency domain is filtered by setting the amplitude of the noise cluster components to zero. The filtered amplitude spectrum is visualized in Figure (6-15), and the inverse process of the 2D C-IRLS-FT is applied to the filtered data to obtain the denoised seismic data in the spatial domain.
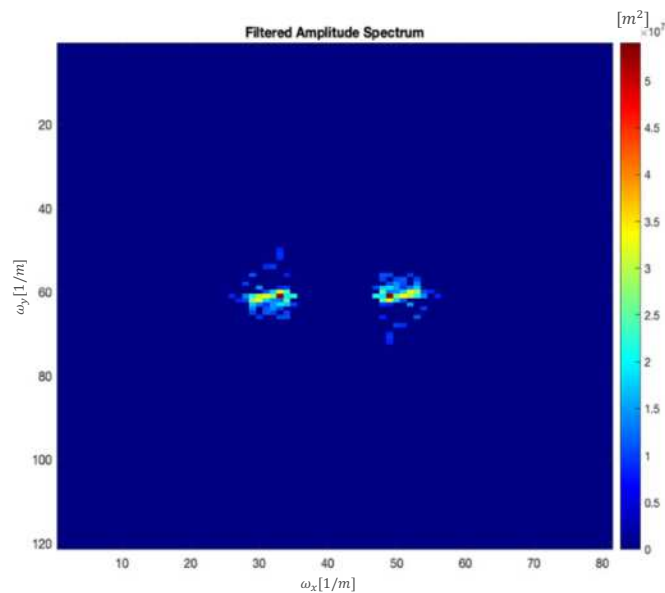
Figure 6-15: The filtered amplitude spectrum of the noisy seismic data

The proposed method demonstrates effective denoising of the seismic data, as evidenced side-by-side comparison of the original and denoised seismic sections, as shown in Figure (6-16).

The amplitude spectra visualizations show a clear distinction between the signal and noise components, and the filtered data retains the significant spectral components while eliminating the noise.
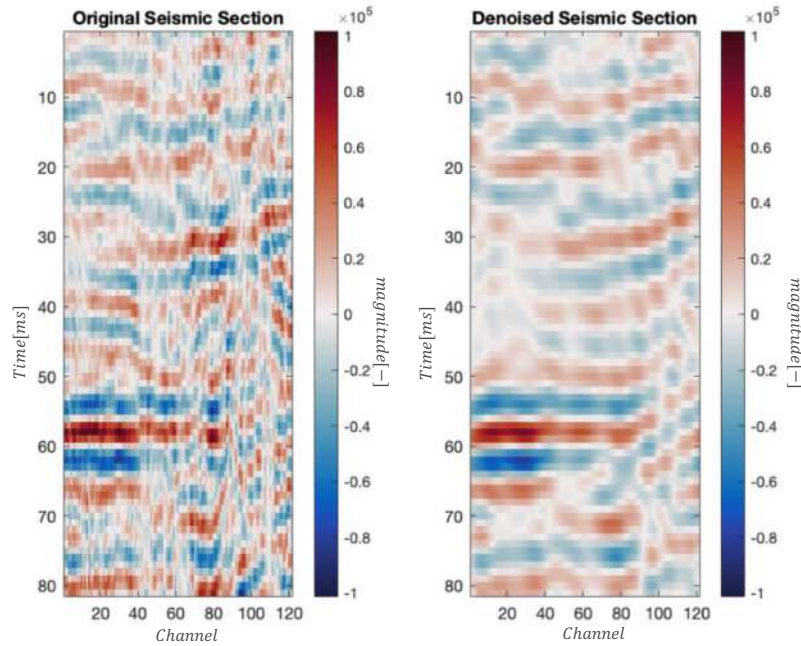
Figure 6-16: The orginal seismic section and the denoised one using 2D C-IRLS-FT with clustering

Thesis Five: In the concluding analysis of the research conducted in the eastern part of region Syria, we introduced a more advanced method, the 2D C-IRLS-FT, particularly in the low-pass filtering of gravity data. The inversion-based Fourier transform using the Chebyshev function as basis function (2D C-IRLS-FT method) played a pivotal role in this procedure, enabling the decomposition of the gravity data into its frequency components. Subsequent spectral analysis illuminated the necessity of applying the 2D C-IRLS-FT method to a chosen data segment for noise reduction, pole reduction, and overall enhancement of data quality. The pole-reduction technique further amplified the coherence of the geological structures within the dataset, leading to a noticeable improvement in data quality. The sufficient reduction of noise and outliers from the data yielded a more distinct representation of the structure, thereby facilitating a more reliable interpretation. As an extension of our current research scope, a novel technique that combines the 2D C-IRLS-FT method and k-means clustering was deployed for denoising seismic data. This was specifically undertaken within a seismic section that harbored a significant noise component. This method hinges on the transformation of seismic data into the frequency domain.

Chapter 7

7. CONCLUSION

The research problem addressed in this PhD thesis revolves around the persistent challenge of noise and outlier contamination in geophysical data, specifically focusing on seismic, gravity, and magnetic measurements, as well as in other 2D images that encompass geophysical data. Noise and outliers in geophysical datasets can originate from various sources, such as instrumental errors, environmental factors, and processing artefacts. The presence of these unwanted signals can significantly impact the accuracy and reliability of interpretations and output results, ultimately affecting our understanding of subsurface structures, resource exploration, and hazard assessment. This, in turn, may lead to suboptimal decision-making in the context of natural resource management, environmental protection, and infrastructure development. In light of these challenges, there is a pressing need to develop effective methodologies for identifying and mitigating noise and outlier-related issues, intending to enhance the quality and reliability of geophysical data interpretation. This thesis aims to contribute to this area of research by exploring novel techniques and algorithms for noise and outlier detection and reduction by assessing their performance in various geophysical applications. We recognize the importance of Fourier transformation as a widely utilized tool in geophysical data processing, particularly for enhancing the quality of datasets and providing a comprehensive picture of subsurface geology, but the traditional Discrete Fourier Transform (DFT) approaches have limitations in processing outlier noisy data. From this point, to address these challenges, this thesis introduces an inversion-based 1D and 2D Fourier transformation (C-IRLS-FT) algorithm. These research findings demonstrated significant improvements in both the space and frequency domains, providing a strong foundation for the development and application of the C-IRLS-FT algorithm in the context of noise and outlier detection and reduction in geophysical data.

To comprehensively evaluate the applicability and effectiveness of the C-IRLS-FT method, three of the objectives of this thesis are centred around reducing the sensitivity to outliers by implementing the inversion-based Fourier transformation on synthetic 1D and 2D datasets, as well as real field measurements. The overarching aim is to improve the reliability and accuracy of

geophysical data interpretation, paving the way for more informed decision-making in resource exploration. To achieve this goal, the proposed C-IRLS-FT inversion approach is primarily grounded in the iteratively reweighted least-squares Fourier transformation. This iterative process allows for the fine-tuning of the analysis and enhances the ability to address noise and outlier-related issues. The series expansion technique is employed to discretize the Fourier frequency spectrum, which provides a more manageable and computationally efficient representation of the data. The expansion coefficients are then estimated as the solution to the over-determined inverse problem, further refining the analysis and ensuring that the method is robust in the face of diverse challenges. The Chebychev polynomials are used as basis functions for this approach. This choice enables rapid and accurate computation of the elements of the Jacobian matrix, streamlining the overall process and increasing the method's efficiency. Furthermore, the most frequent value (MFV) method is employed to address the issue of scale parameters, which can be a critical factor in the analysis of geophysical data. By iteratively determining the Cauchy-Steiner weights through an internal iteration loop, the MFV method minimizes data loss and contributes to the robustness of the C-IRLS-FT method. Overall, this carefully crafted combination of methods and techniques results in a powerful and effective approach to handling noise and outlier challenges in geophysical datasets.

Initially, the inversion-based Fourier transformation method using Chebychev polynomials (C-IRLS-FT) is scrutinized using one-dimensional synthetic datasets. In order to evaluate its noise rejection capabilities, a comparison is made between the proposed inversion technique and the conventional Discrete Fourier Transformation (DFT) method, referencing the research by Dobroka et al. (2012). A time-domain 1D synthetic wavelet is created for this purpose. To gauge the sensitivity to noise, the initially noise-free wavelet is intentionally contaminated with both Gaussian and Cauchy noise, which mimics real field data measurements. Subsequently, the real and imaginary components of the Fourier frequency spectrum are calculated using both the traditional DFT and the inversion-based C-IRLS-FT methods, converting the noisy datasets from the time domain to the frequency domain. The results indicate that the inversion-based Fourier transformation (C-IRLS-FT) method demonstrates robustness and a significant ability to mitigate the influence of outliers compared to the outcomes achieved using the conventional DFT method. This observation is further substantiated by numerical data derived from data and model distance values. The superior performance of the C-IRLS-FT method in addressing noise and outlier

contamination underscores its potential for improving the accuracy and dependability of geophysical data interpretation across various applications.

Owing to the effectiveness of the IRLS inversion technique in processing 1D datasets, the inversion algorithm is further refined for application to 2D synthetic data. The 2D synthetic data is generated to encompass a rectangular region with dimensions spanning from -1 to 1 unit for both the x and y directions. Additionally, this dataset features an anomaly located centrally with dimensions of -0.2 to 0.2 units for both x and y directions. The sampling intervals are established at 0.04 units for both directions, resulting in 101x101 data points. To assess the outlier sensitivity, the dataset is contaminated with random Gaussian and Cauchy noise, similar to the 1D case. To achieve inversion accuracy and stability, the Chebychev polynomials were used as a basis function of order (Mx=My=26) and selected for data processing calculations. Upon data examination, it becomes clear that the 2D C-IRLS-FT technique demonstrates superior noise reduction capabilities in comparison to the traditional 2D DFT approach. However, it is crucial to develop a more robust and effective method for filtering out random noise and outliers, considering the limits of the Least Squares and 2D DFT methods. Therefore, it is highly recommended to explore the utilization of the 2D C-IRLS-FT method for improved noise reduction in 2D geophysical data.

Building upon the foundation of previous research, this study has further explored the critical role of attributes in seismic data processing and interpretation. Attributes are essential tools for extracting crucial information, such as physical and geometric parameters, which may not be obtainable through conventional means. Since the seminal work by Taner et al. in 1979, the field has experienced significant growth, resulting in a diverse range of attributes that can be classified into various categories, such as physical, geometric, and procedural. In the realm of developing processing methods, optimizing the signal-to-noise ratio is of utmost importance. Fourier transforms often play a pivotal role in this process, and our previous work proposed a robust Fourier transform inversion (C-IRLS-FT) procedure that effectively suppresses outliers and significantly improves the signal-to-noise ratio. On these findings, we further explored the application of this procedure to computing attributes, with a particular focus on the Hilbert transform, which serves as a vital component of the robust/resistant inversion framework. The Hilbert transform allows for the analysis of the phase and frequency components of a signal. In this study, we have demonstrated that the newly developed method, referred to as C-IRLS-HT,

outperforms the traditional Hilbert transform method in reducing noise for two seismic attributes. This superior performance is evidenced by the lower data distance values obtained through its application. Specifically, when Method C-IRLS-HT was applied to a dataset with high levels of generated noise for the first attribute (Reflection Coefficient), it demonstrated a significant reduction in error, resulting in more accurate and reliable data compared to the traditional Hilbert transform method. Similarly, for the second attribute (Instantaneous Phase), Method C-IRLS-HT showed greater effectiveness in reducing error when applied to a dataset with the same noise levels, as opposed to the traditional Hilbert transform. These findings not only underscore the effectiveness of Method C-IRLS-HT as a more robust and reliable method for data analysis in situations where noise reduction is essential, but they also point to the potential for further advancements in the field of seismic data processing and interpretation through the development and application of innovative techniques.

Expanding on the foundational principles of the one-dimensional Hilbert Transform, the two-dimensional (2D) Hilbert Transform has demonstrated its potential as an influential instrument in image processing. By enabling the efficient analysis and extraction of spatial information and the identification of features that may be hidden by traditional methods, the 2D Hilbert Transform has made significant strides within the realm of earth science-related image processing, particularly in edge detection, noise reduction, and image enhancement. The 2D Hilbert Transform can be characterized as a convolution operation involving the image and a 2D Hilbert kernel, which is frequently derived from the product of two 1D Hilbert kernels. This operation can be extended to generate a 2D analytic signal that encompasses both the original image data and its Hilbert Transform, providing valuable insights into the image's spatial frequency content and phase information. In this research, we sought to evaluate the efficacy of the conventional 2D Hilbert method and the newly introduced both 2D C-IRLS-HT and 2D H-IRLS-HT method (using Hermite functions as basis functions) in detecting anomalies within a noisy dataset. The traditional DFT-based 2D Hilbert transform encountered numerous challenges, including substantial noise levels, less distinct edges, and the presence of outliers, all of which impeded accurate anomaly detection and characterization in complex datasets. These findings emphasize the necessity of refining existing techniques and exploring alternative approaches to enhance the reliability and robustness of anomaly detection methods. Conversely, when the 2D H-IRLS-HT method was applied (due to speed efficiency) to the noisy synthetic dataset, the results exhibited marked improvements in

clarity, quality, and noise reduction, which led to a more accurate representation of the underlying spatial patterns. Additionally, the elimination of outliers significantly enhanced the interpretability of the data, enabling the clear identification and delineation of anomaly boundaries. The superior performance of the 2D H-IRLS-HT and 2D C-IRLS-HT methods underscores their potential as a robust and reliable alternative for anomaly detection, especially in challenging environments where noise and outliers might compromise the effectiveness of conventional techniques.

In the concluding analysis of the research conducted in the eastern part of region Syria, spanning approximately 128 km^2, gravity, and magnetic geophysical measurements were performed by an exploration company. The use of the Scintrex CG gravity meter, paired with GPS locating, allowed for the collection of good measurements. These measurements underwent a series of necessary corrections considering the rock density of the study area. The intention was to utilize these gravitational measurements to yield valuable subsurface geological information through intricate interpretation processes. In the prior stages of my last research, we employed the Surfer software to identify the structure within the dataset. As the research evolved, we introduced a more advanced method, the 2D C-IRLS-FT, particularly in the low-pass filter process. The study systematically applied the 2D C-IRLS-FT method to evaluate the gravity measurement dataset. Our main objective was to examine the effectiveness of this method and its impact on the data before proceeding with any comprehensive analysis using a two-dimensional low-pass Butterworth filter on the data set. The Fourier Transform, using the 2D C-IRLS-FT method, played a pivotal role in this procedure, enabling the decomposition of the gravity data into its frequency components. In essence, the 2D C-IRLS-FT method proved to be an invaluable addition to the analytical toolkit, particularly when calculating the low pass filter. It facilitated a more thorough interpretation of the subsurface structures, thereby contributing to a comprehensive understanding of the geological context of the study area. Proceeding with the investigation, the use of the MINIMAG MMPOS device, a programmable independent magnetic station, was crucial in capturing readings every twenty seconds within a significant range, despite occasional inaccuracies. Subsequent spectral analysis illuminated the necessity of applying the 2D C-IRLS-FT method to a chosen data segment for noise reduction, pole reduction, and overall enhancement of data quality. The successful implementation of this method facilitated a significant decrease in noise levels, effectively improving the interpretability of the dataset. Particularly, the pole reduction process, a standard in magnetic anomalies analysis, proved instrumental in minimizing

the effects of varying magnetic inclinations and declinations, thus ensuring the accuracy of the evaluation of the underlying geological structures and properties. The pole-to-pole reduction technique further amplified the coherence of the geological structures within the dataset, leading to a noticeable improvement in data quality. The elimination of noise and outliers from the data yielded a more distinct representation of the structure, thereby facilitating a more reliable interpretation. As a result, it became possible to identify key geological characteristics and assess the underlying spatial patterns and relationships more accurately. In essence, this study underscores the synergistic impact of noise reduction methods and pole-to-pole reduction techniques in refining the dataset, thereby fostering robust and accurate subsequent interpretations. As an extension of our current research scope, a novel technique that combines the 2D C-IRLS-FT method and k-means clustering was deployed for denoising seismic data. This was specifically undertaken within a seismic section that harboured a significant noise component. This method hinges on the transformation of seismic data into the frequency domain, followed by the identification of substantial spectral components via clustering and, ultimately, the filtering out of noise components. The procedure initiates the transformation of seismic data into the frequency domain via the 2D C-IRLS-FT method. The amplitude spectrum, computed as the absolute value of the transformed seismic data, is then reshaped into a unidimensional array. This reformed data serves as the input feature for the k-means clustering algorithm. To provide a comparative assessment of the denoising technique, the amplitude spectrum of seismic data was computed without applying the proposed approach, thus serving as a benchmark for evaluating the effectiveness of the combined 2D C-IRLS-FT and k-means clustering technique in amplifying the amplitude spectrum and diminishing the noise components. The results of this advanced method underscored its efficacy in denoising seismic data, as evidenced by the side-by-side comparison of the original and denoised seismic sections. The amplitude spectra visualization elucidated a clear demarcation between the signal and noise components, with the filtered data preserving the significant spectral components while eliminating the noise, thereby underscoring the superior performance and potential applicability of this technique.

8. Acknowledgment

First and foremost, I would like to express my deepest gratitude to God for his countless blessings, guidance, and wisdom that have been bestowed upon me throughout this endeavor. I am truly grateful for the strength and perseverance that have carried me through challenging moments.

I would also like to express my heartfelt appreciation to my distinguished professor, Prof. Dr. Mihály Dobróka. The journey of knowledge and discovery under your mentorship has been both transformative and enlightening. Your dedication, profound insights, and unwavering belief in my capabilities have significantly contributed to my intellectual growth and the successful completion of this project. Your tireless guidance and expertise, coupled with your patience and sincerity, have been a source of constant support and inspiration. I am eternally thankful for your mentorship and trust. A special note of gratitude is extended to the President of Department Prof. Dr. Norbert Szabó. Your leadership, wisdom, and commitment to fostering a conducive and supportive learning environment have played a pivotal role in my academic journey. I am greatly appreciative of your support and encouragement.

I cannot express enough thanks to my family for their continual support and unwavering faith in me. The love, understanding, and sacrifices that you have rendered throughout my journey have been my pillar of strength and motivation. To my diamond, your love, patience, and unwavering support have been my guiding light throughout this journey. You have been my constant source of inspiration and motivation, and for that, I am forever grateful.

Lastly, I would like to extend my sincere thanks to my friends, whose constant encouragement, valuable insights, and unwavering friendship have been instrumental in this journey. Yours camaraderie and support have been truly invaluable. In conclusion, I am deeply grateful to everyone who has been part of my journey and helped make this achievement possible. I am reminded that this achievement is not solely mine, but a collective effort of all those who have supported and believed in me. Thank you.

9. REFERENCES

Abdelaziz, M.I., Dobróka, M., 2020. TESTING THE NOISE REJECTION CAPABILITY OF THE INVERSION BASED FOURIER TRANSFORMATION ALGORITHM APPLIED TO 2D SYNTHETIC GEOMAGNETIC DATASETS. Geosci. Eng. 8, 70–82.

Aggarwal, C.C., 2013. Applications of Outlier Analysis, in: Aggarwal, C.C. (Ed.), Outlier Analysis. Springer, New York, NY, pp. 373–400. https://doi.org/10.1007/978-1-4614-6396-2_12

Aki, K., Richards, P.G., 2002. Quantitative seismology.

Al-Heety, E.M.S., Al-Mufarji, M.A., Al Esho, L.H., others, 2017. Qualitative interpretation of gravity and aeromagnetic data in west of Tikrit City and surroundings, Iraq. Int. J. Geosci. 8, 151.

Alsdorf, D., Barazaugi, M., Litak, R., Seber, D., Sawaf, T., AI-Saad, D., others, 1995. The intraplate Euphrates fault system-Palmyrides mountain belt junction and relationship to Arabian plate boundary tectonics.

Alsouki, M., Taifour, R., 2015. The tectono-depositional evolution of the Syrian Euphrates Graben Area using the 3D seismic data. Arab. J. Geosci. 8, 7577–7587.

Askari, R., Siahkoohi, H.R., 2008. Ground roll attenuation using the S and x-f-k transforms. Geophys. Prospect. 56, 105–114.

Auken, E., Christiansen, A.V., Jacobsen, B.H., Foged, N., Sørensen, K.I., 2005. Piecewise 1D laterally constrained inversion of resistivity data. Geophys. Prospect. 53, 497–506. https://doi.org/10.1111/j.1365-2478.2005.00486.x

Backus, G., 1970. Inference from Inadequate and Inaccurate Data, I*. Proc. Natl. Acad. Sci. 65, 1–7. https://doi.org/10.1073/pnas.65.1.1

Backus, G.E., Gilbert, J.F., 1967. Numerical Applications of a Formalism for Geophysical Inverse Problems. Geophys. J. Int. 13, 247–276. https://doi.org/10.1111/j.1365-246X.1967.tb02159.x

Barnes, E., 2000. Attributes For Automating Seismic Facies Analysis | SEG International Exposition and Annual Meeting | OnePetro [WWW Document]. URL https://onepetro.org/SEGAM/proceedings/SEG00/All-SEG00/SEG-2000-0553/87680 (accessed 5.16.23).

Barnett, V., Lewis, T., 1994. Barnett, V., and Lewis T.: Outliers in Statistical Data. 3rd edition. J. Wiley & Sons 1994, XVII. 582 pp., £49.95. Biom. J. 37, 256–256. https://doi.org/10.1002/bimj.4710370219

Berteussen, K.A., Ursin, B., 1983. Approximate computation of the acoustic impedance from seismic data. GEOPHYSICS 48, 1351–1358. https://doi.org/10.1190/1.1441415

Brew, G., Barazangi, M., Al-Maleh, A.K., Sawaf, T., 2001. Tectonic and geologic evolution of Syria. GeoArabia 6, 573–616.

Butler, K., Russell, R., 2003. Cancellation of multiple harmonic noise series in geophysical records. Geophysics 68. https://doi.org/10.1190/1.1581080

Butler, K.E., Russell, R.D., 1993. Subtraction of powerline harmonics from geophysical records. GEOPHYSICS 58, 898–903. https://doi.org/10.1190/1.1443474

Butterworth, C., 1930. Filter approximation theory. Engineer 7, 536–541.

Chopra, S., Marfurt, K.J., 2005. Seismic attributes — A historical perspective. GEOPHYSICS 70, 3SO-28SO. https://doi.org/10.1190/1.2098670

Cohen, L., 1995. Time-Frequency Analysis: Theory and Applications, Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Cosentino, S., Filoramo, P., Granata, A., Marletta, M., Martino, G., Pelleriti, R., Torrisi, F., Paparo, M., Cosentino, G., Vita, P., Palmisano, G., 2001. An integrated RF transceiver for DECT application, in: Proceedings of the IEEE 2001 Custom Integrated Circuits Conference (Cat. No.01CH37169). Presented at the IEEE 2001 Custom Integrated Circuits Conference, IEEE, San Diego, CA, USA, pp. 519–522. https://doi.org/10.1109/CICC.2001.929833

Deighan, A.J., Watts, D.R., 1997. Ground-roll suppression using the wavelet transform. GEOPHYSICS 62, 1896–1903. https://doi.org/10.1190/1.1444290

Dobróka, M., Szegedi, H., Molnár, J.S., Szűcs, P., 2015. On the Reduced Noise Sensitivity of a New Fourier Transformation Algorithm. Math. Geosci. 47, 679–697. https://doi.org/10.1007/s11004-014-9570-x

Dobróka, M., Szegedi, H., Vass, P., 2017. Inversion-Based Fourier Transform as a New Tool for Noise Rejection, Fourier Transforms - High-tech Application and Current Trends. IntechOpen. https://doi.org/10.5772/66338

Dobróka, M., Szegedi, H., Vass, P., Turai, E., 2012. Fourier transformation as inverse problem — An improved algorithm. Acta Geod. Geophys. Hung. 47, 185–196. https://doi.org/10.1556/AGeod.47.2012.2.7

Gabor, D., 1946. Theory of communication. Part 1: The analysis of information. J. Inst. Electr. Eng. - Part III Radio Commun. Eng. 93, 429.

Gauss, C.F., Davis, C.H., Project, M. of A., 1809. Theory of the motion of the heavenly bodies moving about the sun in conic sections a translation of Gauss's "Theoria motus." With an appendix. Little, Brown and company, Boston. https://doi.org/10.5962/bhl.title.19023

Guanlei, X., Xiaotong, W., Lijia, Z., Xiaogang, X., 2018. Image decomposition and texture analysis via combined bi-dimensional Bedrosian's principles. IET Image Process. 12, 262–273. https://doi.org/10.1049/iet-ipr.2017.0494

Guitton, A., Symes, W.W., 2003. Robust inversion of seismic data using the Huber norm. GEOPHYSICS 68, 1310–1319. https://doi.org/10.1190/1.1598124

Haines, S.S., Guitton, A., Biondi, B., 2007. Seismoelectric data processing for surface surveys of shallow targets. Geophysics 72, G1–G8. https://doi.org/10.1190/1.2424542

Havlicek, Joseph, Havlicek, John, T, N., Bovik, A., 2002. Skewed 2D Hilbert Transforms and Computed AM-FM Models. https://doi.org/10.1109/ICIP.1998.723573

Havlicek, J.P., Havlicek, J.W., Mamuya, N.D., Bovik, A.C., 1998. Skewed 2D Hilbert transforms and computed AM-FM models, in: Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269). IEEE, pp. 602–606.

Hilbert, D., 1917. The Foundations of Physics. Nachrichten Von Ges. Wiss. Zu Gött. Math.-Phys. Kl. 1917, 53–76.

Jackson, D.D., 1979. The use of a priori data to resolve non-uniqueness in linear inversion. Geophys. J. Int. 57, 137–157. https://doi.org/10.1111/j.1365-246X.1979.tb03777.x

JeffryesB, 2002. A method of seismic surveying with overlapping shot times. UK Patent 2 387 226.

Kohlmann, K., 1996. Corner detection in natural images based on the 2-D Hilbert transform. Signal Process. 48, 225–234. https://doi.org/10.1016/0165-1684(95)00138-7

Kovesi, P., others, 1999. Image features from phase congruency. Videre J. Comput. Vis. Res. 1, 1–26.

Litak, R.K., Barazangi, M., Brew, G., Sawaf, T., Al-Imam, A., Al-Youssef, W., 1998. Structure and evolution of the petroliferous Euphrates graben system, southeast Syria. AAPG Bull. 82, 1173–1190.

Lorenzo-Ginori, J.V., 2007. An approach to the 2D hilbert transform for image processing applications, in: International Conference Image Analysis and Recognition. Springer, pp. 157–165.

MacQueen, J., 1967. Classification and analysis of multivariate observations, in: 5th Berkeley Symp. Math. Statist. Probability. University of California Los Angeles LA USA, pp. 281–297.

Marashly, O., Dobroka, M., 2021. Hilbert transform using a robust geostatistical method. IOP Conf. Ser. Earth Environ. Sci. 942, 012029. https://doi.org/10.1088/1755-1315/942/1/012029

Maurya, S.P., Singh, K.H., Singh, N.P., 2019. Qualitative and quantitative comparison of geostatistical techniques of porosity prediction from the seismic and logging data: a case study from the Blackfoot Field, Alberta, Canada. Mar. Geophys. Res. 40, 51–71. https://doi.org/10.1007/s11001-018-9355-6

Maurya, S.P., Singh, N.P., 2020. Effect of Gaussian noise on seismic inversion methods.

McCarthy, P., Sayre, J., Shawyer, B., 1993. Generalized legendre polynomials. J. Math. Anal. Appl. 177, 530–537.

Menke, W., 1984. The resolving power of cross-borehole tomography. Geophys. Res. Lett. 11, 105–108. https://doi.org/10.1029/GL011i002p00105

Meunier, J., Bianchi, T., 2002. Harmonic noise reduction opens the way for array size reduction in vibroseis? operations, in: SEG Technical Program Expanded Abstracts 2002, SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 70–73. https://doi.org/10.1190/1.1817354

Moon, W., Ushah, A., Singh, V., Bruce, B., 1988. Application of 2-D Hilbert Transforai in Geophysical Imaging With Potential Field Data. Geosci. Remote Sens. IEEE Trans. On 26, 502–510. https://doi.org/10.1109/36.7674

Nuamah, D.O.B., Dobroka, M., 2019. Inversion-based fourier transformation used in processing non-equidistantly measured magnetic data. Acta Geod. Geophys. 54, 411–424. https://doi.org/10.1007/s40328-019-00266-4

Nuamah, D.O.B., Dobróka, M., Vass, P., Baracza, M.K., 2021. Legendre polynomial-based robust Fourier transformation and its use in reduction to the pole of magnetic data. Acta Geod. Geophys. 56, 645–666. https://doi.org/10.1007/s40328-021-00357-1

Nyman, D.C., Gaiser, J.E., 1983. Adaptive rejection of high?line contamination, in: SEG Technical Program Expanded Abstracts 1983, SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 321–323. https://doi.org/10.1190/1.1893897

Oldenburg, D.W., Li, Y., 2005. Inversion for applied geophysics: A tutorial. -Surf. Geophys. 89–150.

Osborne, J., Overbay, A., 2004. The Power of Outliers (and Why Researchers Should Always Check for Them). Pr. Assess Res Eval 9.

Parker, R.L., 1977. Understanding Inverse Theory. Annu. Rev. Earth Planet. Sci. 5, 35–64. https://doi.org/10.1146/annurev.ea.05.050177.000343

Pei, S.-C., Ding, J.-J., 2003a. The generalized radial Hilbert transform and its applications to 2D edge detection (any direction or specified directions). Presented at the Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on, p. III–357. https://doi.org/10.1109/ICASSP.2003.1199484

Pei, S.-C., Ding, J.-J., 2003b. The generalized radial Hilbert transform and its applications to 2D edge detection (any direction or specified directions), in: 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). IEEE, p. III–357.

Ricker, N., 1953. The form and laws of propagation of seismic wavelets. Geophysics 18, 10–40. https://doi.org/10.1190/1.1437843

Routh, P., Palacharla, G., Chikichev, I., Lazaratos, S., 2012. Full Wavefield Inversion of Time-Lapse Data for Improved Imaging and Reservoir Characterization, in: SEG Technical Program Expanded Abstracts 2012, SEG Technical Program Expanded Abstracts. Society of Exploration Geophysicists, pp. 1–6. https://doi.org/10.1190/segam2012-1043.1

Saucier, A., Marchant, M., Chouteau, M., 2005. A fast and accurate frequency estimation method for canceling harmonic noise in geophysical records. Geophysics 71, V7–V18. https://doi.org/10.1190/1.2159063

Scales, J.A., Gersztenkorn, A., 1988. Robust methods in inverse theory. Inverse Probl. 4, 1071. https://doi.org/10.1088/0266-5611/4/4/010

Sen, M.K., Stoffa, P.L., 2013. Global Optimization Methods in Geophysical Inversion. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511997570

Sommer, G., Bayro-Corrochano, E., Bülow, T., 1997. Geometric algebra as a framework for the perception—action cycle, in: Solina, F., Kropatsch, W.G., Klette, R., Bajcsy, R. (Eds.), Advances in Computer Vision, Advances in Computing Science. Springer, Vienna, pp. 251–260. https://doi.org/10.1007/978-3-7091-6867-7_26

Steiner, F. (Ed.), 1997. Optimum methods in statistics. Akadémiai Kiadó, Budapest.

Szabó, N., 2015. Hydraulic conductivity explored by factor analysis of borehole geophysical data. Hydrogeol J 23, 869–882. https://doi.org/10.1007/s10040-015-1235-4

Szabó, N.P., 2011. Shale volume estimation based on the factor analysis of well-logging data. Acta Geophys. 59, 935. https://doi.org/10.2478/s11600-011-0034-0

Szegedi H., Dobróka M., 2014. On the use of Steiner's weights in inversion-based Fourier transformation: robustification of a previously published algorithm. ACTA Geod. Geophys. 49, 95–104.

Taner, M.T., Koehler, F., Sheriff, R.E., 1979. Complex seismic trace analysis. GEOPHYSICS 44, 1041–1063. https://doi.org/10.1190/1.1440994

Taner, T., Schuelke, J.S., O'Doherty, R., Baysal, E., 1994. Seismic attributes revisited | SEG Technical Program Expanded Abstracts 1994 [WWW Document]. URL https://library.seg.org/doi/abs/10.1190/1.1822709 (accessed 5.16.23).

Turai, E., 2011. Data processing method developments using TAU-transformation of Time-Domain IP data II. Interpretation results of field measured data. Acta Geod. Geophys. Hung. 46, 391–400. https://doi.org/10.1556/AGeod.46.2011.4.2

Vass, P., 2010. THE FOURIER TRANSFORMATION AS AN INVERSE PROBLEM.

Yilmaz, O., 2001. Seismic data analysis. https://doi.org/10.1190/1.9781560801580

Zhdanov, M.S., 2015. Inverse theory and applications in geophysics. Elsevier.