

THE CONCEPT OF VIRTUAL ARRAYS IN SEISMIC DATA ACQUISITION

A Dissertation

by

IBRAHIM ABDULAZIZ I ALHUKAIL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Geophysics

The Concept of Virtual Arrays in Seismic Data Acquisition

Copyright 2012 Ibrahim Abdulaziz I Alhukail

THE CONCEPT OF VIRTUAL ARRAYS IN SEISMIC DATA ACQUISITION

A Dissertation

by

IBRAHIM ABDULAZIZ I ALHUKAIL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Luc T Ikelle
Committee Members,	Mark Everett
	Jerome Schubert
	Yuefeng Sun
Head of Department,	John R. Giardino

August 2012

Major Subject: Geophysics

ABSTRACT

The Concept of Virtual Arrays in Seismic Data Acquisition. (August 2012)

Ibrahim Abdulaziz I Alhukail, B.S., University of Tulsa;

M.S., King Fahd University of Petroleum and Mineral

Chair of Advisory Committee: Dr. Luc T. Ikelle

We are presenting a new way of improving seismic-array responses. By analyzing the relationship between the covariance matrix forms from the real sensors of seismic arrays and the fourth-order crosscumulants from the same sensors, we find that artificial sensors can be constructed from the real sensors. We have called these artificial sensors *virtual sensors* and the combination of real and virtual sensors a *virtual seismic array*. For example, we can construct from an equally weighted linear array of five sensors, a weighted virtual array of nine sensors. Basically, the virtual sensors allow us to introduce new sensors in the seismic arrays as well as new weightings of the existing real sensors. The key assumption behind this approach is that seismic data are considered nonGaussian; hence the fourth-order crosscumulants of the real sensor responses are nonzero.

DEDICATION

I would like to dedicate my success in completing my Ph.D. in Geophysics from Texas A&M University to my beloved parents; GOD bless their souls. I wish they were here today to share with me this great moment!

I also extend my dedication to my beloved wife and children; Layan, Rakan, Danah, Rashid and Majid. Their love and support has always motivated me. I hope my success today inspires and motivates them to seek success for themselves in the future.

Also, I extend it to my brothers, sisters and their children.

ACKNOWLEDGEMENTS

Since the day I was admitted to this University until today, my committee chair, Dr. Luc Ikelle has always been so inspiring. His hard work and numerous achievements have always amazed me. Regardless of the many hard days I went through, I think back, now, and I realize the great progress made in advancing my intellects and technical capabilities. In addition to teaching me to become a specialist in my field of study, Dr. Ikelle also taught me the meaning of thinking and working hard. He taught me the persistence in seeking knowledge and the independence in searching for the solution. These are the fundamental perceptions needed for true and continuous success. No matter how much, I try to thank him today; my words are not enough to reward one of the greatest geophysicists of our modern time.

I also thank Saudi Aramco for sponsoring me during my study at Texas A&M University. I believe that my time, spent here, had helped me to improve not only my technical knowledge, but also my interpersonal skills and my experience.

Finally, I thank my wife for her patience and love. She has been a wonderful wife and a great mother.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	xi
 CHAPTER	
I INTRODUCTION:.....	1
II VIRTUAL SEISMIC ARRAY.....	6
Introduction	6
Formulation of the virtual seismic array	7
The case of linear array	13
The case of areal array.....	14
III MODELING OF ARRAY DATA	20
Introduction	20
Elastodynamic wave equation	21
Pseudo-spectral finite-difference modeling.....	27
Comparison of staggered-grid and pseudo-spectral finite-difference modeling techniques.....	36
Numerical example (groundroll)	64
IV CONCLUSIONS AND RECOMMENDATIONS.....	82
REFERENCES.....	84
APPENDIX A	86

APPENDIX B.....	90
APPENDIX C.....	121
APPENDIX D	123

LIST OF FIGURES

FIGURE		Page
1	A virtual array obtained from a uniformly spaced linear array of five sensors.	16
2	A comparison of real and virtual array responses.....	16
3	A real array of 16 sensors.....	17
4	The virtual array of the real array in Figure 3.....	18
5	An areal virtual array: 49 sensors (33 virtual sensors and 16 real sensors).....	18
6	The weights of the virtual array in Figure 5.....	19
7	Array responses of the virtual array (red) versus the real array (blue).....	19
8	Element volume and illustration of stress components in the x-direction, on opposite faces of the volume.....	23
9	A 2D geology model used to generate seismic data in Figures 11 and 12...	34
10	Raypaths of seismic events in Figures 11 and 12.	35
11	A shot-gather generated with the staggered-grid finite-difference modeling.....	35
12	A shot-gather generated with the pseudo-spectral finite-difference modeling.	36
13	Staggered-grid and pseudo-spectral finite-difference of 4-meter, 2-meter and 1-meter grid spacings for a homogenous model.	45
14	The same as Figure 13 with large scaling.....	46
15	Crosscorrelation records for similar grid-spacing from both techniques.....	48

FIGURE	Page
16	The crosscorrelation records in Figure 15, after normalization..... 51
17	The normalized crosscorrelation in Figure 16, with the <i>negative-dB</i> values (top) were analyzed aside from the <i>positive- dB</i> value (bottom)... 53
18	Six different runs generated by the pseudo-spectral finite-difference for the 4m grid-spacing of the same model, but with different timesteps ranging from 0.5-0.00625ms..... 56
19	A thin layer of 60m thick imbedded in a homogenous medium..... 61
20	A shot-record for the model in Figure 19, generated by the staggered-grid finite-difference..... 62
21	The same shot-record as in Figure 20, except the source bandwidth was broadened, which violated the grid-dispersion condition..... 62
22	A seismic shot-record, generated by the pseudo-spectral finite-difference, for the model in Figure 19..... 63
23	Homogenous medium of 500m thick, overlain by 200m thick air-filled layer..... 65
24	A snapshot recorded at 0.25s for the air-filled layer model showing a P-wave direct arrival and a P-wave reflection of the air-filled layer, which acted as an artificial free-surface..... 66
25	A snapshot recorded at 0.25s for the homogenous layer model showing a P-wave direct arrival and a P-wave reflection from the free-surface..... 66
26	Seismic shot-record of one-second long, showing a P-wave direct arrival and a P-wave reflection from the air-filled layer, which acted as an artificial free-surface..... 67
27	Seismic shot-record of one-second long, showing a P-wave direct arrival and a P-wave reflection from the free-surface..... 67
28	A snapshot recorded at 0.4s for homogenous medium overlain by an inclined air-filled layer that acted as an artificial free-surface..... 69

FIGURE		Page
29	A snapshot recorded at 0.32s for homogenous medium overlain by a wiggly air-filled layer that acted as an artificial free-surface.....	69
30	A snapshot, recorded at one second, showing the vertical component of the particle-velocity for a homogenous layer below a free-surface.....	72
31	A snapshot, recorded at 0.8 second, showing the vertical component of the particle velocity, for the same model as in Figure 30, with the exception of adding a 10m thick layer below the free-surface.....	73
32	A shot-record, of one-second long, showing the vertical component of the particle velocity, for the same model as in Figure 30, with the exception of adding a 10m thick layer below the free-surface.....	73
33	A land model made of (4096X2048) spatial cells that consisted of 7 layers.....	75
34	A zoom-in of the land model in Figure 33 that shows the random wiggles of the air-filled layer and the random distribution of the elastic parameters of the thin layer.	76
35	A seismic short-record, simulating seismic land data, of two-second long sampled at 4ms generated by the pseudo-spectral finite-difference.....	78
36	A large-scale model of several wiggly layers, along with its elastic parameters to be used as the input model for the pseudo-spectral finite-difference.....	79
37	A seismic shot-record simulated with the pseudo-spectral finite-difference for the model in Figure 36.	80

LIST OF TABLES

TABLE		Page
1	A comparison of the staggered-grid and the pseudo-spectral finite-difference techniques, with respect to the stability and grid-dispersion conditions.....	37
2	Finite-difference modeling parameters.....	43
3	Six runs of the same model generated by the pseudo-spectral finite-difference with the 4.0m grid-spacing.	55
4	The implications of three different elastic model parameters on the parameters of both the finite-difference techniques.....	58

CHAPTER I

INTRODUCTION

The goal of petroleum seismology is to image the subsurface geology for oil and gas exploration and production. Petroleum seismology has two components: (1) seismic data acquisition, which aims to produce an earth response from manmade wave sources and (2) imaging of the recorded data, which aims to produce a model of the subsurface geology. The focus of this thesis is on seismic data acquisition.

In seismic data acquisition, receivers are used to record the earth response. In practice, a receiver is an array of sensors whose number can vary between 6 and 24. So, seismic responses are collected at each sensor of the array and then summed to produce the seismic response associated with one receiver (or seismic trace). This summation is aimed to improve the signal-to-noise ratio (Ikelle and Amundsen, 2005). Our objective in this thesis is to develop a concept that may further improve the signal-to-noise ratio of the array responses, reduce the number of sensors used in a seismic array or both, by constructing additional artificial sensors from the actual sensors (real sensors). We have called these artificial sensors “virtual sensors.” Moreover, we have called the combination of the real and virtual sensors a “virtual seismic array”.

The basic idea behind the concept of a virtual seismic array, which will be described in

This dissertation follows the style and format of Geophysics.

this thesis, is that seismic data, like many real-life signals and processes, are non-Gaussian rather than Gaussian. In other words, the analyses of seismic responses in this thesis will not be limited to the classical second-order statistics tools like covariances. We will also consider higher-order statistics tools like crosscumulants and autocumulants. In this thesis, we will carry out analyses of array responses based on fourth-order statistics. By comparing the current analyses of array responses based on second-order statistics with new analyses based on fourth-order statistics, we will reveal the construction of the virtual sensors from the real sensors.

The next step in my research is to numerically evaluate the concept of virtual seismic arrays. I decided to use the finite-difference modeling tool, which is the most accurate way to generate seismic data (Ikelle and Amundsen, 2005). The finite-difference modeling tool allows us to model diffraction, reflection, multiple diffractions and reflections at the free surface and in the subsurface as well as surface waves, including groundroll, airwaves, etc. In our group, we have a finite-difference modeling code, which works in the time-space (T-X) domain. This code is based on a technique known as the staggered-grid finite-difference, initially developed by Raul Madariaga for modeling dynamic earthquake rupture, then Virieux used it for modeling seismic wave propagation due to a point source (Madariaga, 1976; Virieux, 1986). However, this code approximates spatial derivative by a fourth-order difference operator, instead of a second-order difference operator, as in the case for Madariaga and Virieux. Approximating spatial derivative by a fourth-order difference operator provides more accuracy, efficiency and requires less computer memory than the second-order

differencing (Levander, 1988). Previous theses in our group have demonstrated the effectiveness of this code in simulating seismic data (Singh, 2005; Wilson, 2002).

However, there is one significant difference between data that I need to simulate in order to evaluate virtual seismic arrays and the ones simulated before. I need to simulate data at the array level instead of the receiver level. This requirement implies that the spacing between receivers is no longer 12.5 or 25 meters, as in previous simulations; rather it is on the order of 1 meter or less. To illustrate the effects of generating seismic data at the array level on other finite-difference parameters imposed when the staggered-grid finite-difference is used, I will consider the following example. Suppose, for example, that we have a typical geology model with velocities ranging between 200 *m/s* and 5000 *m/s* and a maximum frequency of 60 *Hz*. The staggered-grid finite-difference criteria for selecting the time interval between snapshots, Δt , and the grid spacing of the geology, Δx , (assuming square-gridding of the geology), are given by two formulas. The first one is the grid-dispersion formula, which aims to reduce velocity dispersion that results when truncating the Taylor series that are used for approximating the spatial derivatives. The condition for avoiding the grid dispersion is related to the number of grid points per wavelength. It requires $\Delta x < \frac{V_{min}}{5 * f_{max}}$, where V_{min} is the minimum velocity traveling within the medium and f_{max} is the maximum source frequency (Ikelle and Amundsen, 2005). The second formula is referred to as the stability condition. The staggered-grid finite-difference equations characterizing the wave motion are computed recursively, timestep by timestep. For example, computing the components of the particle velocity at

timestep $(n + \frac{1}{2})$ and at timestep $(n + 1)$ for the stress components requires the previous timestep $(n - \frac{1}{2})$ of the components of particle velocity and the timestep (n) of the stress components. However, this recursive computation (timestep by timestep) can be a source of numerical instability. In fact, errors introduced by the numerical solution can propagate and be magnified during the timestepping of the finite-difference scheme, causing significant instabilities during the computation and artifacts in the resulting data. Such an instability is very unlikely to occur if the ratio between the temporal and spatial sampling interval is constrained as follows: $\Delta t < 0.606 \frac{\Delta x}{V_{max}}$, where V_{max} is the maximum wave speed in the 2D model (Ikelle and Amundsen, 2005). These constraints make the staggered finite-difference not the best modeling tool for simulating seismic data at the array level, when the receiver spacing is on the order of 1 meter or less. Therefore, the objective of Chapter III of this thesis is to investigate an alternative finite-difference technique that can be used to simulate seismic data with receiver spacing on the order of 1 meter or less and can be less sensitive to the grid spacing requirement. Such alternative finite-difference modeling tool could be the wavenumber-based finite difference known also as Pseudo-spectral method (Kosloff et al., 1984; Kosloff and Baysal, 1982). This finite-difference technique uses a spatial numerical grid to compute the spatial derivatives using the fast Fourier transform instead of the fourth-order differencing as is the case for the staggered-grid finite-difference. While time derivatives are still approximated by second-order differencing, just as in the case of the staggered-grid finite-difference. Thus, this may make this finite-difference modeling

tool less sensitive to the grid dispersion imposed when the staggered-grid finite-difference is used. Therefore, maybe more suitable for modeling seismic data at the array level where spacing among sensors is on the order of one meter or less.

The remaining of this thesis is organized into three chapters. In Chapter II, I will derive the virtual seismic array concept and show, for various E&P array configurations, how the locations of virtual sensors can be predicted. To further my study of virtual array so that it can, one day, be used in practice, I need to construct appropriate synthetic data. In Chapter III, I describe possible choices of numerical modeling techniques, mainly finite-difference techniques, for generating synthetic seismic data at the array level. In Chapter IV, I will draw conclusions on the present state of my investigation and provide some recommendations for future work.

CHAPTER II

VIRTUAL SEISMIC ARRAY

INTRODUCTION

A receiver is actually an array of sensors whose number can vary between 6 and 24. So seismic responses are collected at sensors of the array and then summed to produce the seismic response associated with one receiver (or seismic trace). In a number of acquisition systems, the summation is hardwired in such a way that wavefronts recorded by sensors at time t are directly summed, irrespective of the data quality or some sensor malfunctioning. Although very efficient in terms of acquisition turnaround, these types of acquisition systems are prone to errors ranging from noise leakage due to aliasing to improper summation due to some malfunctioning sensors.

An alternative acquisition system, which is more and more commonly adopted today, is to record the whole array of sensors for a certain length of time, filter the noise and aliased data, and correct for any potential sensor malfunctioning before summing the seismic sensors to produce seismic traces.

We here present an additional processing step before forming the arrays. The basic idea is to construct additional sensors from the real sensors. The ability to construct these additional sensors can help us to either improve the resolution of the array response, to reduce the number of sensors used in the seismic array, or both.

FORMULATION OF THE VIRTUAL SEISMIC ARRAY

Suppose that we have I statistically independent signals impinging on an array of L sensors (also known as elements). The array response of this array can be written as follows:

$$D_l(t) = \sum_{k=1}^I S_k(t - \tau_{lk}) \quad (2.1)$$

where $D_l(t)$ is the signal output of the l -th sensor of the array, $S_k(t)$ is the k -th signal response, and τ_{lk} is the propagation delay between the first sensor (that we here consider as the reference) and the l -th sensor for the k -th signal. We are going to rewrite equation (2.1) in terms of the complex envelope of $D_l(t)$ and $S_k(t)$. If $\tilde{D}_l(t)$ and $\tilde{S}_k(t - \tau_{lk})$ are the complex envelopes of $D_l(t)$ and $S_k(t - \tau_{lk})$, respectively, then (2.1) can be written:

$$\tilde{D}_l(t) = \sum_{k=1}^I \tilde{S}_k(t - \tau_{lk}) \quad (2.2)$$

Our next task is to recast equation (2.2) into a series of linear equations in which the linear coefficients are independent of time. We found that we can do so by decomposing our seismic data $D_l(t)$ into a series of narrow-band signals by using, for example, the filter-bank technique (Harris, 2004). A filter-bank is an array of band-pass filters that separates the input signal into several components, each one carrying a single-frequency

subband of the original signal. It is desirable to design the filter-bank in such a way that subbands can be recombined to recover the original signal. The first process is called analysis, and the second is called synthesis. The filter-bank can be used to decompose a wideband signal into different narrow-band signals. This decomposition is useful here because our derivations assume that the incoming signals to the sensors are narrow-band. So we can carry on our derivation with narrow-band signals and assume that this process is repeated for all the narrow-band components of the original signals.

Using the fact that we are dealing with narrow-band signals, we can use the results, which say that $\tilde{S}_k(t - \tau_{lk})$ is a phase shift of $\tilde{S}_k(t)$, to arrive at

$$\tilde{D}_l(t) = \sum_{k=1}^I \exp^{i\omega_c \tau_{lk}} \tilde{S}_k(t) \quad (2.3)$$

where $\tilde{S}_k(t)$ is the complex envelope of $S_k(t)$ and ω_c is its central angular frequency. So we now have an expression for which the mixing coefficients are time-independent. We can write (2.3) in the standard form of linear mixtures as

$$\tilde{D}_l(t) = \sum_{k=1}^I a_{kl} \tilde{S}_k(t) \quad (2.4)$$

with

$$a_{kl} = \exp^{i\omega_c \tau_{lk}} \quad (2.5)$$

The result in (2.4) and (2.5) is only valid for narrow-band signals. For wideband signals, we can use the filterbank to decompose $D_l(t)$ into narrowband signals with each narrowband signal having its own central frequency.

To facilitate the discussion about τ_{lk} , we consider the case of a linear, uniform sensor array, which is the typical choice in modern seismic acquisition. Because we are working at the array level, we can approximate the waves impinging on the array as plane waves. Hence the propagation delays between the reference and the l -th sensor for signal k can be written as follows:

$$\tau_{lk} = \frac{(l-1)\Delta x \sin \alpha_k}{V}, \quad (2.6)$$

where Δx is the distance between two consecutive sensors in an array, α_k is the direction of waves associated with the single-shot response $S_k(t)$, and V is the velocity of the medium in which sensors are located. We will assume that this medium is homogeneous and acoustic. It is customary to work with the intersensor phase-shift θ_k which is

$$\theta_k = \omega_c \frac{(l-1) \Delta x \sin \alpha_k}{V} = 2\pi \frac{(l-1) \Delta x \sin \alpha_k}{\lambda_c} \quad (2.7)$$

where $\lambda_c = f_c/V$, rather than the incidence angle α_k , because the wavelength has small variations around the central wavelength $\lambda_c = f_c/V$. Using the intersensor phase-shift (2.3) becomes as:

$$\tilde{D}_l(t) = \sum_{k=1}^l \exp[i(l-1)\theta_k] \tilde{S}_k(t) \quad (2.8)$$

which can be written in the standard form of linear mixtures as

$$\tilde{D}_l(t) = \sum_{k=1}^l a_{kl} \tilde{S}_k(t) \quad (2.9)$$

with

$$a_{kl} = \exp[i(l-1)\theta_k] \quad (2.10)$$

or in the vector-matrix form as follows:

$$\tilde{D}(t) = A\tilde{S}(t) \quad (2.11)$$

where

$$A = [a(\theta_1), a(\theta_2), \dots, a(\theta_K)] \quad (2.12)$$

where $\tilde{D}(t)$ describes an L -dimensional vector of the array responses, $\tilde{S}(t)$ represents a I -dimensional vector of the signal responses, and A represents the mixing matrix, whose size is $L \times I$ where

$$\mathbf{a}(\theta) = [1, \exp(-i\theta), \dots, \exp(-i(L-1)\theta)]^T \quad (2.13)$$

Let us now use the statistical formulation to derive our formulation of the virtual sensors. It is based on the linear relationship between the covariance matrix D and that S (Ikelle, 2010)

$$\underbrace{C_D^{(2)}}_{L \times L} = \underbrace{A}_{L \times I} \underbrace{C_S^{(2)}}_{I \times I} \underbrace{A^H}_{I \times L} \quad (2.14)$$

where $C_D^{(2)}$ and $C_S^{(2)}$ are the covariance matrices of D and S , respectively, and $(A)^H$ denotes the complex conjugate transpose. Notice that $C_D^{(2)}$ is a $L \times L$ matrix, whereas $C_S^{(2)}$ is a $I \times I$ matrix. Alternatively, the covariance matrix and fourth-order cumulant tensor of D are related to those of S as follows (Ikelle and Amundsen, 2005):

$$C_D^{(2)} = \sum_k^K C_S^{(2)}(k, k) \mathbf{a}(\theta_k) \mathbf{a}^H(\theta_k) \quad (2.15)$$

The fourth-order cross-cumulant is

$$C_D^{(4)} = \sum_k^K C_S^{(4)}(k, k, k, k) X [a(\theta_k) \otimes \bar{a}(\theta_k)] \cdot [a(\theta_k) \otimes \bar{a}(\theta_k)]^H \quad (2.16)$$

where \otimes is the Kronecker product (see Appendix C). Let us now use the statistical formulation to derive our virtual sensors. The covariance matrix and fourth-order cross-cumulant tensor of D are related to those of S, as follows (Ikelle and Amundsen, 2005):

$$C_D^{(2)} = \sum_k^K C_S^{(2)}(k, k) a(\theta_k) a^H(\theta_k) \quad (2.17)$$

The fourth-order crosscumulant is

$$C_D^{(4)} = \sum_k^K C_S^{(4)}(k, k, k, k) X [a(\theta_k) \otimes \bar{a}(\theta_k)] \cdot [a(\theta_k) \otimes \bar{a}(\theta_k)]^H \quad (2.18)$$

where \otimes is the Kronecker product and the subscript H denotes the Hermitian transpose. We can note that the $L \times L$ matrix $C_D^{(2)}$ and the $L^2 \times L^2$ matrix $C_D^{(4)}$ have the same algebraic structure. The auto-cumulant $C_S^{(4)}(k, k, k, k)$ and the vector $[a(\theta_k) \otimes \bar{a}(\theta_k)]$ play in $C_D^{(4)}$ the roles that $C_S^{(2)}(k, k)$ and $a(\theta_k)$ play for $C_D^{(2)}$, respectively. Thus the L^2 -vector $[a(\theta_k) \otimes \bar{a}(\theta_k)]$ can be considered to be the equivalent or virtual steering vector for the k -th signal for a true array of L sensors with coordinates X_l . So, there are actually L^2 sensors with L of them being real sensors and the others are virtual sensors.

THE CASE OF LINEAR ARRAY

For a linear array, it is always possible to choose a coordinate system in which the l -th sensor has coordinates in the form $(X_l, 0, 0)$ with $1 \leq l \leq L$. The vectors $\mathbf{a}(\theta_k)$ for a linear array are:

$$\mathbf{a}(\theta_k) = \exp\left\{i \frac{2\pi}{\lambda} X_l \cos(\theta_k)\right\} \quad (2.19)$$

Note that $m = r + L(q - 1)$, where $1 \leq r, q \leq L$. The m -th component of the vector $[\mathbf{a}(\theta_k) \otimes \bar{\mathbf{a}}(\theta_k)]$, which is denoted $[\mathbf{a}(\theta_k) \otimes \bar{\mathbf{a}}(\theta_k)]_m$ or $[\mathbf{a}(\theta_k) \otimes \bar{\mathbf{a}}(\theta_k)]_{r,q}$ can be written:

$$[\mathbf{a}(\theta_k) \otimes \bar{\mathbf{a}}(\theta_k)]_{r,q} = \exp\left\{i \frac{2\pi}{\lambda} [(X_r - X_q) \cos(\theta_k)]\right\} \quad (2.20)$$

Comparing (2.18) with (2.19), we deduce that the vector $[\mathbf{a}(\theta_k) \otimes \bar{\mathbf{a}}(\theta_k)]$ can also be considered a true steering vector of the k -th signal for the equivalent or virtual array of L^2 equivalent or virtual sensors with coordinates $X_{r,q} = (X_r - X_q)$, $1 \leq r, q \leq L$, which introduces the virtual array concept for the fourth-order direction finding problem.

Thus, the virtual array of L identical sensors has $L_V = 2L - 1$. This result shows that the virtual array is also a linear array.

For a uniformly spaced linear array, $X_l = (l - 1)\Delta x$, where Δx is the interelement spacing, the virtual array is composed of the sensors whose coordinates are given by $x_{r,q} = x_r - x_q = (r - q)\Delta x$, $1 \leq r, q \leq L$. This result shows that the virtual

array is also a uniformly spaced linear array. Thus the virtual array of L identical sensors has $L_V = 2L - 1$ different sensors. In other words, with a uniformly spaced linear array of L sensors, the fourth-order direction-finding methods are able to process $L_V - 1 = 2(L - 1)$ independent sources. Figure 1 shows an illustration of this case for $L = 5$.

Notice that the sensor of the virtual array at the coordinate $x_{r,q}$ has a multiplicity of order $L - |r - q|$, which is then equivalent to a sensor that is weighted in amplitude by a factor $L - |r - q|$. Thus the virtual array of a uniformly spaced linear array with L sensors is an amplitude-tapered uniformly spaced linear array of $2L - 1$ sensors. This amplitude tapering explains, in particular, the reason why the bandwidth of the virtual array is not twice as narrow as that of the real array, despite the fact that the physical size of the virtual array is two times greater than that of the real array. Figure 2 provides an illustration of the virtual-array response of an initial array of five uniformly spaced sensors of equal weights.

THE CASE OF AREAL ARRAY

For an areal arrays, it is always possible to choose a coordinate system in which the l -th sensor has coordinates in this form $(X_l, Y_l, 0)$ with $1 \leq l \leq L$. The vectors $\mathbf{a}(\theta_k)$ for an areal array are the following:

$$\mathbf{a}(\theta_k) = \exp \left\{ i \frac{2\pi}{\lambda} [X_l \cos(\theta_k) + Y_l \sin(\theta_k)] \right\} \quad (2.21)$$

The m^{-th} component of the vector $[a(\theta_k) \otimes \bar{a}(\theta_k)]$, which is denoted $[a(\theta_k) \otimes \bar{a}(\theta_k)]_{r,q}$ for an areal array is obtained by the following:

$$\begin{aligned}
 [a(\theta_k) \otimes \bar{a}(\theta_k)]_{r,q} &= \exp \left\{ i \frac{2\pi}{\lambda} \left[[(X_r - X_q) \cos(\theta_k)] + [(Y_r - Y_q) \sin(\theta_k)] \right] \right\} \quad (2.22)
 \end{aligned}$$

where $(X_r - X_q, Y_r - Y_q)$ is the coordinates of the r, q sensor.

Note that the L pairs r, q such that $r = q, 1 \leq r \leq L$, give the same virtual sensor of coordinates $(0,0,0)$. We will say that this virtual sensor is of multiplicity L .

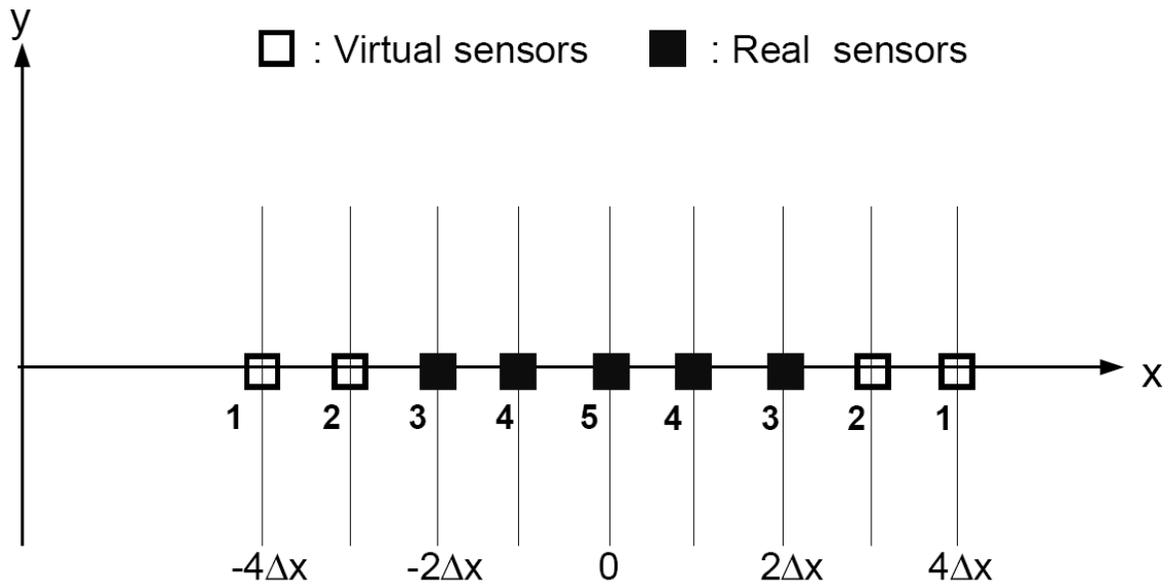


Figure 1. A virtual array obtained from a uniformly spaced linear array of five sensors. Notice that the virtual array is now weighted in amplitude while the original array is not.

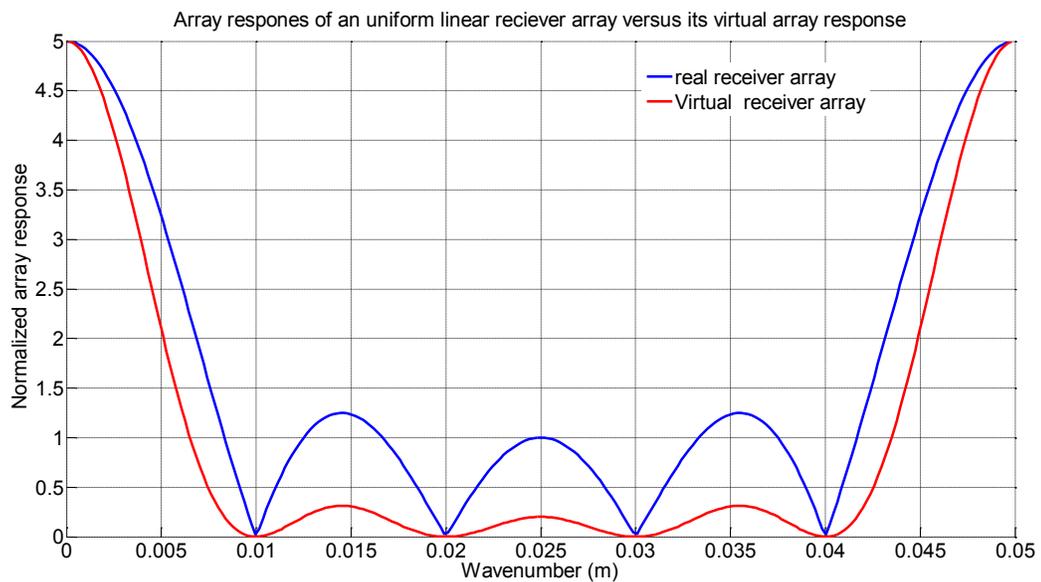


Figure 2. A comparison of real and virtual array responses. We have the linear array described in Figure-1 for this comparison. Response of a real array of uniformly spaced five sensors (blue) and the response of the corresponding virtual array (red).

The example is for an areal array of 16 sensors as described in Figure 3. By using the concept that we have just described, we can construct 256 virtual sensors from these 16 real sensors as shown in Figure 4. By comparing these two figures, we can see that a number of virtual sensors share the same locations with the real sensors. Actually, only 33 new sensor locations are created in this case. The resulting virtual array has 49 sensors (33 virtual sensors and 16 real sensors) as depicted in Figure 5. The other 223 virtual sensors affect the weights of sensors of the virtual array as illustrated in Figure 6 and through the array responses in Figure 7.

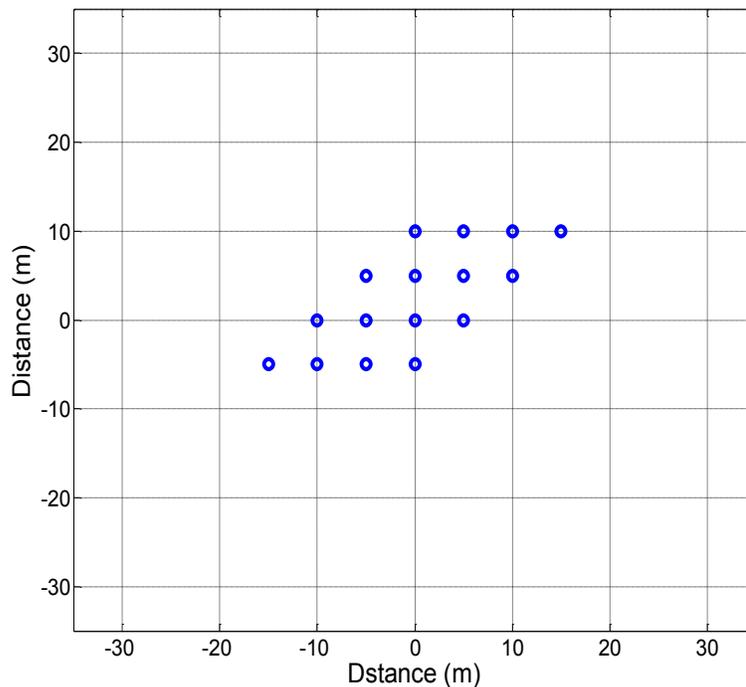


Figure 3. A real array of 16 sensors.

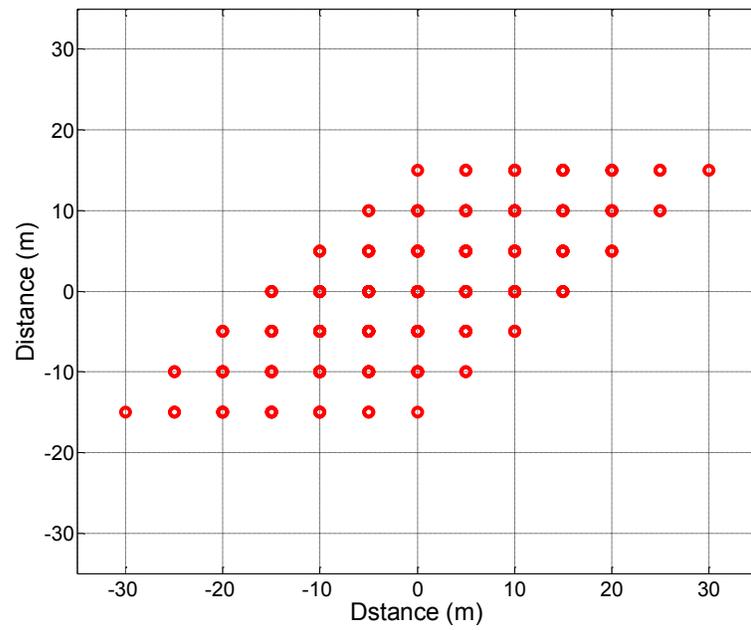


Figure 4. The virtual array of the real array in Figure 3.

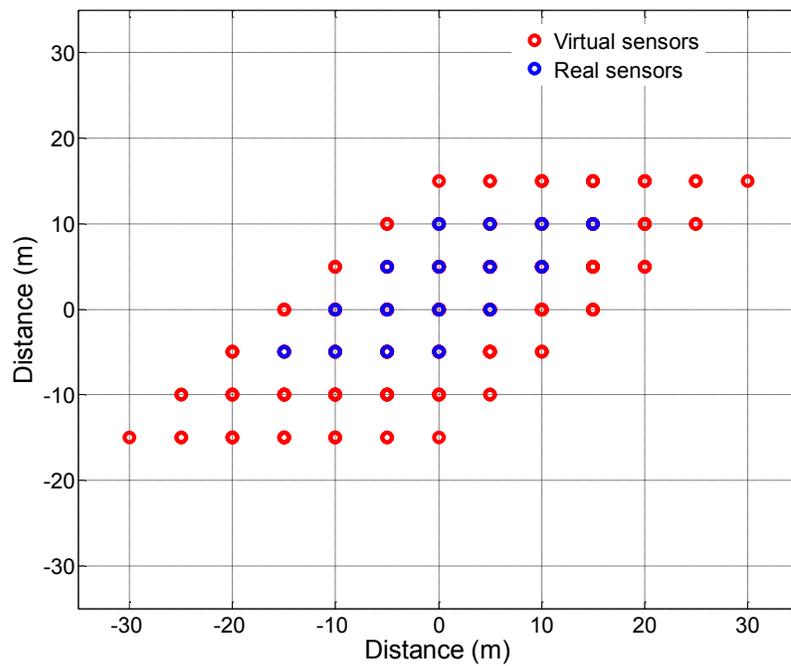


Figure 5: An areal virtual array: 49 sensors (33 virtual sensors and 16 real sensors). Notice that weights of sensors of the virtual array are not included here.

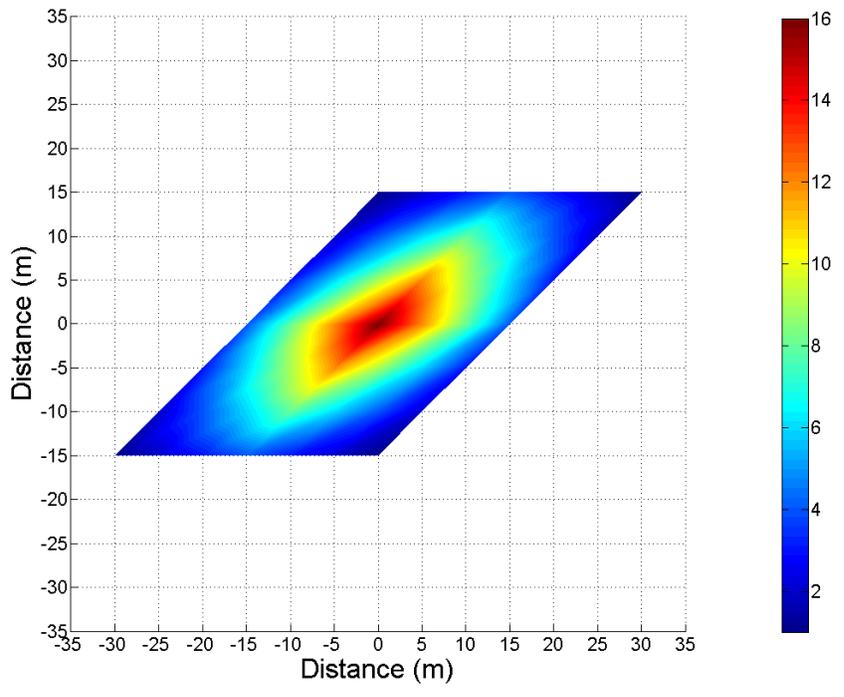


Figure 6. The weights of the virtual array in Figure 5.

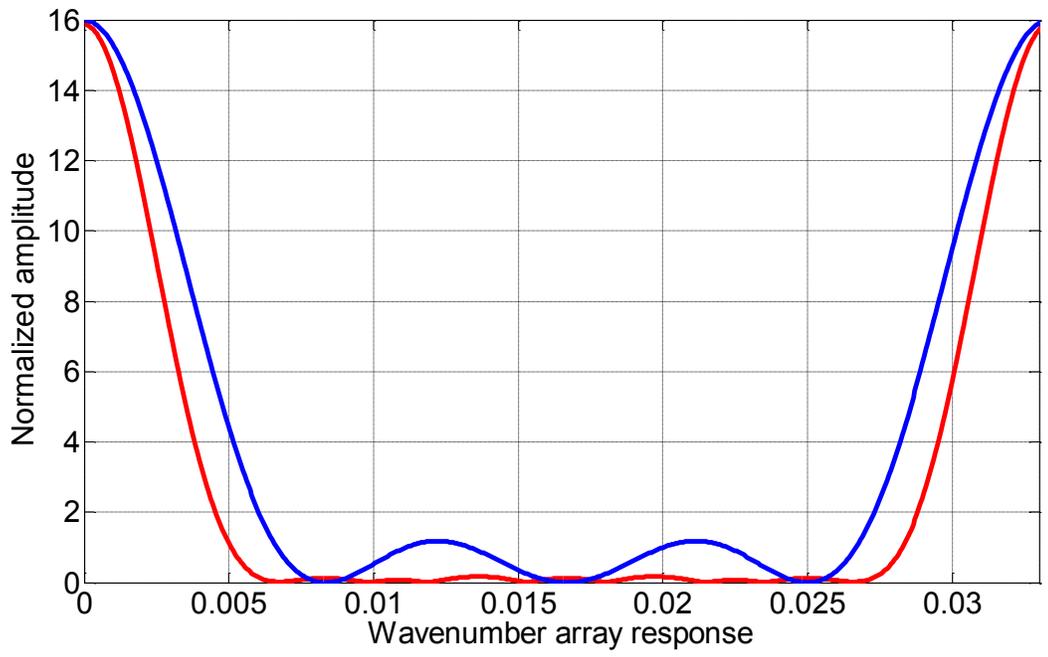


Figure 7. Array responses of the virtual array (red) versus the real array (blue). The virtual array used here is shown in Figure 5 and the corresponding real array is shown in Figure 3.

CHAPTER III

MODELING OF ARRAY DATA

INTRODUCTION

My objective in this chapter is to investigate a way of modeling seismic data, as accurately as possible, at the array level. This implies that the sensor spacing is of 1 meter or less. I selected to work with the finite-difference modeling technique because it is, by far, the most accurate technique for modeling seismic data. The classical formulation of the finite-difference modeling, known as the staggered-grid finite-difference, is carried-out in the time-space (T-X) domain. It basically consists of approximating the spatial derivatives of the partial differential of the wave equations by the first four terms of the Taylor series and the time derivative by the first two terms of the Taylor series. The Taylor series terms used in these computations are chosen in the staggered-grid formulation as proposed by (Madariaga, 1976) to avoid differentiating the geological model.

For modeling array data, with sensor spacing of 1 meter, the staggered-grid finite-difference becomes costly prohibitive because the grid-spacing of the geology model has to be four times, or more, smaller than the sensor-spacing. Here, I investigate the use of the pseudo-spectral finite-difference (Kosloff and Baysal, 1982) as an alternative solution to model array data. The pseudo-spectral finite-difference technique is carried-out in the time-wavenumber (T-K) domain. We now compute the spatial

derivatives of the partial differential of the wave equations by using the Fourier transform.

The organization of this chapter is as follows: in the second section, I will recall the partial differential equation that control wave propagation in the subsurface, known as the elastodynamic wave equation. The finite-difference modeling tools are based on this equation. In the third section, I will review the pseudo-spectral finite-difference formulation. In the fourth section, I will provide a table, which is based on theoretical aspects of comparing the staggered-grid with the pseudo-spectral finite-difference techniques, in terms of their stability and grid dispersion conditions requirements. Then, I will discuss the implications of having different elastic parameters on the finite-difference parameters for both techniques. In the last section, I will discuss a numerical example of using the pseudo-spectral finite-difference to model array data for land seismic.

ELASTODYNAMIC WAVE EQUATION

The elastodynamic wave equation, which governs wave propagation in the subsurface, is derived from Newton's second law of motion and Hooke's law. Newton's second law of motion, also known as the equation of conservation of momentum, states that force equals the product of mass and acceleration. Hooke's law states that the stress is linearly proportional to strain. The components of force, on the other side, are derived from the components of the stress acting on particles treated as elemental volumes. Therefore, the stresses are described as a set of surface forces acting on faces of the

elemental volume. Three-component surface stress or traction is required for each face of the elemental volume. For instance, Figure-1 shows an example of an elemental volume centered in three dimensional Cartesian coordinates. The traction stress acting on the area element facing +x direction is given by $T_x = T_x(\mathbf{x}, t) = (\tau_{xx}, \tau_{yx}, \tau_{zx})^T$. In these components, the second subscript refers to the direction of the normal vector of the surface, the traction is acting on. Similarly, the traction stresses on unit-area elements facing the +y and +z directions are given by $T_y = (\tau_{xy}, \tau_{yy}, \tau_{zy})^T$ and $T_z = (\tau_{xz}, \tau_{yz}, \tau_{zz})^T$ respectively. The stress components in any given direction are those that are resolved from the different stress tractions. Thus, in the x -direction for example, they are given by these components τ_{xx} , τ_{xy} , τ_{xz} . So, in Figure 8, the stress component in the x -direction on the face a-b-c-d is τ_{xx} , and on its opposite face, e-f-g-h, it is $(\tau_{xx} + \frac{\partial \tau_{xx}}{\partial x} \Delta x) - \tau_{xx} = \frac{\partial \tau_{xx}}{\partial x} \Delta x$. The resulting force is the stress multiplied by the surface area on which it acts on ($\Delta y \Delta z$). Hence, the resulting force is $\frac{\partial \tau_{xx}}{\partial x} \Delta x \Delta y \Delta z$. The stress component in the x -direction on the other faces are given by $\frac{\partial \tau_{xy}}{\partial y} \Delta y \Delta x \Delta z$ and $\frac{\partial \tau_{xz}}{\partial z} \Delta z \Delta x \Delta y$. Therefore, the total force component in the x -direction is given by $f_x = \left(\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \right) \Delta x \Delta y \Delta z$. Substituting f_x into Newton's law of motion, where the mass is given by $m = \rho \Delta x \Delta y \Delta z$ and the acceleration in the x -direction is given by $\gamma_x = \frac{\partial^2 u_x}{\partial t^2}$, where u_x is the displacement in x -direction and t is time. This yields to the first equation of motion in x -direction.

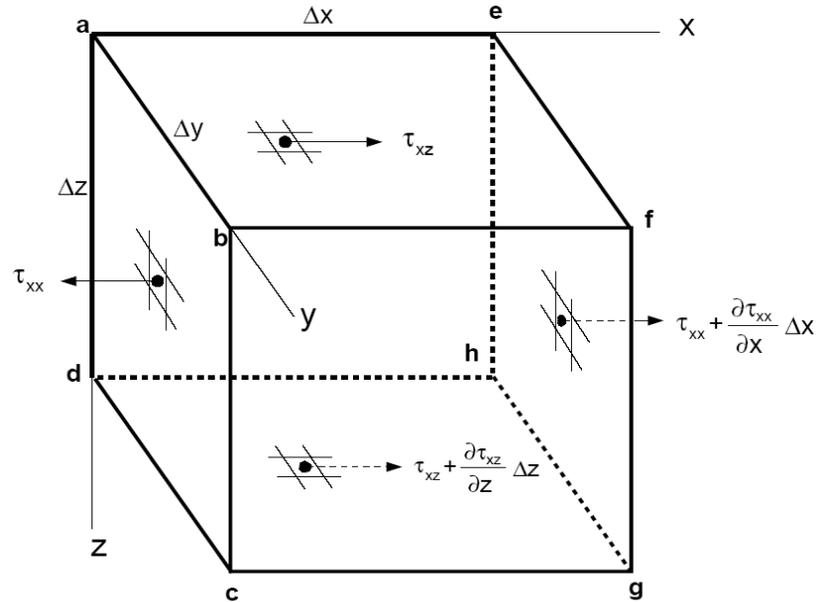


Figure 8. Element volume and illustration of stress components in the x-direction, on opposite faces of the volume (Ikelle and Amundsen, 2005).

$$\rho \frac{\partial^2 u_x}{\partial t^2} = \left(\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \right) + F_x(\mathbf{x}, t) \quad (3.1)$$

Similarly, stresses in the y- and z-directions lead to force components in the y- and z-directions, which would produce two other equations of motion in y- and z-directions respectively.

$$\rho \frac{\partial^2 u_y}{\partial t^2} = \left(\frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} \right) + F_y(\mathbf{x}, t) \quad (3.2)$$

$$\rho \frac{\partial^2 u_z}{\partial t^2} = \left(\frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \right) + F_z(\mathbf{x}, t)$$

In these equations, ρ is the mass per unit volume, $\mathbf{u} = (u_x, u_y, u_z)$ are the components of the particle displacement, $\tau = (\tau_{xx}, \tau_{yy}, \tau_{zz}, \tau_{xy}, \tau_{xz}, \tau_{yz})$ are the components of the stress field and $\mathbf{F} = (F_x, F_y, F_z)$ are the components of the body force per unit volume. (3.1) and (3.2) state that spatially changing surface forces with or without the presence of the body force cause local forces in the medium that give rise to local particle accelerations.

For a given point, \mathbf{x} at time t , Hooke's law states that stress is linearly proportional to strain. This linear relationship is formulated as follow:

$$\tau_{ij}(\mathbf{x}, t) = c_{ijkl}(\mathbf{x})e_{kl}(\mathbf{x}, t) + \mathbf{I}_{ij}(\mathbf{x}, t), \quad \text{where } i, j, k, l = x, y, z \quad (3.3)$$

The constant of proportionality, $c_{ijkl} = c_{ijkl}(\mathbf{x})$, are known as “elastic moduli” or “stiffness constants”. They define the elastic properties or resistance to deformation. The elastic moduli c_{ijkl} , may have up to 21 different parameters for anisotropic material, but for an isotropic material they reduce into two parameters known as the Lamé elastic coefficients. $\mathbf{I}_{ij}(\mathbf{x}, t)$ represent the external stresses and $e_{kl}(\mathbf{x}, t)$ are the strain. Strain is related to particle displacement by the following relationship:

$$e_{ij}(\mathbf{x}, t) = \frac{1}{2} \left(\frac{\partial u_i(\mathbf{x}, t)}{\partial x_j} + \frac{\partial u_j(\mathbf{x}, t)}{\partial x_i} \right), \text{ where } i, j = x, y, z \quad (3.4)$$

For the isotropic case, the elastic moduli c_{ijkl} as function of the Lamé elastic coefficients is generalized as follow:

$$c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) \quad (3.5)$$

The Lamé elastic coefficients $\mu = \mu(x)$ measures a material's resistance to shear (i.e. changes in shape without changes in volume) and $\lambda = \lambda(x)$ has no simple physical meaning, but simplifies Hooke's law greatly and δ_{ij} is the Kronecker delta function defined as:

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases} \quad i, j = 1, 2, \quad (3.6)$$

The relation between stress and strain for isotropic case produces the following equations.

$$\begin{aligned} \tau_{xx} &= (\lambda + 2\mu) \frac{\partial u_x}{\partial x} + \lambda \left(\frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z} \right) + I_{xx}(\mathbf{x}, t) \\ \tau_{yy} &= (\lambda + 2\mu) \frac{\partial u_y}{\partial y} + \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z} \right) + I_{yy}(\mathbf{x}, t) \end{aligned} \quad (3.7)$$

$$\tau_{zz} = (\lambda + 2\mu) \frac{\partial u_z}{\partial z} + \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) + I_{zz}(\mathbf{x}, t)$$

$$\tau_{yz} = \mu \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) + I_{yz}(\mathbf{x}, t)$$

$$\tau_{zx} = \mu \left(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z} \right) + I_{zx}(\mathbf{x}, t)$$

$$\tau_{xy} = \mu \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) + I_{xy}(\mathbf{x}, t)$$

In these equations, $(\tau_{xx}, \tau_{yy}, \tau_{zz})$ are the stress components, (u_x, u_y, u_z) are the displacement components, $(x, y, z)^T$ are the positions and λ and μ are the Lamé elastic coefficients and $(I_{xx}, I_{yy}, I_{zz}, I_{yz}, I_{zx}, I_{xy})$ are the components of the stress force. By combining Hooke's law and the equations of wave motion, we derive the elastodynamic wave equation that governs wave propagation in the subsurface.

$$\rho(\mathbf{x}) \frac{\partial^2 u_i(\mathbf{x}, t)}{\partial t^2} - \frac{\partial}{\partial x_j} \left[c_{ijkl}(\mathbf{x}) \frac{\partial u_k(\mathbf{x}, t)}{\partial x_l} \right] = F_i(\mathbf{x}, t) + \frac{\partial I_{ij}(\mathbf{x}, t)}{\partial x_j} \quad (3.8)$$

where

$$i, j, k, l = x, y, z$$

$\mathbf{F} = \mathbf{F}_i(\mathbf{x}, t)$ and $\mathbf{I} = \mathbf{I}_{ij}(\mathbf{x}, t)$ are the source terms and $\mathbf{u} = u_i(\mathbf{x}, t)$ is the particle displacements.

PSEUDO-SPECTRAL FINITE-DIFFERENCE MODELING

The pseudo-spectral finite-difference, also known as the wavenumber-based finite-difference, is an alternative method to numerically model seismic data. In this scheme, the partial differential of the wave equations temporal derivatives are approximated by a second-order difference operator, similar to the staggered-grid finite-difference. However, derivatives with respect to the spatial coordinates are no longer approximated by differencing (Abramowitz and Stegun, 1972); rather computed by spatially Fourier-transforming the physical quantities, like the stresses and the components of the particles velocities, into the wavenumber domain. The resulted spectrums are then multiplied by a complex wavenumber for a particular direction; depending on the direction of the spatial derivatives. The spatial derivatives are then obtained by applying the inverse Fourier-transform to go back to the spatial domain. This process uses the fact that differentiation in the spatial domain is equivalent to multiplying the spectrum, in the wavenumber domain, with a complex wavenumber. For instance, the Fourier-transform of any arbitrary function, $g(x)$, in the spatial domain is defined as follows:

$$g(k_x) = \int_{-\infty}^{\infty} g(x) \exp(-ik_x x) dx \quad (3.9)$$

where $g(k_x)$ is the spectrum in the wavenumber domain with respect to the x – axis, after it has been Fourier-transformed. The inverse Fourier-transform is defined as follows:

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(k_x) \exp(ik_x x) dk_x \quad (3.10)$$

Rather than defining new symbols to express the functions after they have been Fourier-transformed, the same symbols with different arguments were used, as the context unambiguously indicates the quantity currently under consideration. Derivation in the spatial domain is equivalent in the wavenumber domain to the followings:

$$\begin{aligned}
\frac{\partial g(x)}{\partial x} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(k_x) \frac{\partial}{\partial x} [\exp(ik_x x)] dk_x \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} g(k_x) ik_x \exp(ik_x x) dk_x \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} (ig(k_x)k_x) \exp(ik_x x) dk_x \\
\frac{\partial g(x)}{\partial x} &= F^{-1}(ig(k_x)k_x)
\end{aligned} \tag{3.11}$$

Therefore, differentiation in the spatial domain is equivalent to taking the inverse Fourier-transform of the wavenumber spectrum, $g(k_x)$, after it has been multiplied by the complex wavenumber, ik_x .

Time-wavenumber Formulation

The first-order equations of conservation of momentum for 2D in the time-wavenumber domain (T- \mathbf{K}) are:

$$\begin{aligned}
\rho(\mathbf{x}) \partial_t v_x(\mathbf{x}, t) - F^{-1}\{ik_x \tau_{xx}(\mathbf{k}, t) + ik_z \tau_{xz}(\mathbf{k}, t)\} &= f_x(\mathbf{x}, t) \\
\rho(\mathbf{x}) \partial_t v_z(\mathbf{x}, t) - F^{-1}\{ik_x \tau_{xz}(\mathbf{k}, t) + ik_z \tau_{zz}(\mathbf{k}, t)\} &= f_z(\mathbf{x}, t)
\end{aligned} \tag{3.12}$$

In these equations, $\rho(\mathbf{x})$ is the density, $\mathbf{v} = (v_x, v_z)$ are the components of the particle velocity, F^{-1} is the inverse Fourier-transformed, $\boldsymbol{\tau} = (\tau_{xx}, \tau_{zz}, \tau_{xz})$ are the stress components, $\mathbf{f} = (f_x, f_z)$ are the components of the body force and (ik_x, ik_z) are the complex wavenumber in the x and z directions respectively. The stress-strain relations for an isotropic elastic medium in the time-wavenumber domain (T-K) are:

$$\begin{aligned} \partial_t \tau_{xx}(\mathbf{x}, t) &= [\lambda(\mathbf{x}) + 2\mu(\mathbf{x})]F^{-1}(ik_x v_x(\mathbf{k}, t)) \\ &\quad + \lambda(\mathbf{x})F^{-1}(ik_z v_z(\mathbf{k}, t)) + I_{xx}(\mathbf{x}, t) \\ \partial_t \tau_{zz}(\mathbf{x}, t) &= [\lambda(\mathbf{x}) + 2\mu(\mathbf{x})]F^{-1}(ik_z v_z(\mathbf{k}, t)) \\ &\quad + \lambda(\mathbf{x})F^{-1}(ik_x v_x(\mathbf{k}, t)) + I_{zz}(\mathbf{x}, t) \\ \partial_t \tau_{xz}(\mathbf{x}, t) &= \mu(\mathbf{x})F^{-1}(ik_z v_x(\mathbf{k}, t)) + F^{-1}(ik_x v_z(\mathbf{k}, t)) + I_{xz}(\mathbf{x}, t) \end{aligned} \tag{3.13}$$

The density, the lamè coefficients and the physical quantities in the pseudo-spectral finite-difference are specified at each grid points; rather than staggered, as in the staggered-grid finite-difference. For each timestep, the algorithm starts by computing the wavenumber in the x and z directions respectively and turning them into complex vectors. Then, the values of the normal and shear stress fields in the z direction ($\tau_{zz}\tau_{xz}$) are extracted and Fourier-transformed to the wavenumber domain. The same step is also repeated for the ($\tau_{xx}\tau_{xz}$) in x direction. The resulted spectrums are multiplied by their corresponding complex wavenumber vectors and Fourier-transformed back to the spatial

domain to obtain the spatial derivatives. The components of the particle velocity are then updated by adding the particle velocity from the previous timestep to the spatially differentiated stresses. Numerically, this step is formulated as follows:

$$\begin{aligned}
 [v_x]_{i,k}^{n+1/2} &= [v_x]_{i,k}^{n-1/2} \\
 &\quad + [\Delta t b_x (F^{-1}(ik_x F(\tau_{xx})) + F^{-1}(ik_z F(\tau_{xz})) + f_x)]_{i,k}^n
 \end{aligned}
 \tag{3.14}$$

$$\begin{aligned}
 [v_z]_{i,k}^{n+1/2} &= [v_z]_{i,k}^{n-1/2} \\
 &\quad + [\Delta t b_z (F^{-1}(ik_x F(\tau_{xz})) + F^{-1}(ik_z F(\tau_{zz})) + f_z)]_{i,k}^n
 \end{aligned}$$

where F and F^{-1} stand for the Fourier transform and the inverse Fourier transform respectively. Similarly, the stresses are numerically formulated as follow:

$$\begin{aligned}
[\tau_{xx}]_{i,k}^{n+1} &= [\tau_{xx}]_{i,k}^n \\
&\quad + \Delta t [(\lambda + 2\mu)F^{-1}(ik_x F(v_x)) + (\lambda)F^{-1}(ik_z F(v_z))] \\
&\quad + I_{xx}]_{i,k}^{n+1/2} \\
[\tau_{zz}]_{i,k}^{n+1} &= [\tau_{zz}]_{i,k}^n \\
&\quad + \Delta t [(\lambda + 2\mu)F^{-1}(ik_z F(v_z)) + (\lambda)F^{-1}(ik_x F(v_x))] \\
&\quad + I_{zz}]_{i,k}^{n+1/2}
\end{aligned} \tag{3.15}$$

$$\begin{aligned}
[\tau_{xz}]_{i,k}^{n+1} &= [\tau_{xz}]_{i,k}^n \\
&\quad + \Delta t \left[\mu_{xz} \left(F^{-1}(ik_z F(v_x)) + \left(F^{-1}(ik_x F(v_z)) \right) \right) \right] \\
&\quad + I_{xz}]_{i,k}^{n+1/2}
\end{aligned}$$

In comparison of the above formulations (3.14) and (3.15), which outline the pseudo-spectral finite-difference algorithm for each timestep, to the formulations of the staggered-grid finite-difference, (Ikelle and Amundsen, 2005), it is clear that now all spatial derivatives at each grid-point are computed independently from the neighboring grid-points. This makes the derivatives more representative to their locations and eliminates the weight-averaging effect associated with the use of neighboring grid-points to approximate the derivatives by the differencing scheme for each grid-point. The other

benefit of this scheme is that spatial derivatives are no longer estimated grid-point by grid-point; rather computed all in one single step. All grid-points that pertain to a particular direction are extracted, transformed to the wavenumber domain, multiplied by a complex wavenumber vector for that direction and then transformed back to the spatial domain. This makes algorithm more efficient and requires less computation to achieve comparable results.

Numerical Example

Here, I use the staggered-grid and the pseudo-spectral finite-difference techniques to simulate a seismic shot-record from the same model. The objective is to show that the staggered-grid and the pseudo-spectra finite-difference techniques are both capable of modeling all type of seismic events that can be present in real seismic shot-records. This example is a one-dimensional heterogeneous model, where the elastic parameters change only along the z-axis, (Ikelle and Amundsen, 2005). The model consists of homogenous medium, 500 meters thick, imbedded between two half-spaces. The top half-space represents an acoustic medium and the bottom half-space represents an elastic medium. Absorbing boundary condition, (Cerjan et al., 1985) is applied on all sides and edges of the model, including the free-surface, in order to limit the number of reflections that can be modeled and interpreted. An explosive source was placed in the center laterally and at 240 meters above the imbedded homogenous layer. Two kilometers of offset along each side of the source were populated with pressure sensors, at 4 meters apart, and placed at the same depth as the source, Figure 9. Two-second long

seismic shot-record, with sampling rate of 4.0 millisecond and 25 Hertz central frequency for the source, was generated. The shot-record shows P-wave direct arrival, various P-wave primaries and internal multiple reflections from the imbedded medium top and bottom interfaces as well as refractions. A sketch, which shows the seismic events that were generated from this model, is shown in Figure 10. The seismic shot-records are shown in Figure 11 and Figure 12 for the model in Figure 9, simulated by the staggered-grid and the pseudo-spectral finite-difference techniques, respectively. Observing the two shot-records shows hardly any difference between the two shot-records, which reconfirm that the pseudo-spectral finite-difference is also very power numerical tool to model seismic data.

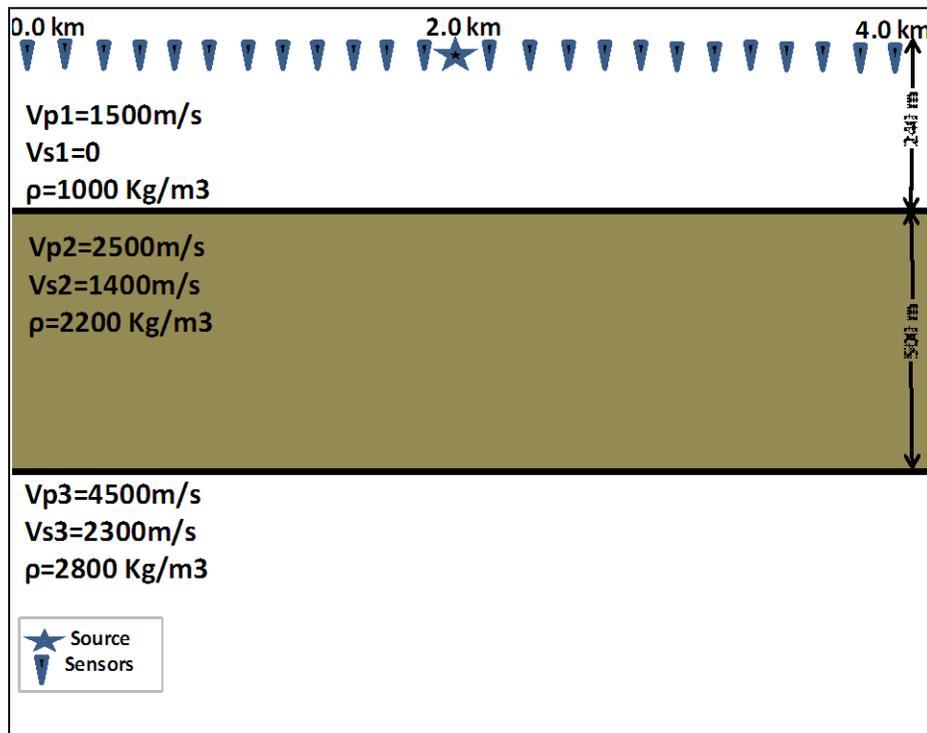


Figure 9. A 2D geology model used to generate the seismic data in Figures 11 and 12.

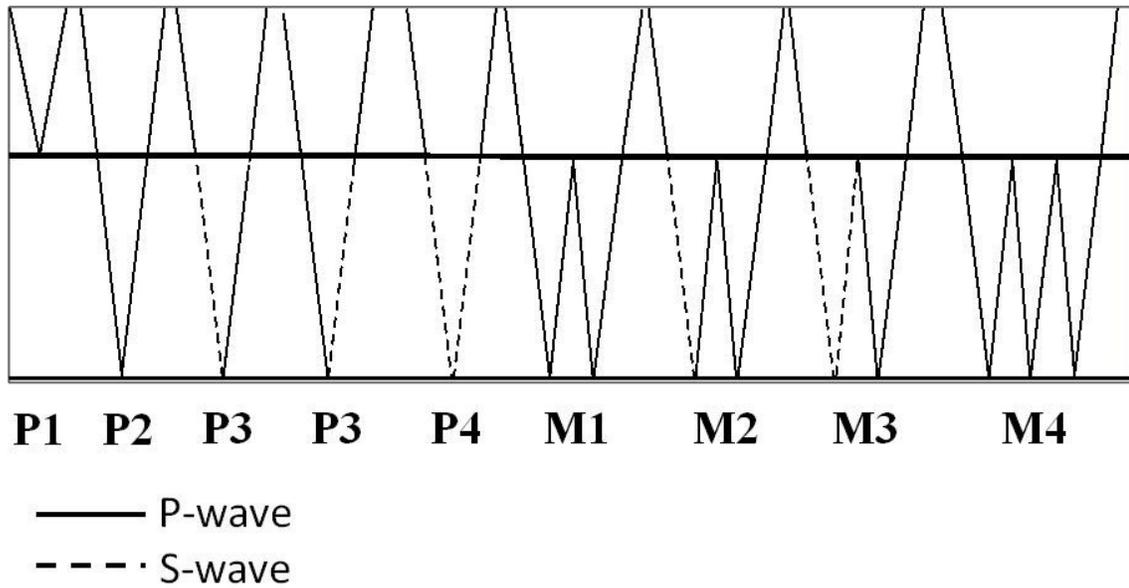


Figure 10. Raypaths of the seismic events in Figures 11 and 12.

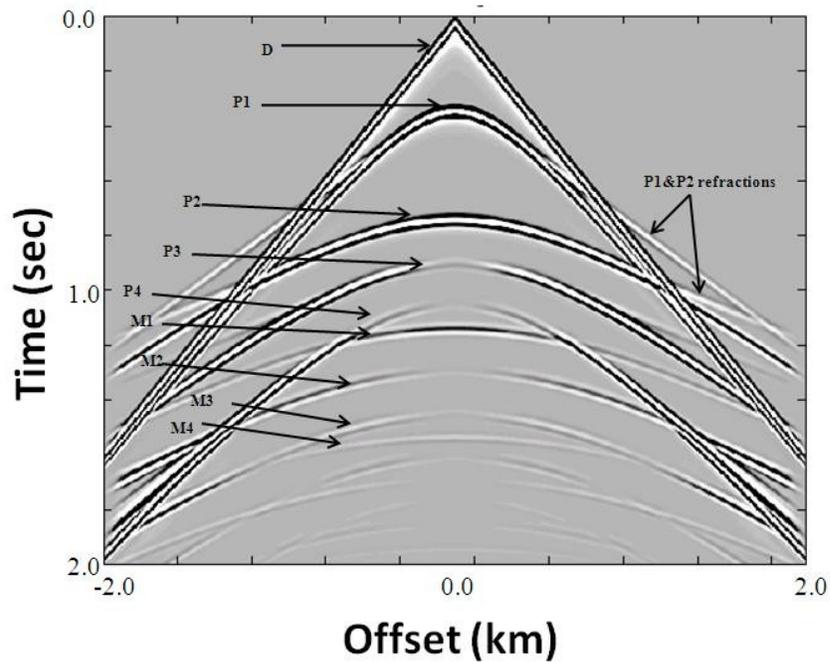


Figure 11. A shot-gather generated with the staggered-grid finite-difference modeling. The geology model is of the acquisition geometry associated with this shot-gather is described in Figure 9. The nomenclature of the seismic events in this figure is also given in Figure 10.

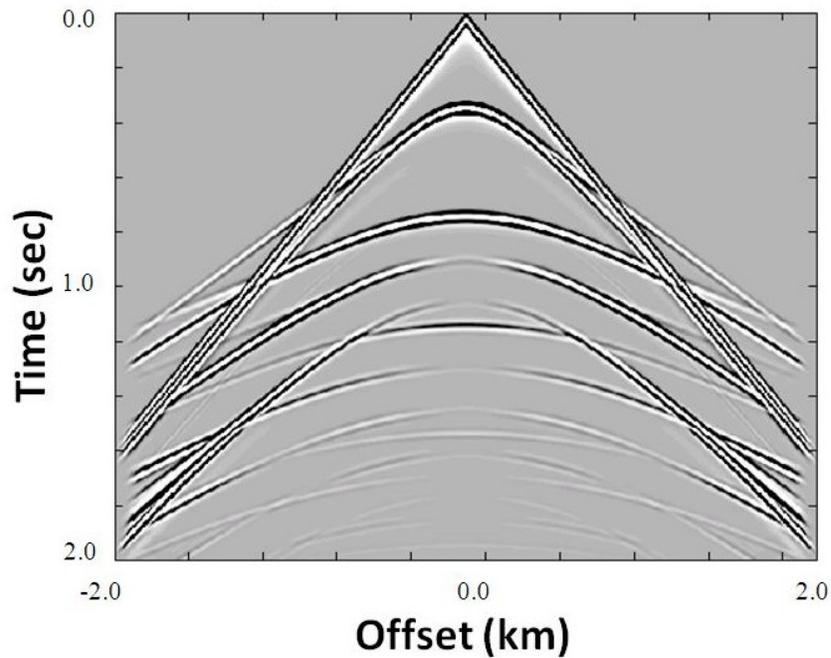


Figure 12. A shot-gather generated with the pseudo-spectral finite-difference modeling. The geology model is of the acquisition geometry associated with this shot-gather is described in Figure 9.

COMPARISON OF STAGGERED-GRID AND PSEUDO-SPECTRAL FINITE-DIFFERENCE MODELING TECHNIQUES

In this section, I compare the staggered-grid and pseudo-spectral finite-difference techniques. The key features of my comparison are summarized in Table 1. Basically, my comparison is carried out with respect to errors associated with the timestepping process, which is known as the stability condition, and with respect to the errors associated with the Taylor series approximation, known as the grid-dispersion condition. Finite-difference parameters such as the size of the model in grid-points, grid-spacing, timesteps and the number of timesteps are key parameters that need to be carefully chosen, such that the stability and grid-dispersion conditions are satisfied.

Table 1. A comparison of the staggered-grid and the pseudo-spectral finite-difference techniques, with respect to the stability and grid-dispersion conditions. Δt is the timestep of the finite-difference, Δx is the grid-spacing for the finite-difference, V_{max} is the maximum velocity of the medium, V_{min} is the minimum velocity of the medium, and f_{max} is the maximum frequency of the source.

Theoretical comparison	Staggered-grid finite-difference	Pseudo-spectral finite-difference
Spatial derivatives	Fourth-order difference operator	Fast Fourier-transform
Temporal derivatives	Second-order difference operator	Second-order difference operator
Stability condition	$\Delta t < 0.606 \frac{\Delta x}{V_{max}}$	$\Delta t < 0.4502 \frac{\Delta x}{V_{max}}$
Grid-dispersion	$\Delta x < \frac{V_{min}}{5f_{max}}$	$\Delta x < \frac{V_{min}}{2f_{max}}$

The staggered-grid finite-difference is carried out in (T-X) domain, in which both the temporal and spatial partial derivatives are carried out with the Taylor series approximation,(Yang, 2008). The pseudo-spectral finite-difference is carried-out in (T-K) domain, in which the partial spatial derivatives are calculated by Fourier transforms while the temporal derivatives are carried out in with Taylor series approximation, just like the staggered-grid finite-difference, (Gazdag, 1981; Kosloff et al., 1984; Kosloff and Baysal, 1982). In other words, the key difference between the two methods lies in the computation of the spatial derivatives. Therefore, the grid-dispersion condition, which is associated with errors due to the truncation of the Taylor series in the computation of the spatial derivatives, is different between the two techniques, as described in Table 1.

The grid-dispersion condition of the staggered-grid finite-difference, in this Table is based on a fourth-order difference operator, (Ikelle, 2005; Levander, 1988;

Virieux, 1986). In contrast, by computing the spatial derivatives with Fourier transforms, as done in pseudo-spectral finite-difference, an infinite order of accuracy can be achieved. More precisely, only two grid-points per wavelength are required in the pseudo-spectral finite-difference, to satisfy the grid-dispersion, whereas 5 grid-points per wavelength are required for the staggered-grid finite-difference, with a fourth-order difference operator. The two grid-points per wavelength, required in the grid-dispersion condition for the pseudo-spectral finite-difference, results from the Nyquist sampling theorem, which is a seismic acquisition requirement, to avoid aliasing the wavefield, rather than a requirement of the pseudo-spectral finite-difference.

On the other side, the staggered-grid and pseudo-spectral finite-difference techniques, both, use a second-order difference operator, which is only the first two terms of the Taylor series, to approximate the temporal derivatives. Therefore, the stability conditions, which control the timestep of the finite-difference, for both finite-difference techniques, are expected to be identical. However, the stability condition of the staggered-grid finite-difference is different from the stability condition of the pseudo-spectral finite-difference, as described in Table 1. Actually, the timestep, for the pseudo-spectral finite-difference, has to be smaller than the timestep, for the staggered-grid finite-difference, for models with identical maximum velocities and grid-spacing, as can be noticed from their formulas in Table 1. The difference between the two stability conditions comes from the fact that the stability condition is related the grid-dispersion condition. This relation manifests itself from the fact that the stability condition is a function of the grid-spacing and the grid-spacing is obtained from the grid-dispersion

condition. Therefore, in the staggered-grid finite-difference, when a second-order difference operator is used to approximate the spatial derivatives, the timestep obtained from the stability condition, which is given by $\Delta t < 0.707 \frac{\Delta x}{v_{max}}$ (Virieux, 1986), becomes even larger than the timestep resulting from the stability condition, when a fourth-order difference operator is used. Therefore, when there are fewer terms of the Taylor series used, to approximate the spatial derivatives, the less accurate the spatial derivatives becomes, the larger the timesteps are that result from the stability conditions. On other hand, when there are more terms of the Taylor series used, to approximate the spatial derivatives, the more accurate the spatial derivatives become, the smaller the timesteps are that results from the stability conditions. Therefore, because the pseudo-spectral finite-difference uses the Fourier-transforms to compute the spatial derivatives, which has high order of accuracy, the spatial derivatives are very accurate, thus the stability condition is expected to be more constrained than the staggered-grid finite-difference, despite the fact that both finite-difference techniques use a second-order difference operator to approximate the time derivatives. However, the difference between the two stability conditions is relatively small, when compared to the difference in the grid-dispersion conditions, for both finite-difference techniques; the staggered-grid and pseudo-spectral finite-difference.

Failing to satisfy the stability condition, for both the finite-difference techniques, is one of the most common sources of errors because the stability conditions described in Table 1, for the staggered-grid and pseudo-spectral finite-difference techniques, were derived for homogenous mediums,(Ikelle and Amundsen, 2005). However, when

modeling heterogeneous mediums, I found that selecting timesteps, which are smaller than what the stability conditions required, can overcome errors that might be associated with the fact that the stability conditions, described in Table 1, were derived for homogenous models.

Numerical Comparison

As discussed in the previous section, the spatial derivatives computed in the pseudo-spectral finite-difference have high-order of accuracy, thus only two grid-points per wavelength is required for very accurate simulation. On the other side, spatial derivatives computed in the staggered-grid finite-difference, which depends on the number of the Taylor series terms used by the difference operator, have limited order of accuracy, thus five or more grid-points per length may be required. However, satisfying the grid-dispersion conditions, for the staggered-grid finite-difference, which was described in Table 1, was only for avoiding dispersion errors in the numerical results. Hence, satisfying the grid-dispersion condition, in the staggered-grid finite-difference, might not be enough, and may necessitate more grid-points per wavelength or in other words, smaller grid-spacing, to improve the accuracy of the spatial derivatives. On the other side, spatial derivatives computed in the pseudo-spectral finite-difference are highly accurate by only satisfying the stability condition.

In this section, my objective is to compare the staggered-grid and pseudo-spectral finite-difference techniques, with respect to their accuracy and efficiency to model seismic data, at the array level. So, the accuracy aspect of this comparison is

investigated by simulating shot-records for the same model, by both finite-difference techniques, but with different grid-spacing. Then their differences can be evaluated by comparing their correlation energy and time delays from their crosscorrelations. On the other side, the efficiency aspect of this comparison is investigated by considering different medium elastic parameters and discussing their implications on selecting the staggered-grid and pseudo-spectral finite-difference parameters. The objective is to highlight the benefits of the pseudo-spectral finite-difference in scenarios pertaining to modeling array data.

In the process of evaluating the best approach to investigate the accuracy of both finite-difference techniques, I realized that using complex and heterogeneous models were very challenging, in cases where the shot-records needed to be regenerated with different grid-spacing, because their proportionalities were hard to maintain. For example, I concluded that modeling wiggly interfaces with different grid-spacing generate differences in the shot-records, as a consequence of the wiggly interface becoming coarser with larger grid-spacing and smoother with finer grid-spacing models. Therefore, the coarser the grid-spacing, the more scattering and diffractions are introduced in the numerical results, and the finer the grid-spacing, the less scattering and diffractions are introduced in the numerical results. Thus, numerical results would suffer from more differences that would affect the analysis. Therefore, I found that analyzing a simple homogeneous model, which can be simulated over various grid-spacing for both finite-difference techniques, is an effective and informal approach to analyze their differences. Also, conclusions could be safely extended to more complex models.

Therefore, my analysis was conducted over a model, which was made of homogenous medium that had P- and S-wave velocities of 2700 *m/s* and 1400 *m/s* respectively and density of 1600 kilograms/meter³. A vertical seismic source, with 10 Hertz central frequency placed in the center laterally and at a depth of 401 meters from the top of the model, was used to generate P- and S-wavefronts. Sensors that recorded the vertical component of the particle-velocity field were placed at a depth of 405 meters from the top of the model with four-meter sensor-spacing. Absorbing boundaries were applied over all edges of the model, including the free-surface to minimize artifact and free-surface reflections introduced into the simulations. For each technique, the staggered-grid and the pseudo-spectral, three simulations of the same model but with different grid-spacing were generated; four- two- and one-meter. Each time the grid-spacing is decreased, the size of the model had to increase. This was important so that the proportionality of the model's physical size and all other parameters like the source and receivers' physical locations were maintained. The objective was to isolate all other changes in parameters except the way each finite-difference technique computed the spatial derivatives. So, in the case of the two-meter grid-spacing, the model had to double and quadruple for the one-meter grid-spacing in relative to the four-meter grid-spacing. Also, the timestep had to decrease for each decrease in grid-spacing to ensure the stability condition is satisfied. Table 2 summarizes the size of the mode, for each grid-spacing, and the other parameters effected as a result of changing the grid-spacing. For each simulation, the stability and grid-dispersion conditions (Table 1) were well satisfied. The timesteps, for similar grid-spacing of both techniques, were kept the same

for consistency. Seismic-records of one-second long for each grid-spacing and for each finite-difference technique were generated that showed a P- and S-waves arrival, Figure 13.

Table 2. Finite-difference modeling parameters.

Grid-spacing	4 meters	2 meters	1 meters
Size of the model	(1024X512)	(2048X1024)	(4096X2048)
Timestep	5.0×10^{-4}	2.5×10^{-4}	1.0×10^{-4}
Number of timesteps	2000	4000	10000
Number of grid-points (absorbing boundaries)	100	200	400
Total record time	1 <i>second</i>	1 <i>second</i>	1 <i>second</i>

While these seismic records looked alike to some extent, plotting the amplitude several times, showed that there were differences, Figure 14. The seismic-records produced by the pseudo-spectral finite-difference looked superior over the ones produced by the staggered-grid finite-difference. They were not affected by any grid-dispersion or scattering regardless of the high post in their amplitude. On the other side, some grid-dispersion appeared in the coarser grid-spacing, for the shot-records produced by the staggered-grid finite-difference, even though, the grid-dispersion condition was satisfied, particularly for the S-wave because of its lower velocity. Reducing the grid-spacing had improved the quality of the seismic-records and reduced the grid-dispersion and that scattering that appeared in the coarser grid-spacing. For the seismic-records produced by the pseudo-spectral finite-difference, there were some differences evident

among their shot-records and finer grid-spacing produced better quality. However, this improvement was not resulting from finer grid-spacing but from finer timesteps; which were imposed by the stability condition as consequences of using finer grid-spacing. I will provide more elaboration and discussion of my findings regarding the improvement that may be obtained from using finer timesteps in the section to follow.

The analysis, aforementioned, relied only on visual observations and did not provide detailed and reliable explanation to these differences. A more robust method that would allow to measure and quantify the differences was essential. I found that using the normalized crosscorrelation as a tool to link the shot-records in a way that could reveal their differences was very useful. The normalized crosscorrelation is defined as the ratio of the crosscorrelation between two fields over the autocorrelation of a reference field (Ikelle, 2005). At each point of the seismic-record, (x, t) , a portion of the data made of J times-samples and I traces around a point (x, t) , was extracted from one seismic record and crosscorrelated with its portion counterpart from another seismic record. The numbers corresponding to the maximum value of the crosscorrelation and its correspondent time-lag were allocated at the same location as the original point (x, t)

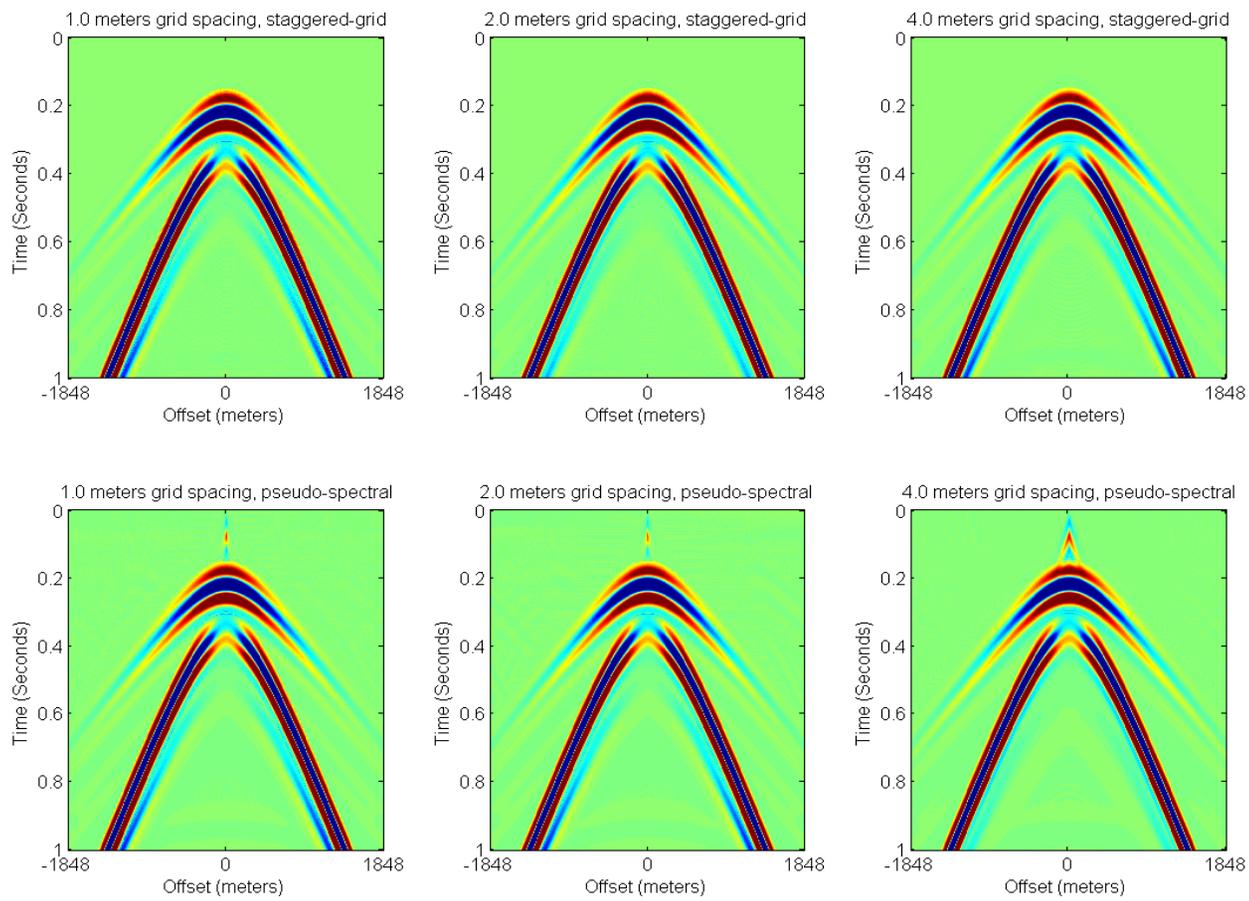


Figure 13. Staggered-grid and pseudo-spectral finite-difference of 4-meter, 2-meter and 1-meter grid spacings for a homogenous model. The physical quantity displayed here is the vertical component of the particle-velocity and the seismic source is a vertical force with 10 Hz central frequency.

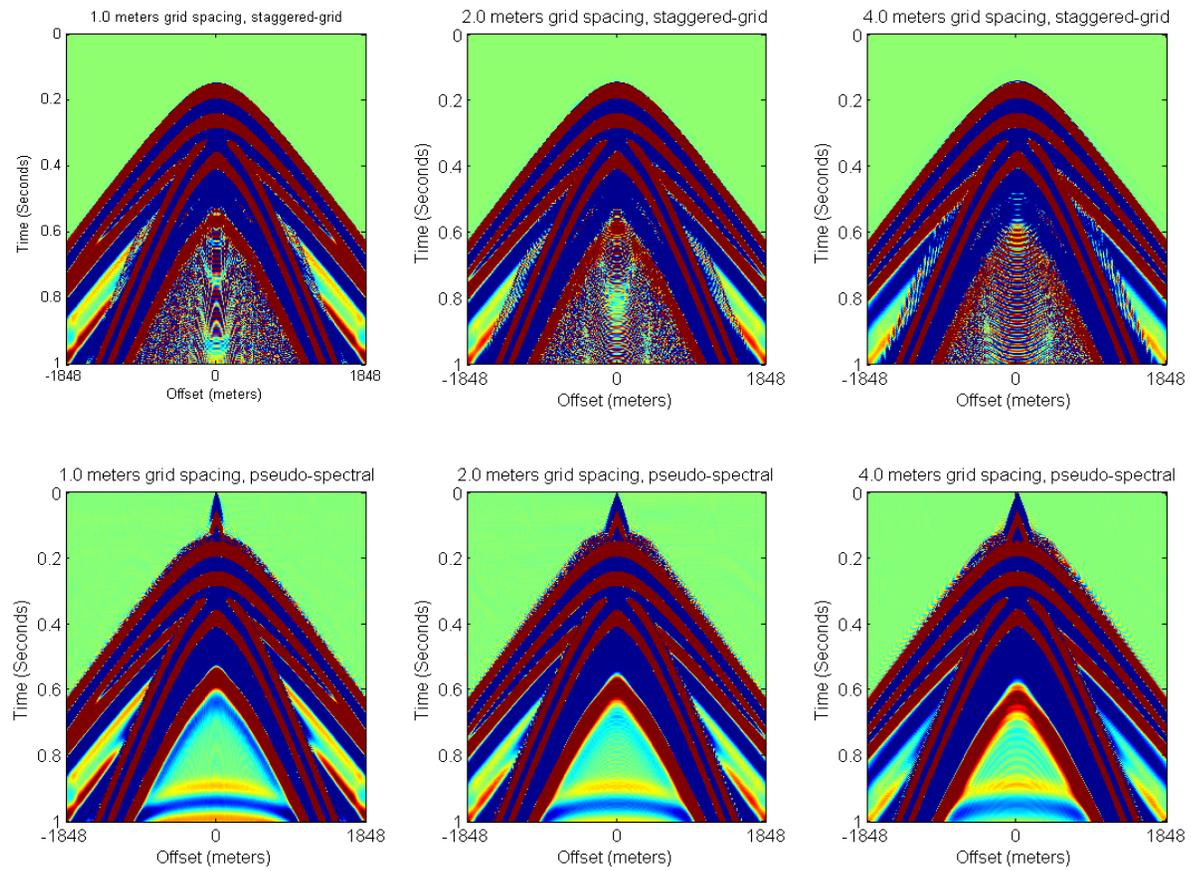


Figure 14. The same as Figure 13 with large scaling.

in an array. Then, the process kept on repeating for other portions of the data, after incrementing one data-point each time, until the whole records were covered. Finally, the maximum values of the crosscorrelation were normalized by the maximum values of autocorrelation and converted to decibels (dB) scale. For cases in which the crosscorrelation was too small or almost null, the process was voided, so that no “Nan” or “Infinity” introduced into the results. Figure 15 shows the maximum values of the crosscorrelation for the similar grid-spacing from both techniques. So, the one-meter grid-spacing from the pseudo-spectral was crosscorrelated with the one-meter grid-spacing from the staggered-grid and so on for the two-meter and four-meter grid-spacing. The size of the windows used to compute the crosscorrelations were 60 milliseconds by 10 traces. Observing Figure 13, showed that the seismic-records were reproduced, but in terms of their maximum crosscorrelation values rather than their

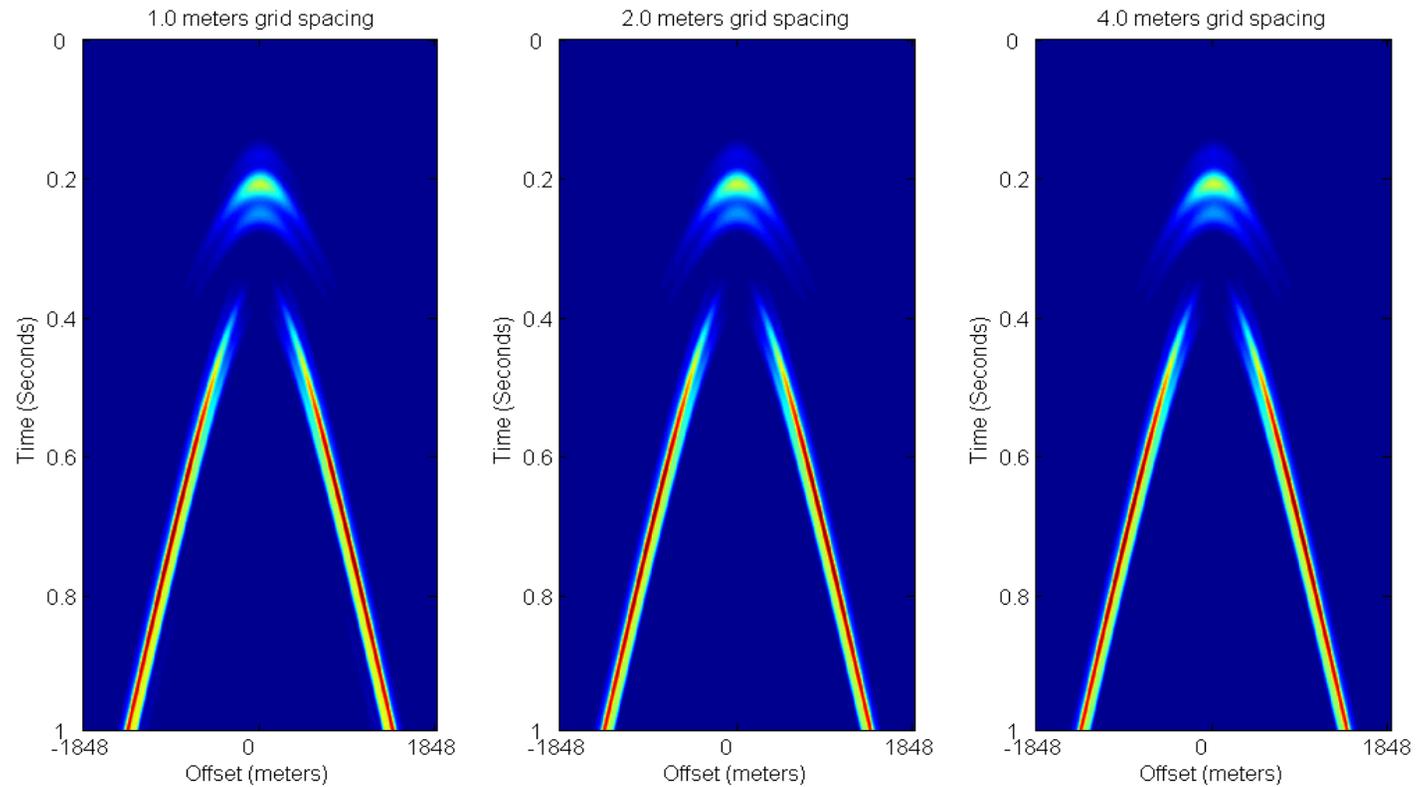


Figure 15. Crosscorrelation records for similar grid-spacing from both techniques. The one-meter grid-spacing from the pseudo-spectral was crosscorrelated with the one-meter from the staggered-grid and so on for the two-meter and four-meter grid-spacing. The size of the windows used to compute the crosscorrelations were 60 milliseconds by 10 traces. The seismic-records were reproduced, but in terms of their maximum crosscorrelation values rather than their absolute amplitude.

absolute amplitude values, thus they were connected in way that can help to compare them. However, these crosscorrelation records ought to be normalized in order to post the details and show their differences. Figure 16 showed the same crosscorrelation records as in Figure 15, but after they had been normalized, by the maximum values of the pseudo-spectral seismic-records autocorrelations, and converted to decibels (dB), thus revealing their correlation energy, along with their crosscorrelation time-lags. Now, the normalized crosscorrelation records provided a mean to observe and analysis their differences. I decided to limit the scale shown to be in the range of $-10dB$ to $5 dB$, so that better observations could be made. *Zero-dB*, at any point of the normalized crosscorrelation records, suggested no difference between the two seismic-records that were crosscorrelated. *Negative-dB* suggested lower amplitude of the staggered-grid seismic-records in relative to the pseudo-spectral seismic-records and *positive-dB* suggested higher amplitude of the staggered-grid seismic-records in relative the pseudo-spectral seismic-records. With that being said, observing the three records in Figure 16, showed that, as the grid-spacing became finer; from four-meter to two-meter to one-meter grid-spacing; the difference between the pseudo-spectral seismic-records and the staggered-grid seismic-records was minimized. This was evident from the decibel values going toward the *zero-dB* with finer grid-spacing. In other words, as the grid-spacing decreased, the difference between the two finite-difference techniques was reduced substantially. Also, the crosscorrelation time-lags records showed that time-lags decreased with decreasing grid-spacing. Another difference was the grid-dispersion,

which was still evident, particularly at the four-meter grid-spacing and disappeared as the grid-spacing became finer.

In order to understand what attributed to these differences, I looked at the *negative-dB* values aside from the *positive-dB* values, Figure 17. The top ones are the *negative-dB* records and the bottom ones are *positive-dB* records. *Negative-dB* values were indications of stronger anomalies in the pseudo-spectral seismic-records than the staggered-grid seismic records or totally existed in the pseudo-spectral seismic-records and not in the staggered-grid seismic records while the *positive-dB* values were indications of the contraries. Hence, it was easily observed that the *negative-dB* records contained most of the differences, particularly for the P- and S-wave reflections, while the *positive-dB* records showed almost no difference for the two seismic events. Furthermore, the *positive-dB* records showed the grid-dispersion that the four-meter grid-spacing seismic record, generated by the staggered-grid finite-difference, suffered from. These observations confirmed that the Pseudo-spectral finite-difference could achieve better seismic simulations even with larger grid-spacing while the staggered-grid finite-difference would need to have much smaller grid-spacing to start producing comparable simulations. These were evident from the *negative-dB* records where the differences were diminishing as the grid-spacing decreased. Due to the limitation in computing resources that deterred me from repeating these simulations with additional reduction in the grid-spacing, because of the exponential increase in the size of the

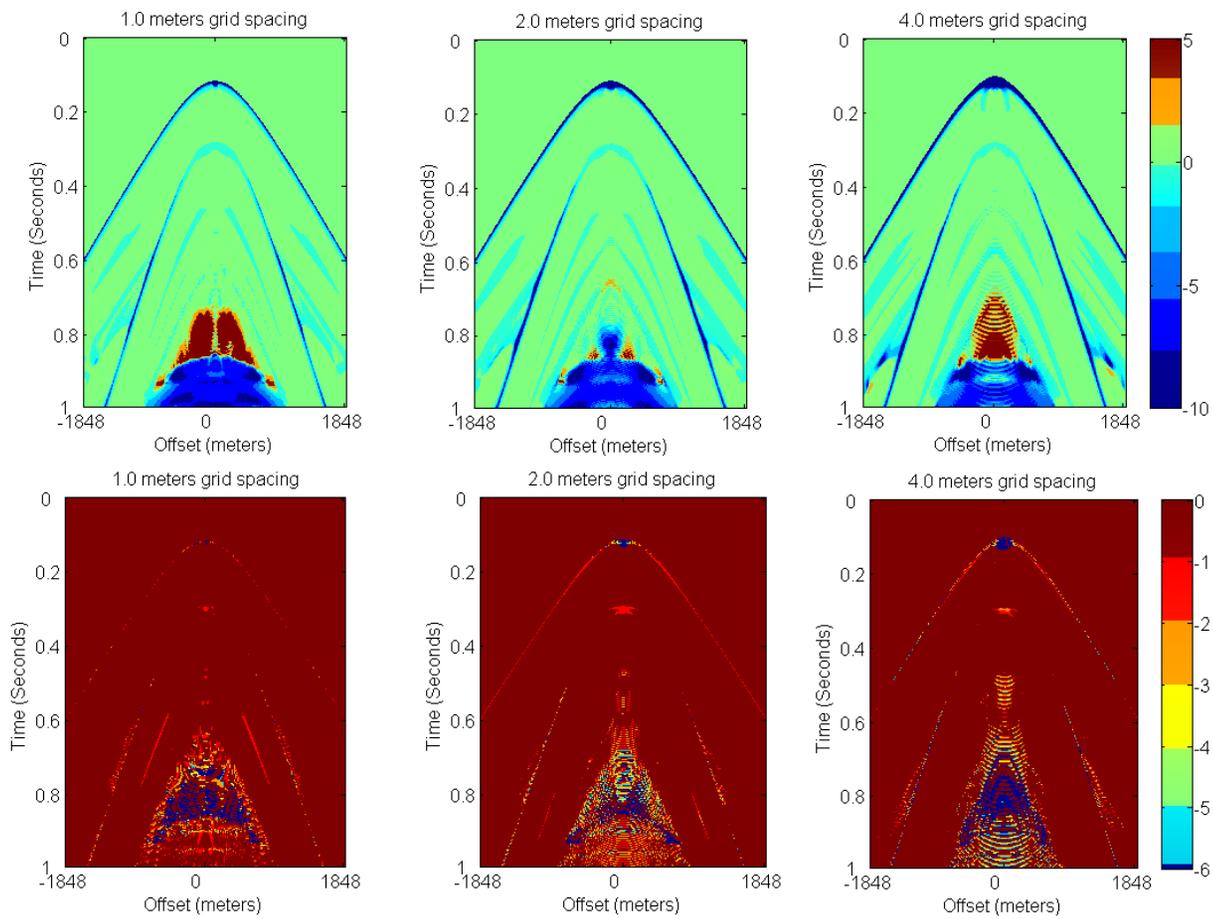


Figure 16. The crosscorrelation records in Figure 15, after normalization. The correlation energy (top) and the time-lags (bottom).

model and the associated decrease in the timesteps, which would make the simulations very costly prohibitive. I believe there would be a threshold where further decrease in the grid-spacing would not produce anymore differences between the two finite-difference numerical results, for my model here; with the assumed elastic parameters. Different models and different elastic parameters might produce different thresholds.

As I have highlighted during the review on the staggered-grid finite-difference, the spatial derivatives are estimated by a fourth-order difference operator that uses the first four terms of the Taylor series approximation and truncate higher-orders. As a result, the spatial derivatives are only approximated and never computed accurately, as in the case of the pseudo-spectral finite-difference; where the spatial derivatives are computed accurately by the method of the Fourier-transforms, thus having an infinite order of accuracy. Improving the accuracy of approximating the spatial derivatives, in the staggered-grid finite-difference, would either requires adding more Taylor terms to the approximation, thus allowing the use of larger grid-spacing or decreasing the grid-spacing substantially. As more terms are added, the Taylor series approximation is improved along the function. This would suggest that using a six-order difference operator, would definitely improve the accuracy of approximating the spatial derivatives with larger grid-spacing, which would help to reduce the size of the model and the timestep; resulting in cost effective simulation of the seismic data. However, the staggered-grid finite-difference would become much more computationally intensive with very lengthy difference operators that might become complicated in its formulations. Also, the free-surface boundary condition is much more difficult to

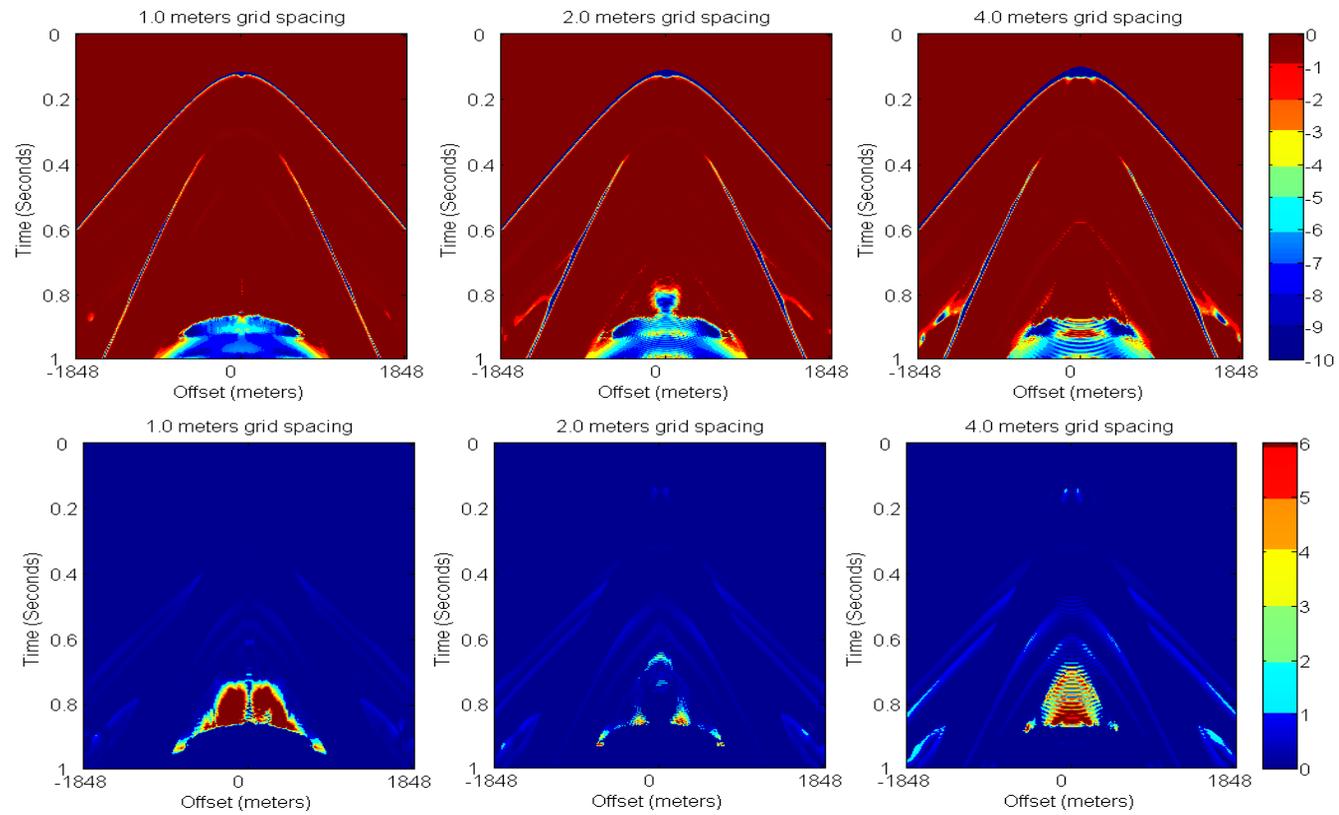


Figure 17. The normalized crosscorrelation in Figure 16, with the *negative-dB* values (top) were analyzed aside from the *positive-dB*-value (bottom).

solve for long spatial difference operators. Computing the spatial derivatives for any physical quantity, such as the particle velocity, with six-order difference operator, at any point in the model, would require using the stresses, from the previous timestep, for six neighboring points in each direction. Alternatively, decreasing the grid-spacing substantially would improve the accuracy of approximating the spatial derivatives with fewer Taylor terms, but the simulations become costly prohibitive due to the significant increase in the size of the models and consequently the sharp decrease in the timestep.

Before I proceed on discussing the implication of reducing the grid-spacing, in the staggered-grid finite-difference, on the cost of modeling array data, I will briefly highlight the improvement that maybe achieved by simulating seismic data with finer timesteps. Here, I discuss an example of reproducing the same seismic shot-record, generated by the pseudo-spectral finite-difference for the 4.0 meter grid-spacing for the same model as above, but with finer timesteps. Six different runs were reproduced with various timesteps ranging from 0.5-0.00625 milliseconds, which required number of timesteps ranging from 2000-180,000 to generate the one-second long seismic-records, Table 3. Observing the shot-records showed substantial improvement in the P- and S-wave events including the artifact reflections from the top air-filled and bottom layers of the model with finer timesteps. These artifact reflections were due to the fact that the absorbing sponge that was applied along the edges of the model in order to eliminate reflections from the edges was not a perfect attenuator. Also, the source outburst, which resulted from the source interacting with the surrounding medium, which had been observed after posting the amplitude in the pseudo-spectral shot-records, was reduced

Table 3: Six runs of the same model generated by the pseudo-spectral finite-difference with the 4.0m grid-spacing.

Shot-record	Timestep (millisecond)	Number of timesteps	Length of shot-records
1	0.5	2,000	One-second long
2	0.1	10,000	One-second long
3	0.05	20,000	One-second long
4	0.025	40,000	One-second long
5	0.0125	80,000	One-second long
6	0.00625	160,000	One-second long

substantially with finer timesteps, Figure 18. These observations reconfirm that for the pseudo-spectral finite-difference, the grid-spacing has no impact on the quality of the simulated seismic data. The differences were only due to the change in the timesteps. Therefore, the differences that were observed in Figure 14, for the pseudo-spectral shot-records, were totally due to the finer timesteps, which were imposed as a consequence of reducing the grid-spacing.

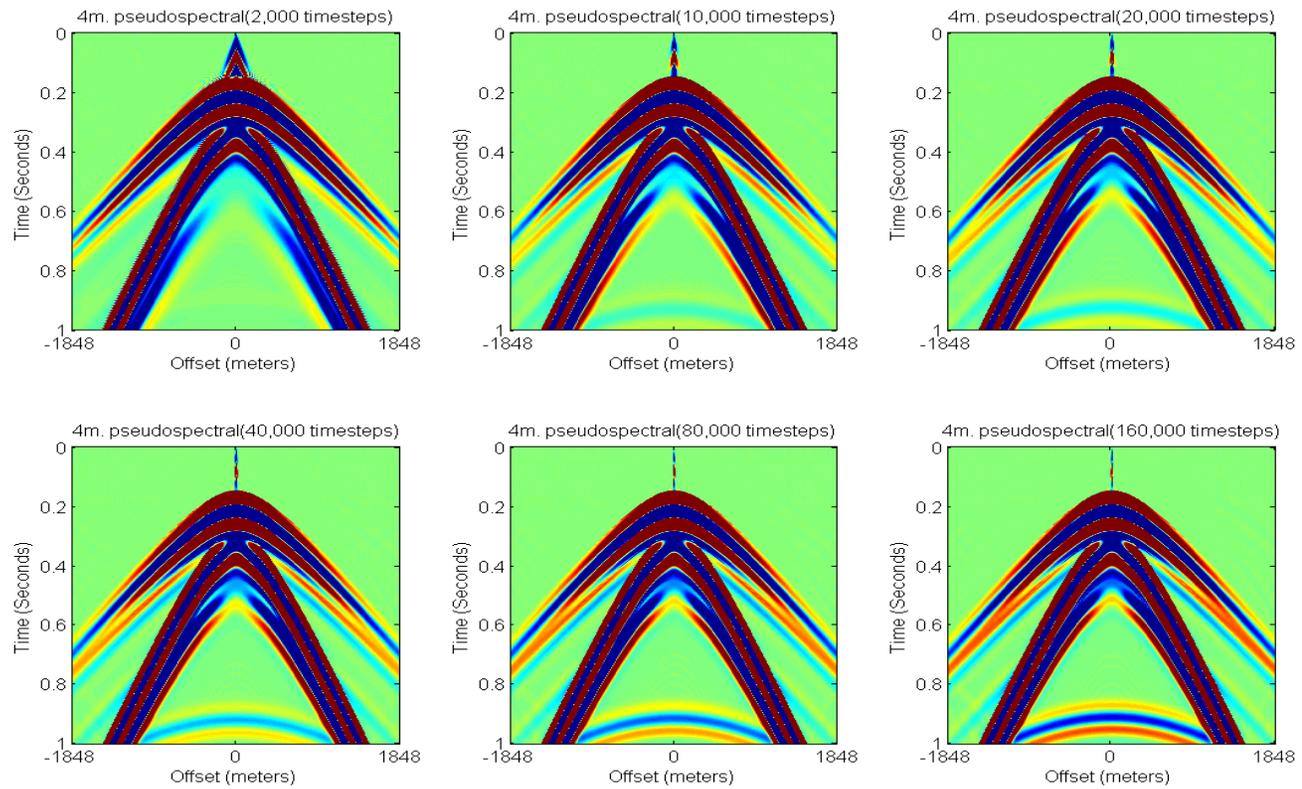


Figure 18. Six different runs generated by the pseudo-spectral finite-difference for the 4m grid-spacing of the same model, but with different timesteps ranging from 0.5-0.00625ms.

Here, I compare the staggered-grid and the pseudo-spectral finite-difference techniques with respect to their efficiency to model three different examples. Basically, I look at the implication of the model elastic parameters over the finite-difference parameters of both techniques. Finite-difference Parameters, like grid-spacing and timestep, are selected based on the elastic parameters of the models like the minimum and maximum seismic velocities and the resolution bandwidth needed to model the seismic response. Finite-difference parameter, like the model size and the total number of grid-points are byproducts of the grid-spacing and the timestep. Therefore, the model size is selected based on the number of grid-points needed to cover the physical size of the model and the total number of timesteps is selected based on the two-way propagation time needed to simulate the seismic response of the model. These are two key parameters that control the efficiency of executing any finite-difference techniques. The smaller the size of the finite-difference model and the fewer the required timesteps are, the faster the execution and the more efficient the finite-difference is.

Table 4 describes the elastic parameters for three examples and the associated finite-difference parameters for the staggered-grid and pseudo-spectral and finite-difference techniques. While the finite-difference parameters for the pseudo-spectral finite-difference, in Table 4, were computed mostly based on the stability and grid-dispersion conditions, described in Table 1, the grid-spacing for the staggered-grid finite-difference was computed based on 8 grid-points per wavelength because the

Table 4. The implications of three different elastic model parameters on the parameters of both the finite-difference techniques. In this Table, the finite-difference parameters are Δt , which stands for the timestep of the finite-difference, and Δx , which stands for the grid-spacing for the finite-difference. The medium parameters are V_{max} , which stands for the maximum velocity of the medium, V_{min} , which stands for the minimum velocity of the medium and f_{max} , which stands for the maximum frequency of the source.

Examples		Staggered-grid finite-difference				Pseudo-spectral finite-difference			
Example I									
Geology model (km)	(2.5X2.5)	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps
Record length	2 seconds								
Vmin	1200m/s								
Vmax	5000m/s	(1000X1000)	2.5	0.25	8,000	(250X250)	10	0.8	2,500
Fmax	60 Hz								
Example II									
Geology model (km)	(2.5X2.5)	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps
Record length	2 seconds								
Vmin	300m/s								
Vmax	5000m/s	(5,000X5,000)	0.5	0.05	40,000	(1000X1000)	2.5	0.2	10,000
Fmax	60 Hz								
Example III									
Geology model (km)	(2.5X2.5)	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps	Model size	Grid spacing (m)	Timestep (ms)	Number of timesteps
Record length	2 seconds								
Vmin	300m/s								
Vmax	5000m/s	(10,000X10,000)	0.25	0.025	80,000	(2000X2000)	1.25	0.1	20,000
Fmax	100 Hz								

objective is to compare their efficiency, when their accuracy are similar. Also, in the three examples, the maximum velocity, the physical size of the model and total propagation time were kept constant because the comparison is about the efficiency of the both finite-difference techniques to model array data; where the minimum velocity can be very small. So, in the three examples, the physical size of the model is (2500X2500), maximum seismic velocity 5000 m/s and total propagation time is *2 seconds*.

In the first example, the minimum velocity is 1200 m/s and the maximum source frequency is *60 Hertz*. Therefore, the grid-spacing and timestep in the pseudo-spectral finite-difference are four times larger than the grid-spacing and the timestep of the staggered-grid finite-difference. Thus, only 2,000 timesteps are required for a finite-difference model that is four times smaller, in the pseudo-spectral finite-difference, opposed to 8,000 timesteps for a finite-difference model that is four time bigger, for the staggered-grid finite-difference. Therefore, it is evident that the efficiency of the pseudo-spectral finite-difference in relative to the staggered-grid finite-difference is about 16 times, for the assumed elastic parameters. However, despite the great difference in the efficiency between the two finite-difference techniques, 8,000 timesteps for a model size of (1000X1000) is considered not very heavy in terms of the time to execute this model on average computing tool.

However, in second example, the minimum velocity was reduced to 300 m/s , while keeping all other elastic parameters constant in relative to the first example. Seismic events with these low velocities are typical, when modeling array data. It is

evident the low velocity required very small grid-spacing and timestep, in the staggered-grid finite-difference, which imposes very large finite-difference model and extremely big number of timesteps. This lack of efficiency, for the staggered-grid finite-difference versus the pseudo-spectral finite-difference, is shown even more in the third example, where broader bandwidth is used. Therefore, the staggered-grid finite-difference is extremely cost prohibitive when modeling seismic data at the array level. On the other side, the pseudo-spectral finite-difference, which requires much smaller finite-difference model and very reasonable number of timesteps, appear to be cost effective alternative to the staggered-grid finite-difference, modeling seismic data at the array level.

Another benefit that could be gained from the use of the pseudo-spectral finite-difference in modeling seismic data is the frequency independence of the grid-spacing. From my analysis, I found that the pseudo-spectral finite-difference is not only a reliable modeling tool in cases where smaller grid-spacing is no longer affordable, but also in cases where the vertical resolution is not broad enough to resolve thin interfaces. Unlike the staggered-grid finite-difference, the pseudo-spectral finite-difference has an infinite order of accuracy in computing the spatial derivatives, thus requires only two grid-points per wavelength opposed to eight or sixteen grid-points per wavelength for the same accuracy for the staggered-grid finite-difference. Therefore, broader frequency bandwidth of the seismic source could be used, in the pseudo-spectral finite-difference; to resolve thin interfaces where otherwise would not possible by the staggered-grid finite-difference. To illustrate this benefit, I discuss briefly an example where a thin layer of 60 meters thick was inserted in a homogenous medium, as shown in Figure 19.

With grid-spacing of 4 meters and a model minimum velocity of 1000 meter/second, the staggered-grid finite-difference requires source maximum frequency to be about 60 Hz, as imposed by the grid-dispersion condition; in order to avoid severe grid-dispersion affecting the modeled data. With these assumed parameters, the top and bottom interfaces of the thin layer would not be resolved, thus appeared as one interface, as shown in Figure 20. Broadening the source bandwidth with the assumed parameters

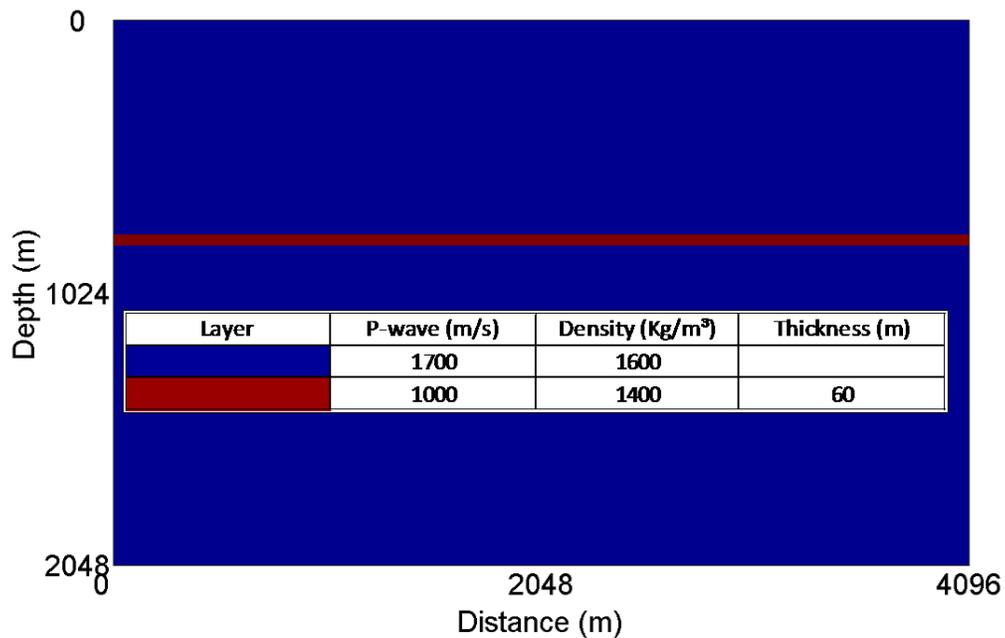


Figure 19. A thin layer of 60m thick imbedded in a homogenous medium.

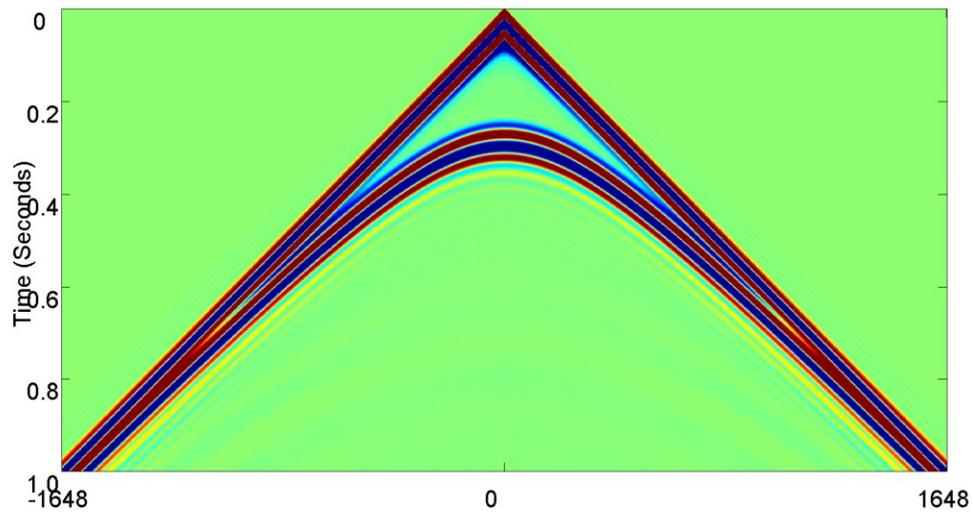


Figure 20. A shot-record for the model in Figure 19, generated by the staggered-grid finite-difference. With 4m grid spacing and minimum velocity of 1000m/s, the grid-dispersion condition imposed maximum source frequency of about 60 Hz, which was not broad enough to resolve the top and bottom interfaces.

may resolve the top and bottom interfaces of the thin layer, but generate severe grid-dispersion, Figure 21. Alternatively, the pseudo-spectral finite-difference can be used

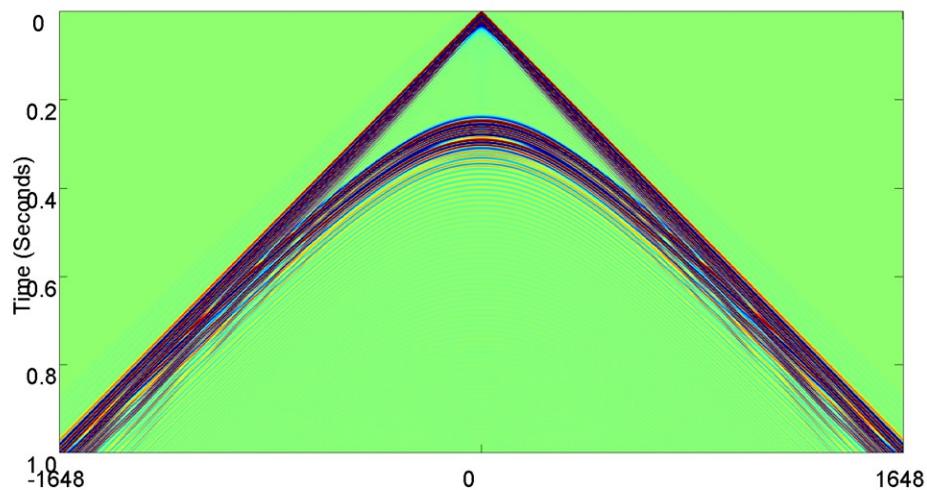


Figure 21. The same shot-record as in Figure 20, except the source bandwidth was broadened, which violated the grid-dispersion condition. Severe grid-dispersion had resulted that effected the simulated data.

with much broader seismic-source bandwidth to resolve the top and bottom interfaces of the thin layer with the assumed model parameters, Figure 22. The pseudo-spectral finite-difference has no grid-dispersion condition that needs to be satisfied except the two grid-points per wavelength, which is a requirement of the Nyquist spatial sampling, rather than by the pseudo-spectral finite-difference. Thus, allowing the use of a broader seismic source bandwidth with no extra cost. Therefore, the pseudo-spectral finite-difference is much reliable and cost effective modeling tool in cases where the higher seismic source bandwidth is desired to model thin interfaces.

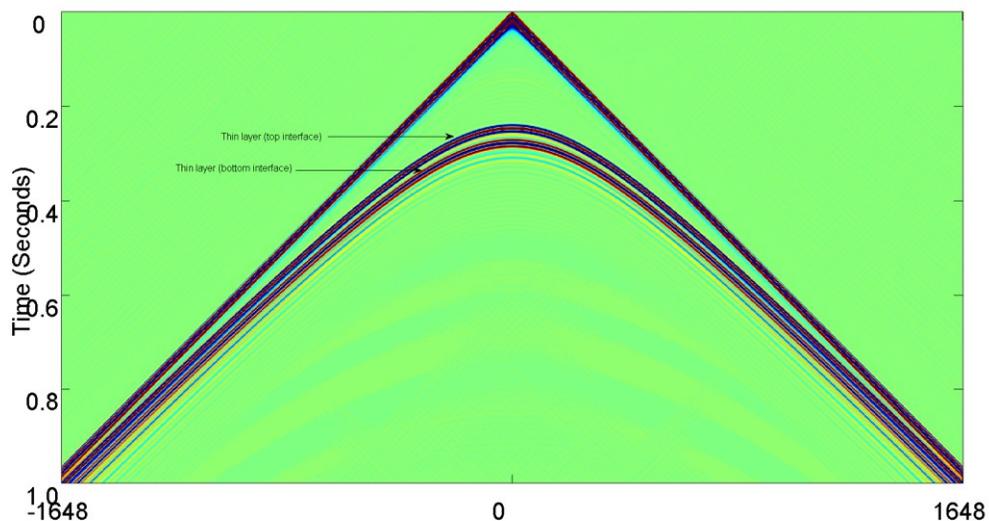


Figure 22. A seismic shot-record, generated by the pseudo-spectral finite-difference, for the model in Figure 19. Unlike the staggered-grid finite-difference, the pseudo-spectral finite-difference has no restriction on the seismic-source maximum frequency; except the two grid-points per wavelength, therefore boarder bandwidth was used to resolve the top and bottom interfaces of the model.

NUMERICAL EXAMPLE (GROUNDROLL)

In this section, I demonstrate the capability of the pseudo-spectral finite-difference to model array data for land example. In this model, I use a thin wiggly air-filled layer as an artificial free-surface along with wiggle interfaces to simulate a real land free-surface. I also use thin layered medium to excite dispersive Rayleigh-waves (ground roll) and backscattering, which are common seismic events in land seismic data. However, before I proceed on describing my model and the parameters of the elastic medium and finite-difference, I briefly review and show an air-filled layer can be used as an artificial free-surface in the finite-difference. Then, I will also discuss briefly how dispersive ground roll are formed; because they are part of the elastic wavefield recorded in any land seismic experiment, (Ikelle and Amundsen, 2005).

Air-filled layer

The free-surface boundary condition can also be satisfied artificially by literally adding an air-filled layer as the first layer of the geological model. It is an easy way to accommodate for more realistic models because the free-surface can be of any topography. By applying an absorbing boundary at the top of the air-filled layer, which can eliminate reflections from the top of the model and by the acoustic impedance contrast between the air-filled layer and the first interface of the medium, an artificial free-surface can be created that can produce comparable modeling accuracy to that of satisfying the free-surface boundary condition.

To examine the ability of the air-filled layer to act as an artificial free-surface, two models were examined. The first model was made of a 500 meter thick homogenous layer overlain by a 200 meter thick air-filled layer, Figures 23. The second model was only made of the homogenous layer. An explosive source and pressure sensors were used to record the pressure wavefield for both models. In both models, the explosive source was placed in the center laterally. However, in the air-filled layer model, the sensors were placed at 300 meters deep from the top of the air-filled layer and 100 meters deep inside the homogenous layer. For the homogenous model, the sensors were placed at 100 meters deep inside the model. Snapshots of both models were recorded at the same instant, 0.25 second, which show the P-wave direct arrival and the P-wave reflection from the air-filled layer interface, Figure 24 and from the free-surface boundary, Figure 25. Both snapshots show comparable results for modeling

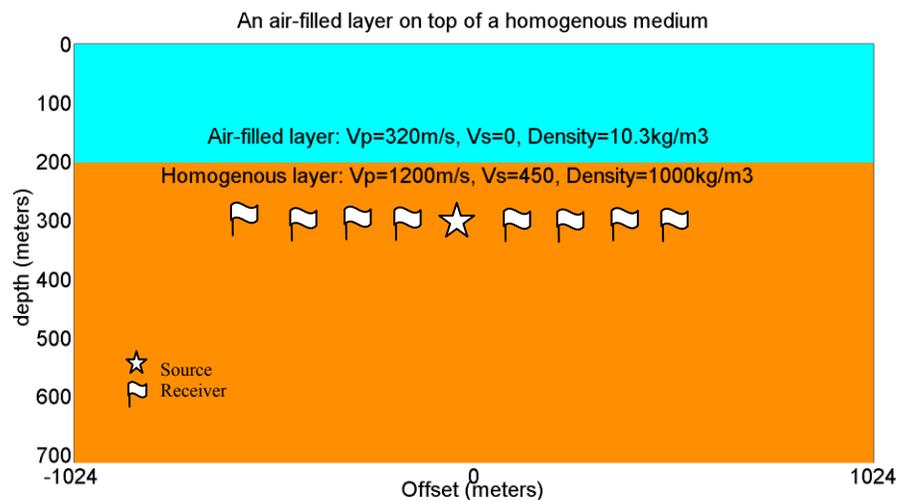


Figure 23. Homogenous medium of 500m thick, overlain by 200m thick air-filled layer. An explosive source was placed in the center of the model at a depth of 300m. Pressure sensors were placed at 321m depth inside the homogenous medium.

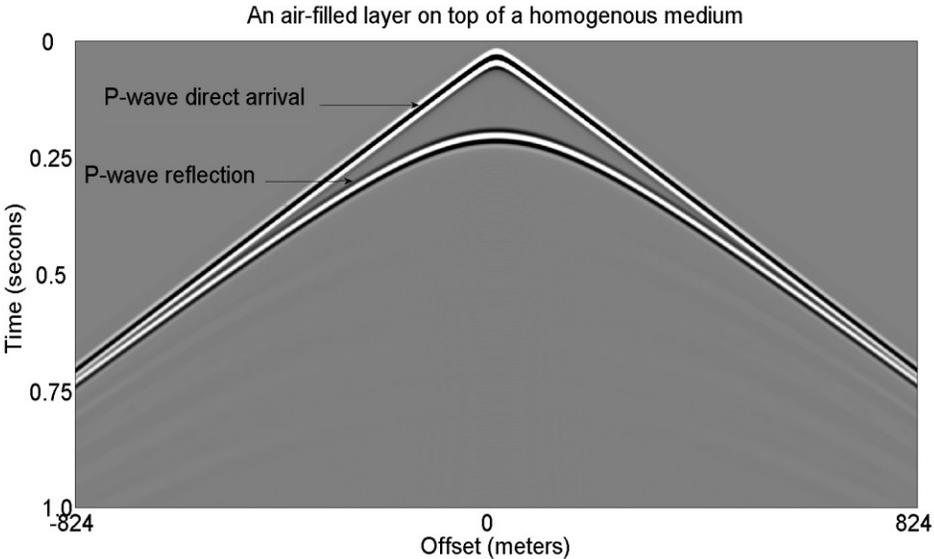


Figure 26. Seismic shot-record of one-second long, showing a P-wave direct arrival and a P-wave reflection from the air-filled layer, which acted as an artificial free-surface.

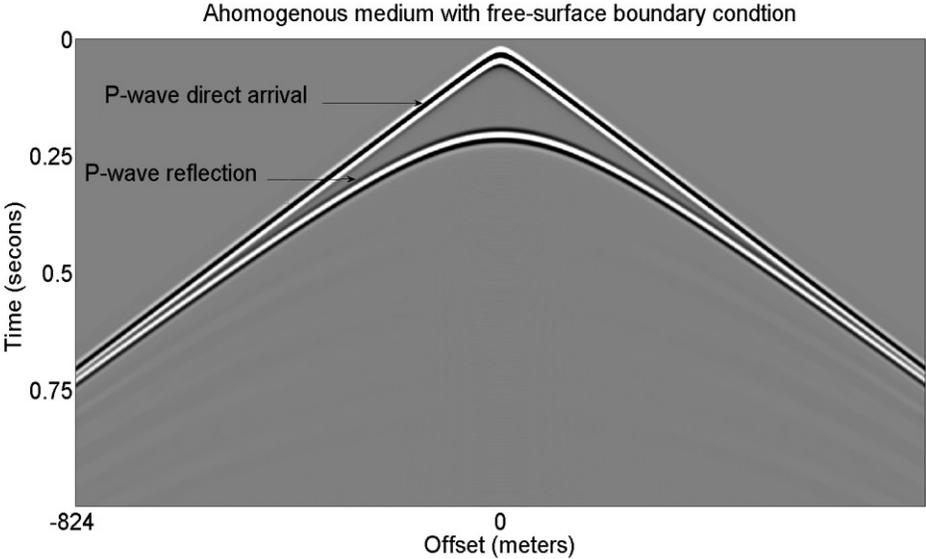


Figure 27. Seismic shot-record of one-second long, showing a P-wave direct arrival and a P-wave reflection from the free-surface.

The numerical results of the previous two models, snapshots and shot-records, have shown that the air-filled layer can be used as an artificial free-surface. It produces comparable modeling accuracy to that of satisfying the boundary condition for a planar free-surface. However, by using the air-filled layer as an artificial free-surface, the constraint on having a planar free-surface is no longer required. This would allow modeling more realistic complex models because the artificial free-surface can be, now, of any topography. For example, Figure 28 shows a snapshot, recorded at 0.4 second, for homogenous medium overlain by an inclined air-filled layer. It is evident that the inclined air-filled layer acted as an artificial free-surface that modeled reflected incident P-wave accurately. Figure 29 is another example of a snapshot recorded at 0.32 second, where the air-filled layer is wiggly. Here, the reflected P-wave is dissipated and scattered as a result of the wiggly air-filled layer. All these models confirm that the air-filled layer is a reliable choice as a free-surface, whenever modeling requires the free-surface to be non-planar. The air-filled layer or the artificial free-surface is used to model more examples in the subsequent sections.

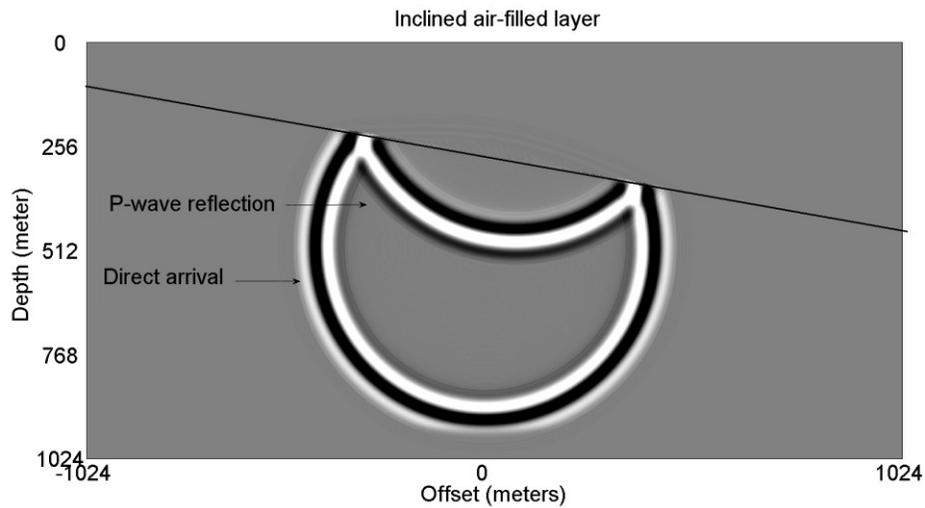


Figure 28. A snapshot recorded at 0.4s for homogenous medium overlain by an inclined air-filled layer that acted as an artificial free-surface.

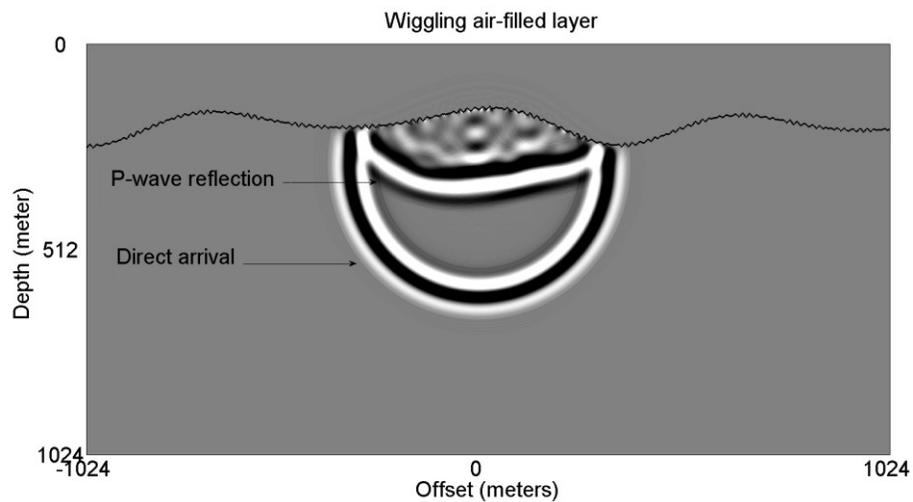


Figure 29. A snapshot recorded at 0.32s for homogenous medium overlain by a wiggly air-filled layer that acted as an artificial free-surface. The reflected P-wave is dissipated and scattered as a result of the wiggly free-surface.

While medium parameters of the air-filled layer such as the density and P-wave velocity are much lower than the rest of the medium, which would normally impose very fine grid-spacing and the timesteps, which are required by the stability and the grid-

dispersion conditions of the finite-difference, errors introduced as results of violating these conditions would not affect the modeled data. These errors would only occur in and near the air-filled layer interface. Thus, as long as the sensors are not placed in or near the air-filled layer, no errors have been observed in the modeled data.

Dispersive groundroll

Ground roll are vertically polarized Rayleigh waves; named after (Rayleigh, 1885). Rayleigh showed that there exists a solution that satisfies the elastic equation of motion for a wave traveling along a free-surface, with retrograde to elliptical particle motion, and decays exponentially with depth. The decay, in amplitude, is inversely proportional to the square-root of the distance from the source. Thus, its amplitude decays more rapidly than a P- or S-wave, which their amplitudes decay inversely proportional to the distance from the source. Its generation comes from diffracting an incoming P- or S-wavefront at the free-surface. A condition necessary for generating high amplitude ground roll is that the source must be on or near the free-surface, so that the wavefront is strongly curved, (Ikelle and Amundsen, 2005). Its velocity is slightly less than the S-wave velocity and only depends on the ratio of the P- and S-wave velocities, or equivalently on the Poisson's ratio. (Knopoff, 1952) derived the ratio of the Rayleigh- and S-wave velocities as a function of the Poisson's ratio. He showed that for typical values of Poisson's ratio, in the range of 0.2 to 0.4, the Rayleigh-wave velocity, as a function of the S-wave velocity, varies from $0.9V_s$ to $0.95V_s$,

To analyze how ground roll are formed, I created a simple model made of homogenous medium, below a free-surface. A vertical source of 20 Hertz central frequency was placed in center laterally and at 6 meters below the free-surface. The P- and S-waves velocities of the medium were 800 m/s and 500 m/s respectively. These velocity values made the Poisson's ratio to be about 0.2, which made the velocity of the ground roll, to be about 0.91 of the S-wave velocity. A snapshot, recorded at one second, of the vertical component of the particle-velocity shows that a homogenous layer below a free-surface is enough to excite ground roll. It also shows the rapid decay, in ground roll amplitude, with depth. Ground roll are lacking behind S-wave is also evident, and it may require large offset to observe the separation between the two seismic events, Figure 30. In this model, the ground roll is not dispersive, however in land; the near-surface, which is immediately below the free-surface, is unconsolidated to semi-consolidated and very heterogeneous in its elastic parameters and thickness, thus it is referred to in petroleum seismology, by the weathered zone or the low-velocity zone. It can cause ground roll to be dispersive. The dispersion happens as a result of various wavelength components, or equivalently various frequency components, travel with different velocities. Consequently, it may cover an envelope-cone instead of a straight line on a seismic shot-record. To show how dispersive ground roll would look like on a snapshot and on a seismic shot-record, we used the same model as in Figure 28; with the exception of adding a second thin layer, right below the free-surface, with a thickness of 10 meters. The P- and S-wave velocities, for the thin layer, were 1050 m/s and 700 m/s respectively. Figures 31 and 32 show a snapshot, recorded at 0.8 second, and a

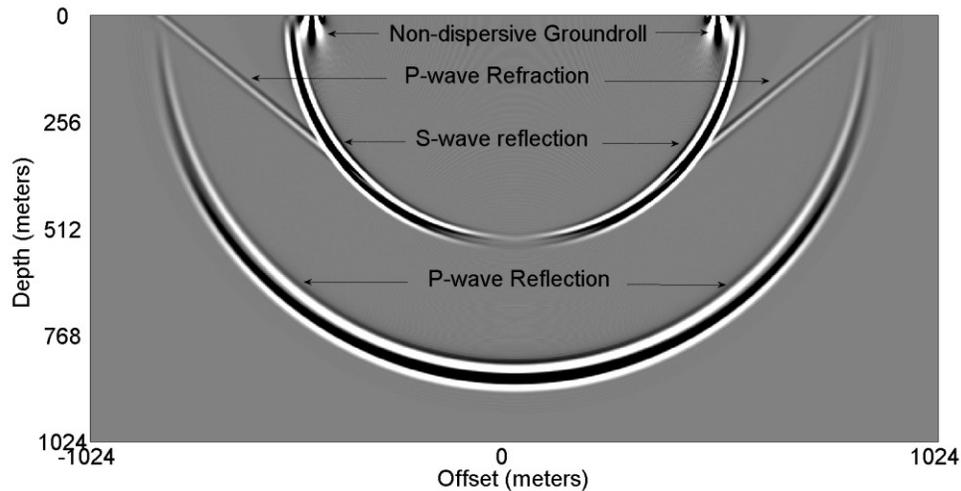


Figure 30. A snapshot, recorded at one second, showing the vertical component of the particle-velocity for a homogenous layer below a free-surface. A vertical source of 20 Hertz central frequency was placed in center laterally and at 6 meters below the free-surface. The P- and S-waves velocities of the medium are 800 m/s and 500 m/s respectively. Non-dispersive ground roll is excited, which decays, in amplitude, with depth.

shot-record, of one second long, for this model, for the vertical component of the particle velocity respectively. Ground roll, now, is dispersive, as a result of the one-dimensional heterogeneity, below the free-surface. As ground roll travels along the free-surface with its retrograde motion, it keeps striking and bouncing from the thin layer, which has different elastic parameters. These cause different wavelengths to travel with different velocities, which generate several ground-roll modes that travel with different velocity. The same phenomena are observed for a partial part of the S-wave, which was trapped in the thin layer. Because of the radiation pattern for the vertical component of the particle velocity, which decreases with wider angles, the trapped part of the P-wave would not show any dispersion.

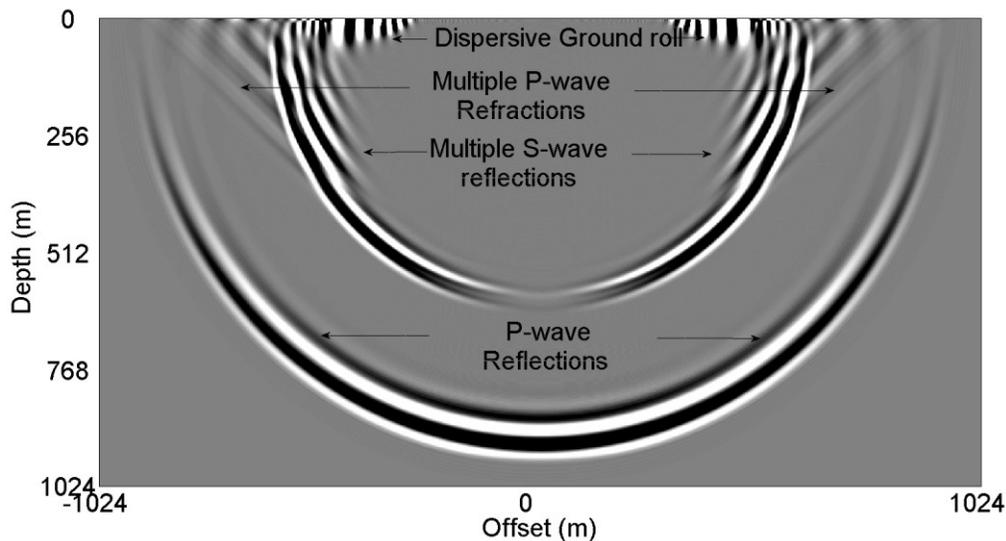


Figure 31. A snapshot, recorded at 0.8 second, showing the vertical component of the particle velocity, for the same model as in Figure 30, with the exception of adding a 10m thick layer below the free-surface. The P- and S-wave velocities, for the thin layer, were 1050 m/s and 700 m/s respectively. Due to the different elastic parameters of the thin layer, ground roll become dispersive; where various wavelength components travel with different velocity.

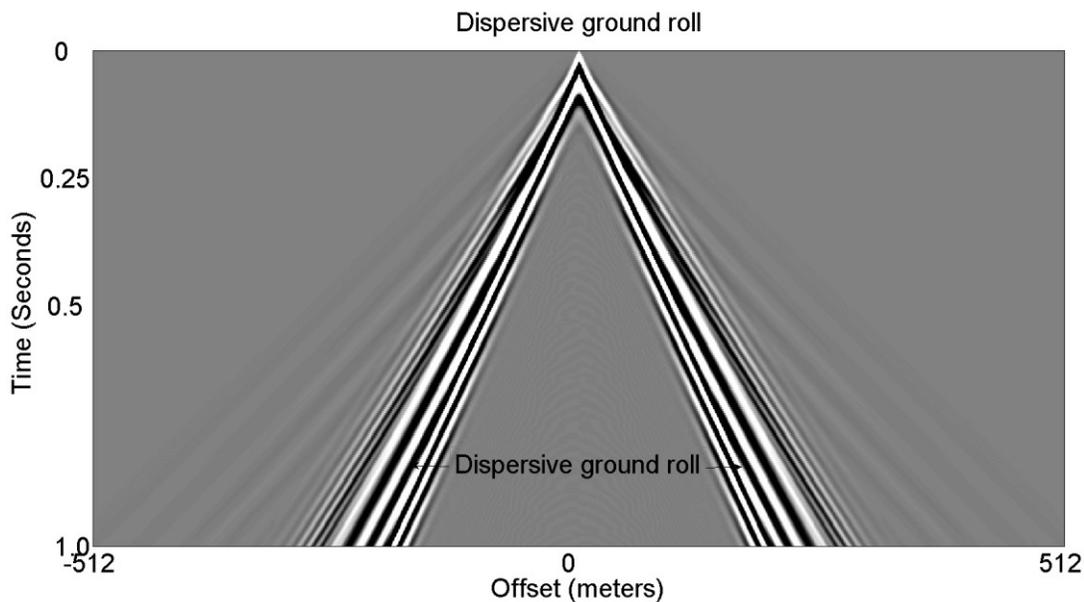


Figure 32. A shot-record, of one-second long, showing the vertical component of the particle velocity, for the same model as in Figure 30, with the exception of adding a 10m thick layer below the free-surface. The P- and S-wave velocities, for the thin layer, were 1050 m/s and 700 m/s respectively. Due to the different elastic parameters of the thin layer, ground roll become dispersive; where various wavelength components travel with different velocity.

After providing an overview of how ground roll are generated and how they can get dispersed, we can now proceed with our land data model example. In order to create a realistic land model, we used the air-filled layer as an artificial free-surface, so that we have a choice on how wiggly our free-surface can be. Generating high amplitude ground roll would also require placing the source immediately right below the free-surface, so that the curved wavefront can diffract off the free-surface and generates ground roll. We use a randomly heterogeneous thin layer right below the air-filled layer to excite dispersive ground roll. Because of the rapid decay in ground roll amplitude with depth, the sensors need to be placed right below the free-surface. The elastic parameters of the heterogeneous thin layer are balanced so that the Poisson's ratio that can show clear distinction between the S-wave and ground roll. Finally, we also use wiggly interfaces to generate diffraction and backscattering events.

This land model has a spatial cell dimension of (4096X2048). The air-filled layer occupied the first 400 cells from the top of the model vertically. The interface of the air-filled layer was created by summing a group of sine waves of various frequencies and amplitudes, in order to make the interface randomly wiggly, so that it simulates an actual land surface. The elastic parameters of the air-filled layer were 320m/s for the P-wave and 10.3 Kg/m³ for the density. The S-wave was null due to the fact that air is an acoustic medium that does not support S-wave propagation. A thin layer of 10 meters, in thickness, was placed right below the air-filled layer, which consisted of two different elastic mediums, distributed randomly across this layer. The objective was to make this layer heterogeneous, to simulate a weathered layer below a free-surface. Four more

layers, of wiggly interfaces, of various thickness and elastic parameters were placed below the thin layer respectively, Figure 33. A zoom-in of this model that shows the wiggly air-filled layer and the random distribution of the elastic parameters across the thin layer are shown in Figure 34. The elastic parameters for this model have been overlain on the figures for simplicity.

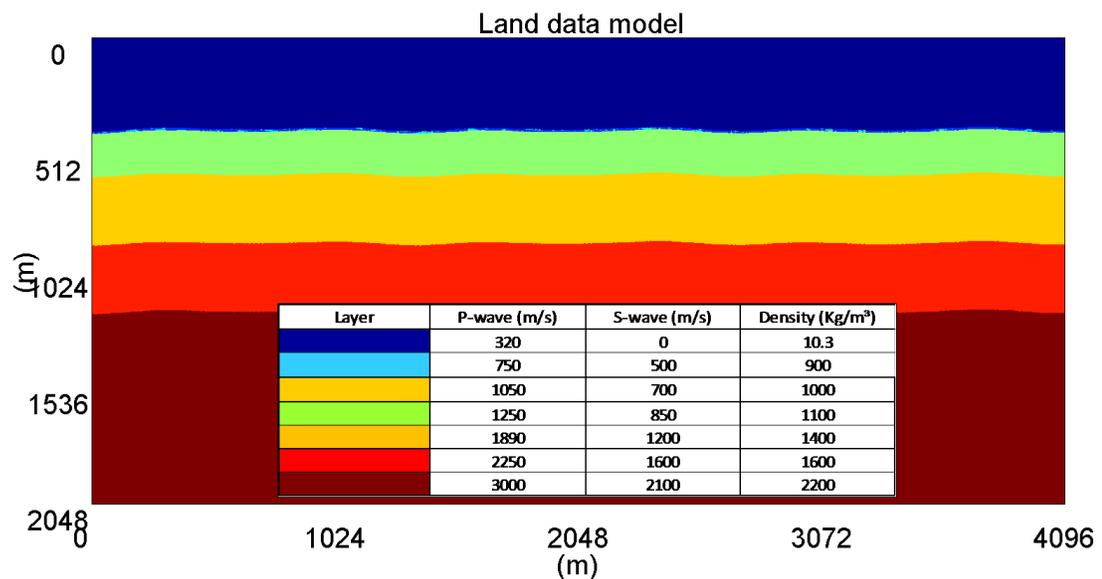


Figure 33. A land model made of (4096X2048) spatial cells that consisted of 7 layers. An air-filled layer was used to act as an artificial free-surface that occupied the first 400 cells from the top of the model vertically. A heterogeneous thin layer of 10m thick was placed right below the air-filled layer with randomly distributed elastic parameters, to simulate a weathered layer.

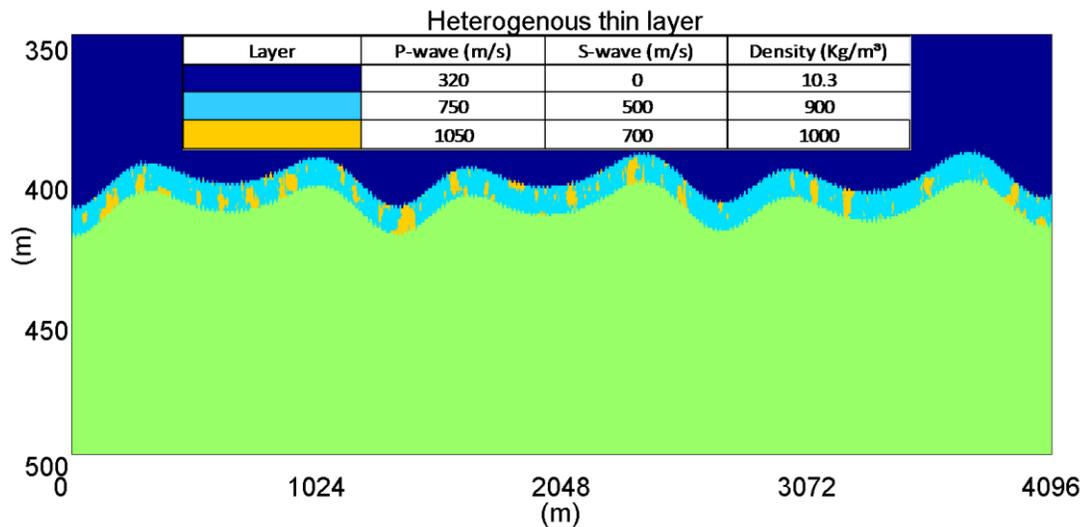


Figure 34. A zoom-in of the land model in Figure 33 that shows the random wiggles of the air-filled layer and the random distribution of the elastic parameters of the thin layer. The elastic parameters of this model have been posted as a table on the model figure for simplicity.

The pseudo-spectral finite-difference tool was used to model the seismic wave propagation through this model. Due to the low elastic parameters of the thin layer and the heterogeneity of our land model, the stability and grid dispersion formulas, discussed and shown in the computational constrains section, would not be specific enough for accurate and stable result because they were derived for homogenous mediums. Therefore, the grid-spacing and the timestep of the finite-difference had to be chosen by trials and errors in a way that could lead to the best results. After going through several iterations of different values, we found that 1 meter for the grid-spacing and 0.05×10^{-3} for the timestep provided accurate and stable numerical results for our objective. An absorbing boundary of 100 grid-points has been applied to all edges of the model, including the top edge, so that no artifact reflections would occur. A vertical source of

20 Hertz central frequency was placed in the center laterally and at 6 meters deep below the air-filled layer. Sensors, at 4 meters apart, which capture the vertical component of the particle velocity, were placed at 8 meters below the free-surface. While outputting every 80th timestep to the shot-record file, a total of 40,000 timesteps was executed. This produced a shot-record of 2 second long, with 4 millisecond sampling-rate. Figure 35 shows the shot-record for the land model. At the far offset, we see that reflections have been masked by dispersive ground roll and the many refracted events. At the near offset, reflections are hardly visible because of the backscattering, which are diffracted events from the wiggly interfaces. The shot-record of this model showed that the pseudo-spectral finite-difference is a powerful tool for simulating complex land models. The more attributes and parameters we account for in our models, the closer the finite-difference simulation would be to the real seismic data. While our land model was simplified, we observed that our simulation looked quite similar to a real land seismic shot-record.

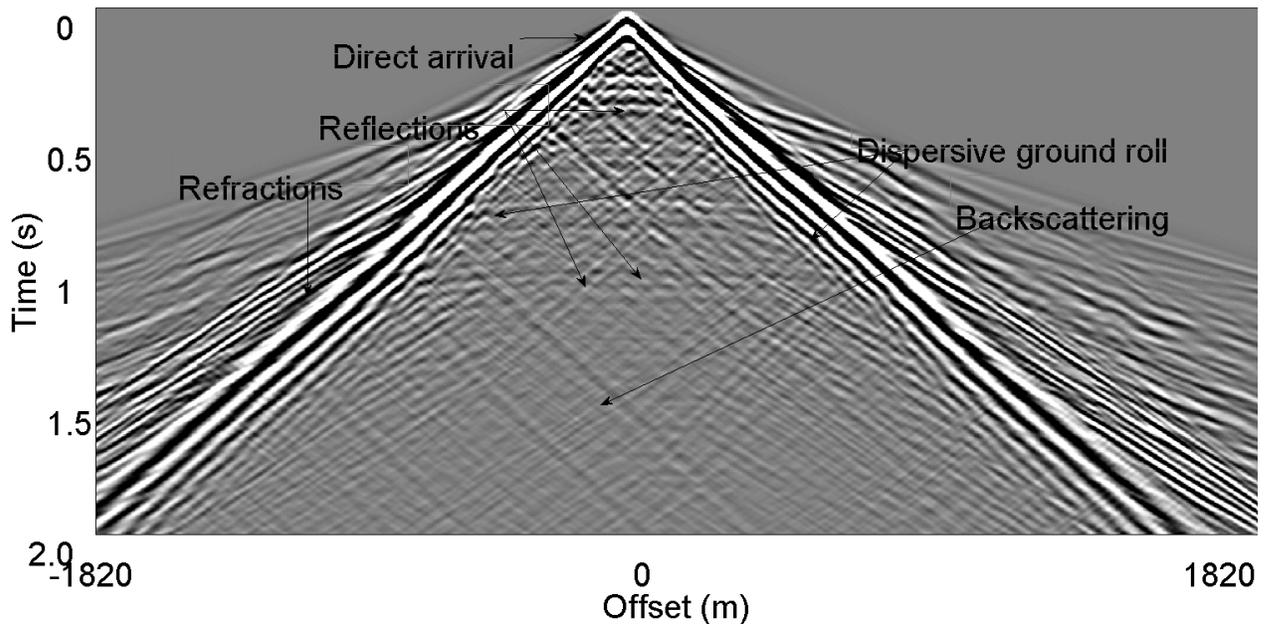


Figure 35. A seismic short-record, simulating seismic land data, of two-second long sampled at 4ms generated by the pseudo-spectral finite-difference.

Another benefit of the pseudo-spectral finite-difference is in cases where large-scale numerical modelings are desired. Large scale numerical modelings are needed in cases like modeling of regional basins or when modeling is aimed beyond sediment columns such as to model faults or large-scale fractures that extend beyond basement. Again, since the pseudo-spectral finite-difference has no constrain on the size of the grid-spacing, except the two grid-points per wavelength, and since the spatial derivatives are computed locally, without the need to use adjacent grid-points, then large grid-points can be used to model large scale mediums. For example, Figure 36 shows a model of several wiggly layers with normal model size of (2048X1024) along with the elastic parameters of its layers. The minimum velocity of the model is 1400 m/s , thus with source maximum frequency of 50 Hz, grid-spacing of 12 meters can be easily used.

Figure 37 shows the seismic shot-record generated with the pseudo-spectral finite-difference for this model. While four-second long was simulated for the shot-record, only 12,000 timesteps were needed for a finite-difference model of normal size. Therefore, the pseudo-spectral finite-difference is an optimum modeling tool for large scale simulation.

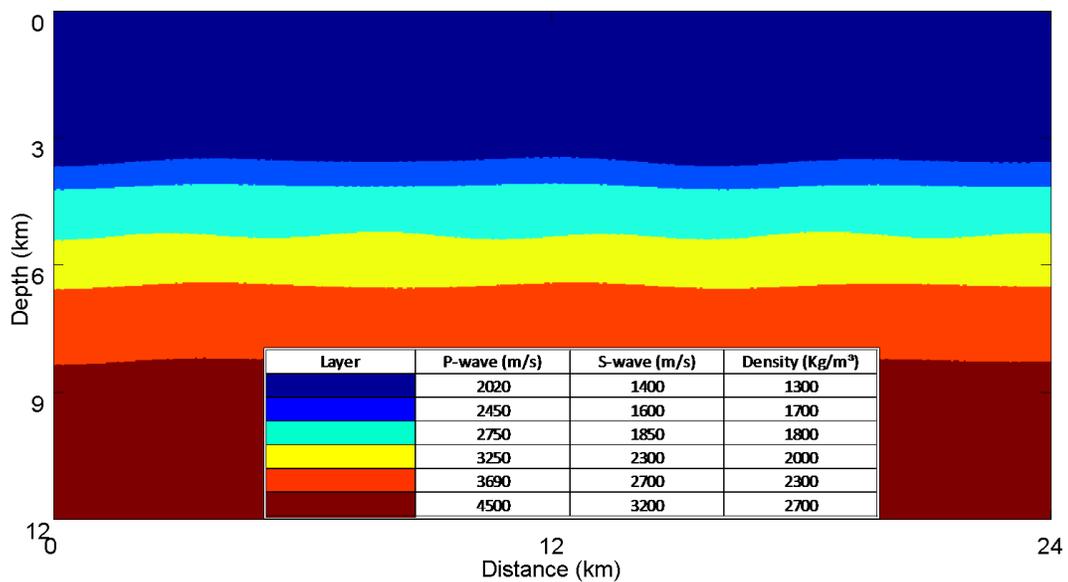


Figure 36. A large-scale model of several wiggly layers, along with its elastic parameters to be used as the input model for the pseudo-spectral finite-difference.

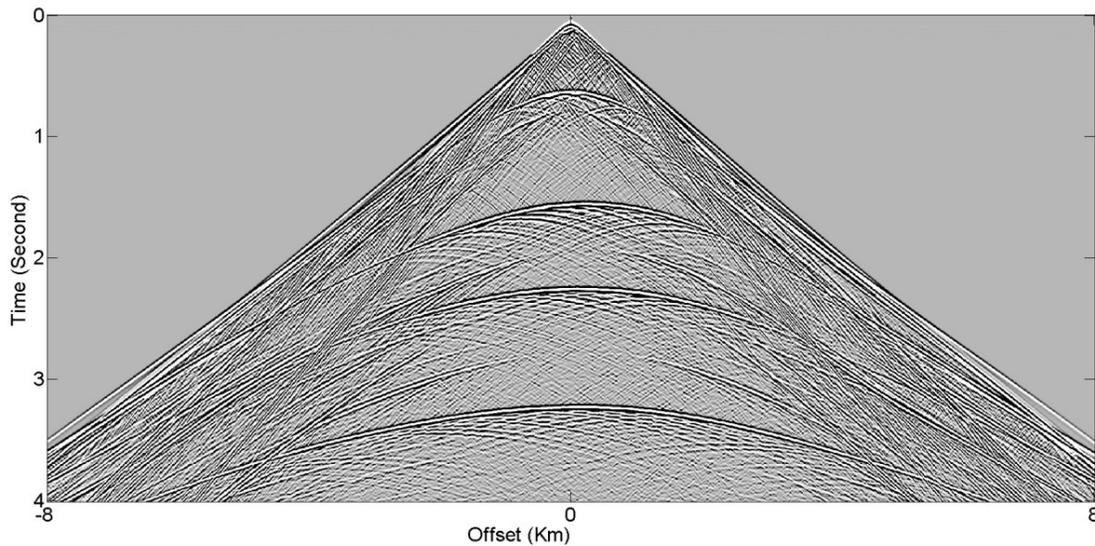


Figure 37. A seismic shot-record simulated with the pseudo-spectral finite-difference for the model in Figure 36. While the shot-record was simulated for 4 second, only 12,000 timesteps were needed for a model of 2048X1024 grid-points.

In conclusion, the pseudo-spectral finite-difference uses the fast Fourier transform, which has an infinite order of accuracy, to compute the spatial derivatives instead of finite differences. As a result, the pseudo-spectral finite-difference only requires two grid-points per wavelength to sample the wavefields, versus 8 grid-points per wavelength for a fourth-order difference operator or 16 grid-points per wavelength for a second-order difference operator, in the staggered-grid finite-difference, to achieve comparable accuracy. The two grid-points per wavelength is a seismic acquisition criteria to prevent aliasing the wavefield, rather than a requirement of the pseudo-spectral finite-difference. Therefore, the pseudo-spectral finite-difference is much more efficient than the staggered-grid finite-difference, particularly for numerical modeling of array data; where the grid-spacing is in the order of one meter or less. Also, for numerical modeling of large inhomogeneities where the size of the grid-spacing is only

constrained by the Nyquist spatial sampling. Also, the pseudo-spectral finite-difference is an optimum numerical modeling tool when broader vertical resolution is needed to resolve thin interfaces or smaller features.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

We have introduced a new concept that we called seismic virtual arrays. By assuming that seismic data are nonGaussian, we can add to our real sensors, virtual sensors, which help us to improve array responses, to reduce the number of real sensors or both. We are in the process of investigating how this new concept of virtual arrays can be integrated into the existing seismic acquisition technology.

One of the key assumptions in the derivation of the virtual seismic array is that array responses are narrowband; that is, the bandwidth of these responses is about 5 Hz or less around central frequency. We know that this assumption does not hold for seismic array responses because the bandwidth of seismic data is about 60 Hz around the central frequency. In other words, seismic data are wideband signals. The future plans, in this study, will include a description of how a wideband signal can be decomposed into narrowband signals, for which the concept of virtual seismic array can be applied. In particular, we will focus on the filterbank technique for decomposing a wideband signal into a group of narrowband signals, and the reconstruction of the wideband signal after constructing the virtual sensors (Harris, 2004).

I have been provided with two finite-difference codes; the staggered-grid and the wavenumber-based (pseudo-spectral) finite differences. I have gained the experience on how to use both codes and made several runs for various geology models. I described a

numbers of them in this thesis. Both finite-difference techniques showed comparable result for modeling various seismic events including first-arrival, primaries, refractions, converted waves and internal multiples. These results confirm that the wavenumber-based finite-difference is also a reliable modeling tool for seismic data. Moreover, I showed that the wavenumber-based finite-difference is more suitable for modeling data at the array level, compared to the staggered-grid finite difference, especially when dealing with land seismic data.

REFERENCES

- Abramowitz, M., and I.A. Stegun, 1972, Handbook of mathematical functions with formulas, graphs, and mathematical tables: Dover.
- Amblard, P.O., and M. Gaeta, 1996, Statistics for complex variables and signals - Part I: Variables: Signal Processing, **53**, 1-13.
- Bracewell, R.N., 2000, The Fourier transform and its applications, McGraw-Hill International.
- Cerjan, C., D. Kosloff, R. Kosloff, and M. Reshef, 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**, 705-708.
- Gazdag, J., 1981, Modeling of the acoustic wave equation with transform methods: Geophysics, **46**, 854-859.
- Harris, F. J., 2004, Multirate signal processing for communication system: Prentice Hall
- Ikelle, L., 2010, Coding and decoding: Seismic data the concept of multishooting: Elsevier.
- Ikelle, L.T., 2005, An analysis of 2D and 3D multiple attenuation for a canonical example: Geophysics, **70**, A13-A28.
- Ikelle, L.T., and L. Amundsen, 2005, Introduction to petroleum seismology: Society of Exploration Geophysicists.
- Knopoff, L., 1952, On rayleigh wave velocities: Bulletin of Seismological Society of America, **42**, 307-308.
- Kosloff, D., M. Reshef, and D. Loewenthal, 1984, Elastic wave calculations by the fouier method: Bulletin of the Seismological Society of America, **74**, 875-891.

- Kosloff, D., and E. Baysal, 1982, Forward modeling by a Fourier method: *Geophysics*, **47**, 1402-1412.
- Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: *Geophysics*, **53**, 1425-1436.
- Madariaga, R., 1976, Dynamics of an expanding circular fault: *Bulletin of the Seismological Society of America* **66**, 639 -666
- Rice, S. O., 1982, Envelopes of narrow-band signals: *Proceedings of the institute of electrical and electronics engineers*, **70**, 692-699.
- Schreier, P.J., and L. L. Scharf, 2006, Higher-order spectral analysis of complex signals: *Signal Processing*, **86**, 3321-3333.
- Singh, S.K., 2005, Sub-Basalt imaging: Modeling and demultiple: M.S.thesis, Texas A&M Univeristy, College Station, Texas.
- Veen, A.-J.d., and G. Leus, 2005, *Signal Processing for Comunication: Delft University of Technology*, et 4147.
- Virieux, J., 1986, P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method: *Geophysics*, **51**, 889-901.
- Wilson, R. J., 2002, Potential impacts of vertical cable seismic: Modeling, resolution and multiple attenuation: M.S.thesis, Texas A&M Univeristy, College Station, Texas.
- Yang, X., 2008, Simulation of seismic real and virtual data using the 3D finite-difference technique and representation theorem: Ph.D. dissertation, Texas A&M University, College Station, Texas.
- Zatnab, M., 1998, How narrow is narrowband: *IEEE Transctions on Signal Processing*, **145**, 85-91.

APPENDIX A

NARROWBAND ARRAY RESPONSE

In this Appendix, I define the complex envelop and show how the real original signal can be recovered. Then, I will derive the array response of the complex envelop for a narrowband signal (Ikelle, 2010; Rice, 1982; Veen and Leus, 2005; Zatnab, 1998).

Complex Envelop

For example, if $S_k(t)$ is a narrow-band real signal, then the complex envelop for $S_k(t)$ with central frequency f_c (where $\omega_c = 2\pi f_c$) can be expressed as follow:

$$\tilde{S}_k(t) = [S_k(t) + i\hat{S}_k(t)]exp[-i\omega_c t] \quad (\text{A.1})$$

where $\tilde{S}_k(t)$ is the complex envelope of $S_k(t)$ and $\hat{S}_k(t)$ is an arbitrary function, which is in practice selected as the Hilbert transform of $S_k(t)$ (Bracewell, 2000). For a function $h(t)$, which is nonsingular at $t = 0$ for $t < 0$, then its Fourier transform, is given by $H(\omega) = R(\omega) \pm iX(\omega)$; where ω is the angular frequency, has the special property known as the Hilbert transform, which is defined as follow:

$$\begin{aligned}
X(\omega) &= -\left(\frac{1}{\pi}\right) R(\omega) * \left(\frac{1}{\omega}\right) \\
&= -\left(\frac{1}{\pi}\right) \mathcal{P} \int R(y) dy / \omega - y
\end{aligned}$$

and (A.2)

$$\begin{aligned}
R(\omega) &= -\left(\frac{1}{\pi}\right) X(\omega) * \left(\frac{1}{\omega}\right) \\
&= -\left(\frac{1}{\pi}\right) \mathcal{P} \int X(y) dy / \omega - y
\end{aligned}$$

Where \mathcal{P} donates the Cauchy principle value at discontinuities. If $H(\omega)$ vanishes for $\omega < 0$, its Fourier transform $h(t) + jx(t)$ has $h(t)$ and $x(t)$ forming a Hilbert transform pair; $h(t)$ and $x(t)$ have the same amplitude spectrum but differ in phase by 90° .

$[h(t) + jx(t)]$ is called the analytic signal belonging to $h(t)$, and $x(t)$ is the quadrature signal corresponding to $h(t)$. So, from the complex envelop of $\hat{S}_k(t)$, the real signal can be reconstructed as:

$$S_k(t) = \Re \left[\hat{S}_k(t) \exp[i\omega_c t] \right] \quad (\text{A.3})$$

where \Re stands for the real part.

Narrowband Array Response

The objective, here, is to prove that the time delay for a narrowband complex enveloped signal is given by (2.3). So the delayed narrowband signal in the time domain $S_k(t - \tau_{lk})$ is given by:

$$S_k(t - \tau_{lk}) = \Re \left[\tilde{S}_k(t - \tau_{lk}) \exp[i\omega_c(t - \tau_{lk})] \right] \quad (\text{A.4})$$

Where $\tilde{S}_k(t - \tau_{lk})$ is the delayed version in the complex envelop and \Re stands for the real part. The Fourier transform of the complex envelop of the delayed narrowband signal is given by:

$$\begin{aligned} \hat{F} \left(\tilde{S}_k(t - \tau_{lk}) \right) &= \int_{-\infty}^{\infty} \tilde{S}_k(t - \tau_{lk}) \exp^{-i\omega t} dt \\ &= \tilde{S}_k(\omega) \exp^{-i\omega \tau_{lk}} \end{aligned} \quad (\text{A.5})$$

The inverse Fourier transform \hat{F}^{-1} of the complex envelop delayed narrowband signal is given by:

$$\begin{aligned}
\hat{F}^{-1}(\tilde{S}_k(t - \tau_{lk})) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{S}_k(\omega) \exp^{-i\omega\tau_{lk}} \exp^{i\omega t} d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{S}_k(\omega) \exp^{i\omega(t - \tau_{lk})} d\omega
\end{aligned} \tag{A.6}$$

Since the signal is a narrowband, then the variable ω can be replaced with $\omega = \omega_c + \hat{\omega}$ where ω_c is the angular central angular frequency and $\hat{\omega}$ is a narrowband signal. Then, (A.6) can be expressed as:

$$= \frac{1}{2\pi} \int_{-\delta\omega}^{\delta\omega} \tilde{S}_k(\omega_c + \hat{\omega}) \exp^{i\omega_c t} \exp^{i\hat{\omega} t} \exp^{i\omega_c \tau_{lk}} \exp^{i\hat{\omega} \tau_{lk}} d\omega \tag{A.7}$$

Assuming that $S_k(t)$ is sufficiently small, then $\exp^{i\hat{\omega} \tau_{lk}} \approx 1$, therefore (A.7) is reduced to the followings:

$$= \frac{1}{2\pi} \exp^{i\omega_c \tau_{lk}} \int_{-\infty}^{\infty} \tilde{S}_k(\omega) \exp^{i\omega t} d\omega = \exp^{i\omega_c \tau_{lk}} \tilde{S}_k(t) \tag{A.8}$$

The well-known conclusion is that, for narrowband signals, time delays shorter than the inverse bandwidth can be represented as phase shift of the complex envelop of the signal.

APPENDIX B

REVIEW ON SECOND- AND HIGHER-ORDER STATISTICS

My objective in this Appendix is to provide a review and some background on statistical averages; mainly on second-order and higher-order statistics. This background is essential to carry-out this research.

I start by defining random variables and discuss an example of representing seismic data as random variable. Then, I show how to statistically characterize random variables by some statistical averages tools such as moments and cumulants. The definitions of these statistical averages tools are then extended to include joint moments and joint cumulants; because we often deal with a number of random variables, rather than a single random variable.

REAL RANDOM VARIABLES

Random Variables

If the outcome from an experiment has multiple possibilities, then the proper description for the experiment is in terms of the different possible outcomes and the probability of each outcome. For instance, suppose the experiment consists of tossing a coin, whose sides are marked as 1 and -1. Then, the possible outcomes from the experiment form a random variable, which may take on either of these values with probability of $\frac{1}{2}$.

In seismic, for example, the different shot-receiver pairs that share the same reflection point, after normal-moveout correction, could be considered as N repeating experiments; where N represents the number of shot-receiver pairs or what is known as fold in seismic data acquisition. Therefore, each timestep, in a common-midpoint gather, after correcting for the normal-moveout, is considered as a random variable that can be characterized by statistical averages, (Ikelle and Amundsen, 2005). Therefore, we could consider the values corresponding to all traces at each timestep as a repeated experiment, or in other words as a random variable, that can be characterized by statistical averages

One of the key statistical averages tools that can be used to characterize random variables is known as moments. Moments are defined as the expectation of any random variable raised to some power. The power depends on the order of moment being calculated. The general formulation of moments is as follows:

$$m_k = E[X^k] = \int_{-\infty}^{\infty} x^k p(x) dx \quad (\text{B.1})$$

where m_k is the k^{th} -order moment and $p(x)$ is the probability density function, which is the frequency of samples appearing in a random variable. A histogram is a graphical representation of a probability density function. The k^{th} order moment is the expected value, where E stands for expectation of the random variable X , raised to some power k . Moments, up to fourth-order, are computed as follow:

$$\text{First – order moment: } m_1 = E[X] = \int_{-\infty}^{\infty} x p(x) dx$$

$$\text{Second – order moment: } m_2 = E[X^2] = \int_{-\infty}^{\infty} x^2 p(x) dx$$

(B.2)

$$\text{Third – order moment: } m_3 = E[X^3] = \int_{-\infty}^{\infty} x^3 p(x) dx$$

$$\text{Fourth – order moment: } m_4 = E[X^4] = \int_{-\infty}^{\infty} x^4 p(x) dx$$

For example, for a Gaussian random variable that has zero mean, moments, up to fourth-order, are computed as follow:

First-order moment:

$$m_1 = E[X] = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x)^2}{2\sigma^2}\right] dx = 0$$

Second-order moment:

(B.3)

$$m_2 = E[X^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x)^2}{2\sigma^2}\right] dx = \sigma^2$$

Third-order moment:

$$m_3 = E[X^3] = \int_{-\infty}^{\infty} x^3 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x)^2}{2\sigma^2}\right] dx = 0$$

Fourth-order moment:

$$m_4 = E[X^4] = \int_{-\infty}^{\infty} x^4 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x)^2}{2\sigma^2}\right] dx = 3\sigma^4$$

where σ^2 is the variance. This example shows that for Gaussian random variables, first- and third-order moments are always zero, while second-order moment is non-zero and gives the variance. The fourth-order moment is always a multiple of the second-order moment. If we keep computing higher-order moments, we will find that for Gaussian random variables, an infinite number of moments exist. Odd moments are always zero while even moments are always non-zero and are multiple of the second-order moments.

Another useful key statistical averages tool is known as cumulants. However, to define cumulants, I need to define other quantities known by first and second characteristic functions. The first characteristic function, $\Phi(\omega)$, is the Fourier transformation of the probability density function, $p(x)$. The second characteristic function, $\Psi(\omega)$, is the natural log of the first characteristic function. Mathematically, they are expressed as follow:

First characteristic function:

$$\Phi(\omega) = \int_{-\infty}^{\infty} \exp[i\omega x]p(x)dx$$

Second characteristic function:

(B.4)

$$\Psi(\omega) = \ln\Phi(\omega)$$

Cumulants are found by differentiating the second characteristic function, $\Psi(\omega)$, with respect to ω at the origin.

$$C_k = (-i)^k \left. \frac{\partial^k \Psi(\omega)}{\partial^k \omega} \right|_{\omega=0} \quad (\text{B.5})$$

Cumulants up to fourth-order are computed as follow:

$$\text{First-order cumulant: } C_1 = (-i) \left. \frac{\partial \Psi(\omega)}{\partial \omega} \right|_{\omega=0}$$

$$\text{Second-order cumulant: } C_2 = (-i)^2 \left. \frac{\partial^2 \Psi(\omega)}{\partial^2 \omega} \right|_{\omega=0} \quad (\text{B.6})$$

$$\text{Third-order cumulant: } C_3 = (-i)^3 \left. \frac{\partial^3 \Psi(\omega)}{\partial^3 \omega} \right|_{\omega=0}$$

$$\text{Fourth-order cumulant: } C_4 = (-i)^4 \left. \frac{\partial^4 \Psi(\omega)}{\partial^4 \omega} \right|_{\omega=0}$$

It is important to mention that I am only discussing and showing moments and cumulants up to fourth-order because it is the highest order of moments and cumulants I would need to carry-out this research. For example, for a random variable X , with Gaussian probability density function given by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\eta)^2}{2\sigma^2}\right] \quad (\text{B.7})$$

The first characteristic function is:

$$\Phi(\omega) = \int_{-\infty}^{\infty} \exp[i\omega x] * \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x^2}{2\sigma^2}\right] dx = \exp\left[\frac{-\omega^2\sigma^2}{2}\right] \quad (\text{B.8})$$

And the second characteristic function is:

$$\Psi(\omega) = \ln\Phi(\omega) = \sigma^2 \frac{(i\omega)^2}{2} \Big|_{\omega=0} \quad (\text{B.9})$$

By differentiating (B.9) with respect to ω at the origin, I can find various orders of cumulants of the random variable, X .

$$c_1 = 0 \quad (\text{B.10})$$

$$c_2 = \sigma^2$$

$$c_k = 0, \text{ for } k > 2$$

This example shows, that for Gaussian random variables, cumulants up to second-order are equivalent to moments up to second-order. However, cumulants higher than second-order are always zero while moments higher than second-order are not necessary zero, as shown in (B.3). It also shows how computing cumulants is effective in characterizing the non-Gaussianity of random variables than moments. When cumulants higher than second-order are zero, the random variables are characterized as Gaussian and when they are not zero, they are characterized as non-Gaussian.

Cumulants and moments are different, although clearly related. Cumulants are not determined directly by integrative processes as are moments rather they are derived from the second characteristic function. However, cumulants can also be found by first estimating moments and vice versa (Ikelle, 2010), according to the following two formulas:

$$m_n^X = \sum_{i=0}^{n-1} \binom{n-1}{i} c_{n-i}^{(X)} m_i^{(X)} \quad (\text{B.11})$$

$$c_n^X = m_n^X - \sum_{i=1}^{n-1} \binom{n-1}{i} c_{n-i}^{(X)} m_i^{(X)} \quad (\text{B.12})$$

where $\binom{n-1}{i}$ is defined as $\frac{(n-1)!}{((n-1)-i)!i!}$.

Actually, one of the classical ways of differentiating between Gaussian and non-Gaussian random variables is to compute their third- or fourth-order cumulants. When cumulants of third-order and higher are zero, the random variables are characterized as Gaussian and when they are nonzero, they are characterized as non-Gaussian random variables. These facts promote the study of the third- and the fourth-order cumulants key roles in analyzing the non-Gaussianity of random variables. In the following sections, I discuss their meaning and what they actually measure for single random variable and for joint random variables.

Third-order cumulant for a single random variable is the skewness. Skewness is a measure of the asymmetry of the distribution of random variables. It can be positive, negative or zero. A positive skew indicates that the tail, on the right-side of the probability density function, is longer than the left-side, which means that the mass of distribution is concentrated on the left-side. On other hand, a negative skew indicates that the tail on the left-side of the probability density function is longer than the right-side, which means that the mass of distribution is concentrated on the right-side. A zero skew indicates that the values are relatively distributed evenly on both sides of the mean; implying a symmetric distribution. Therefore, the third-order cumulant, for any symmetric probability density functions whether Gaussian or non-Gaussian, is null. As a result, the third-order cumulant might not be reliable statistical tool that could evaluate the non-Gaussianity in random variables when the distribution is symmetric. In such case, the fourth-order cumulant must be considered instead.

The fourth-order cumulant for a single random variable is the kurtosis. It is a measure of peakedness of the probability density function. Random variables that have negative kurtosis tend to have their distribution flatten in relative to Gaussian distributions. These types of distributions are known as sub-Gaussian. The Uniform distribution is an example for this class of distributions. On the other hand, random variables that have positive kurtosis tend to have their distributions peaked more than Gaussian distributions. These types of distributions are known as super-Gaussian. The Laplace distribution is an example for this class of distributions. By examining moments and cumulants of these different distributions, I could highlight the followings:

- First of all, for distributions that are centered at the zero along the X-axis and symmetric, such as the Gaussian, the Laplace and the Uniform distributions, first-order moments and first-order cumulants are zero, while for the Rayleigh distribution, which is not centered at zero along the X-axis and is not symmetric, its first-order moment and its first-order cumulant is not zero.
- Secondly, for distributions that are symmetric, such as the Gaussian, Laplace and Uniform, their third-order moment and their third-order cumulant are also zero, while for the Rayleigh distribution is not zero because of its asymmetry.
- Third, for Gaussian distributions, cumulants higher than second order, is null while higher-order moments are not.
- Fourth, for the Laplace distributions, fourth-order cumulant is positive. This distribution is a type of the super-Gaussian. However, for the Uniform

distribution, fourth-order cumulant is negative. This distribution is a type of the sub-Gaussian.

- Finally, for zero mean distributions, such as the Gaussian, the Laplace and the Uniform, first- to third-order moments are equal to first- to third-order cumulant, only at the fourth-order, they become different. While for the Rayleigh distribution, which are not zero mean, its moments and cumulants become different from the second-order.

For zero mean random variables, cumulants are significantly simplified, as a result of the vanishing terms that contains the first-order moments. Table 2.3 shows cumulants up to fourth-order, as functions of moments, for zero mean versus non-zero mean random variables. For the zero mean, cumulants are equivalent to moments up to third-order. Cumulants only begin to differ at the fourth-order.

Joint Random Variables

We often must deal with multiple random variables instead of only a single random variable. Hence, it is useful to extend the definitions of moments and cumulants to a set of random variables.

Joint moments

Given a set of I real random variables $\{X_1, X_2, \dots, X_I\}$, their joint moments of order $r = k_1 + k_2 + \dots + k_I$ are given by their joint expectation:

$$m_{k_1, k_2, \dots, k_I} = \text{Mom}[X_1^{k_1}, X_2^{k_2}, \dots, X_I^{k_I}] \quad (\text{B.13})$$

For instance, for two joint random variables, XY , the joint moment is given by:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy p_{xy}(x, y) dx dy \quad (\text{B.14})$$

where E stands for their joint expectation and $p_{xy}(x, y)$ is their joint probability density function. A scatter plot is a graphical representation of a joint probability density function of two random variables.

Joint cumulants

As I had to introduce the first and second characteristic function in order to define cumulant for a single random variable. I will also have to define the first and second joint characteristic functions to be able define joint cumulants. So, the first joint characteristic function is the Fourier transform of the joint probability density function.

$$\Phi(\omega_1, \omega_2, \dots, \omega_n) = E\{\exp[i(\omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n)]\} \quad (\text{B.15})$$

We can expand (B.15) by expressing it as follow

$$\Phi(\boldsymbol{\omega}) = E\{\exp[i(\boldsymbol{\omega} \cdot \mathbf{x})]\} = \int_{V_x} dV_x \exp[i(\boldsymbol{\omega} \cdot \mathbf{x})] p(\mathbf{x}) \quad (\text{B.16})$$

where the frequencies ω_k and values of random variable x_j are expressed in form of vectors: $\boldsymbol{\omega} = [\omega_1, \omega_2, \dots, \omega_n]^T$, $\mathbf{x} = [x_1, x_2, \dots, x_n]$, $dV_x = dx_1 dx_2 \dots dx_n$ and $p(\mathbf{x})$ is the n -variable joint probability density function of the random variables $\{X_1, X_2, \dots, X_n\}$. So, the joint cumulant of order $r = k_1 + k_2 + \dots + k_n$ is defined as the derivative of the second joint characteristic function at the origin.

$$\text{Cum}[X_1^{k_1}, X_2^{k_2}, X_n^{k_n}] = (-i)^r \frac{\partial^r \ln \Phi(\omega_1, \omega_2, \dots, \omega_n)}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \dots \partial \omega_n^{k_n}} \Bigg|_{\omega_1 = \omega_2 = \dots = \omega_n = 0} \quad (\text{B.17})$$

For zero-mean random variables, $\text{Cum}[X_1] = E[X_1] = 0$, the joint cumulants up to fourth-order are explicitly given by:

$$\begin{aligned} \text{Cum}[X_1, X_2] &= E[X_1 X_2] \\ \text{Cum}[X_1, X_2, X_3] &= E[X_1 X_2 X_3] \\ \text{Cum}[X_1, X_2, X_3, X_4] &= E[X_1 X_2 X_3 X_4] - E[X_1 X_2] E[X_3 X_4] \\ &\quad - E[X_1 X_3] E[X_2 X_4] - E[X_1 X_4] E[X_2 X_3] \end{aligned} \quad (\text{B.18})$$

When $X_1 = X_2 = X_3 = X_4$, then (2.18) becomes as follow:

$$\text{Cum}[X_1, X_2] = E[X^2] = \textit{variance} \quad (\text{B.19})$$

$$\text{Cum}[X_1, X_2, X_3] = E[X^3] = \textit{skewnes}$$

$$\text{Cum}[X_1, X_2, X_3, X_4] = E[X^4] - 3E[X^2] = \textit{kurtosis}$$

Joint second-order statistics

For two zero-mean random variables, $\mathbf{X} = [X_1, X_2]^T$, the second-order joint cumulants are equal to the second-order joint moments.

$$\text{Cum}[X_1, X_2] = \text{Mom}[X_1, X_2] = E[X_1 X_2] \quad (\text{B.20})$$

In practice, for two random variables, there are four second-order joint cumulants;

$\text{Cum}[X_1, X_1]$, $\text{Cum}[X_1, X_2]$, $\text{Cum}[X_2, X_1]$, and $\text{Cum}[X_2, X_2]$. These second-order cumulants are normally grouped into 2X2 second-rank tensor or matrix, called the covariance matrix, and denoted $\mathbf{C}_X^{(2)}$. The diagonals of the covariance matrix, $\mathbf{C}_X^{(2)}$, will give the variances, $\sigma_{X_1}^2$ and $\sigma_{X_2}^2$, of the random variables X_1 and X_2 , since $E[X_1, X_1] = E[X_1^2]$ and $E[X_2, X_2] = E[X_2^2]$, while the off-diagonals are symmetric, since $E[X_1, X_2] = E[X_2, X_1]$, and will give the second-order joint expectation of these two random variables. Therefore, the covariance matrix, $\mathbf{C}_X^{(2)}$, for two random variables, X_1 and X_2 , is expressed as follow:

$$\mathbf{C}_X^{(2)} = \begin{pmatrix} E[X_1, X_1] & E[X_1, X_2] \\ E[X_2, X_1] & E[X_2, X_2] \end{pmatrix} = \begin{pmatrix} \sigma_{X_1}^2 & E[X_1, X_2] \\ E[X_1, X_2] & \sigma_{X_2}^2 \end{pmatrix} \quad (\text{B.21})$$

For I random variables, the covariance matrix, $\mathbf{C}_X^{(2)}$, is $I \times I$ matrix and expressed as follow:

$$\begin{aligned} \mathbf{C}_X^{(2)} &= \begin{pmatrix} E[X_1, X_1] & \cdots & E[X_1, X_I] \\ \vdots & \ddots & \vdots \\ E[X_I, X_1] & \cdots & E[X_I, X_I] \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{X_1}^2 & \cdots & E[X_1, X_I] \\ \vdots & \ddots & \vdots \\ E[X_1, X_I] & \cdots & \sigma_{X_I}^2 \end{pmatrix} \end{aligned} \quad (\text{B.22})$$

Joint fourth-order cumulants

For I random variables, the third-order joint cumulants will produce a third-rank tensor that will have $I \times I \times I$ components. However, for symmetric zero-mean distributions the third-order joint cumulants are always zero. In such a case, the fourth-order joint cumulant is the key statistical quantity for determining whether the random variables are Gaussian or non-Gaussian.

Joint cumulants of fourth-order form a fourth-rank tensor, which is an array of four indices; denoted normally by $\mathbf{C}_X^{(4)}$. The fourth-order joint cumulants are not equal to the fourth-order joint moments, as are the cases for second- and third-order joint

cumulants for zero-mean random variables. For the random vector $\mathbf{x} = [x_1, x_2, \dots, x_l]^T$, the scalar elements of the fourth-order joint cumulant tensor are

$$\begin{aligned} \text{Cum}[X_i, X_j, X_k, X_l] &= E[X_i X_j X_k X_l] - E[X_i X_j]E[X_k X_l] \\ &\quad - E[X_i X_k]E[X_j X_l] - E[X_i X_l]E[X_j X_k] \end{aligned} \quad (\text{B.23})$$

Where i, j, k and l vary from 1 to l . Therefore, the fourth-order joint cumulant tensor is an array of l^4 elements that can be organized into either $l^2 \times l^2$ matrix or into l^2 matrices of $l \times l$ elements for each matrix. For instance, for two random variables, the fourth-order joint cumulants will form a tensor of 16 elements, which can be grouped into either a matrix of 4×4 or into four matrices of 2×2 for each matrix.

In this chapter, I have shown that for a Gaussian random variable, an infinite of moments exist, where the odd moments are zero and the even moments are multiple of the second moments. However, for Gaussian random variable, cumulants higher than second order is zero. Also, for symmetric zero-mean random variables, third-order cumulants is zero. Therefore, fourth-order cumulants is the key statistical tool that must be considered for non-Gaussian random variables. I also have shown that for zero-mean random variables, cumulants and moments only begin to differ at the fourth-order. Finally, I have also shown that for l random variables, the second-order joint cumulants can be grouped into

$l \times l$ matrix, where the diagonals give the variances for each random variable and the off-diagonals give the second-order joint expectations of the random variables.

COMPLEX RANDOM VARIABLES

A complex random variable Z is defined by two real random variables Z_R and Z_I , such that $Z = Z_R + i Z_I$, where Z_R represents the real part and Z_I represent the imaginary part and $i = \sqrt{-1}$. The possible outcomes, for a complex random variable Z , are represented as $z = z_R + iz_I$, where z_R and z_I are real numbers.

Complex random variables are common in a number statistical variables and signal processing including the Hilbert transformation and the complex envelop domain, which is the domain used in the derivation of the virtual seismic array. Thus, it is important that my review of second- and higher-order statistics include a review on complex random variables and their second- and higher-order statistics. So, my objective, in this section, is to redefine moments and cumulants for complex-valued random variables or (complex random variables, for short), and extend these redefinitions to include joint moments and joint cumulants of complex random variables.

There are two basic approaches in deriving moments and cumulants for complex random variables. One approach is to consider the real and imaginary parts of the complex random variables as two-dimensional real random variables. In such case, the formulas, shown and discussed in the previous section, could be easily used for the complex random variables by replacing each complex random variable with a two-

dimensional real random vector. The downside of this approach is that it produces complicated formulas, particularly when defining joint moments and joint cumulants for complex random variables. The other approach is to derive formulas, which directly operate on complex random variables. In such case, the separation between the real and imaginary parts of the complex random variables is avoided. In this approach, formulas for moments, cumulants, joints moments and joint cumulants of complex random variables are much simplified in terms of taking their derivatives and the physical sense related to the complex nature of the data. Here, I present this approach and follow (Amblard and M. Gaeta, 1996; Ikelle, 2010; Schreier and Scharf, 2006), who provided very unique and clear review on complex random variables.

Circular complex random variables

To define moments and cumulants for complex random variables, the probability density function and the first and second characteristic functions need to be defined, as was done for the moments and cumulants of the real random variables. However, the shapes of distributions, for seismic random variables in the complex envelop domain, follow a class of complex random variables known as circular complex random variables. Thus, the review of complex random variables, in this section, is limited to this class of complex random variables. More discussion regarding the circularity of seismic data will follow in the next section.

The probability density functions for circular complex random variables have special characteristics that make them unique. So, my objective, here, is to highlight

these characteristics and show their distributions as 1D function, instead of functions of their real and imaginary parts. Distributions of these two complex random variables are centered at $(z_R =, z_I = 0)$, which means that both z_R and z_I are zero-mean real random variables. Therefore, circular complex random variables are also zero-mean, $E[Z] = 0$ implies $E[Z_R] = E[Z_I]=0$.

Also, the distributions of circular complex random variables are radially isotropic. They only depend on $\rho^2 = z_R^2 + z_I^2$. In a general term, a complex random variable is considered circular if Z and $Z \exp(i\varphi)$ have the same probability density function. In other words, the distributions of circular complex random variables are invariant under complex rotation; they only depend on the modulus of z . Because, the derivations are carried-out directly on complex random variables, then the samples of circular complex random variable z can be represented as:

$$z = \rho \exp(i\varphi) \quad (\text{B.24})$$

where

$$\rho = \sqrt{z_R^2 + z_I^2} \quad (\text{B.25})$$

is the modulus and φ is the phase, then the distribution of Z is equivalent to the distribution of its modulus ρ . The phase φ is always uniformly distributed on $[-\pi \pi]$, thus has no contribution to the type of distribution. Therefore, instead of representing the probability density function of a circular complex random variable as functions of it

its real and imaginary part (see Ikelle, 2010), it can be represented as a single function of ρ . So, the probability density function for a circular Gaussian random variable, as a function of the modulus ρ , follows a Rayleigh function, instead of the Gaussian distribution for real random variables.

Also, when operating directly on the complex random variables, the probability density function of z becomes a joint function of z and its complex conjugate \bar{z} , which is defined as $\bar{z} = z_R^2 - z_I^2$. Thus, for circular complex random variables, the probability density function is not a function of z and \bar{z} , but a function of the product, $z\bar{z} = |z|^2 = z_R^2 + z_I^2$. For example, the probability density function of circular Gaussian complex random variable is a product of its real part function Z_R and its imaginary part function Z_I , which is given by:

$$g_{R,I}(R, I) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{z_R^2}{2\sigma^2}\right) \exp\left(-\frac{z_I^2}{2\sigma^2}\right) \quad (\text{B.26})$$

However, operating directly on the complex random variables make the distribution a product of $z\bar{z}$ only. Hence, the probability density function becomes:

$$g_{z,\bar{z}}(z, \bar{z}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{z\bar{z}}{2\sigma^2}\right) \quad (\text{B.27})$$

In the circular Gaussian case, the probability density function is a product of the real function Z_R and the imaginary function Z_I , thus they are statistically independent; in accordance with the definition of statistical independence for real random variables.

However, for the non-Gaussian case, the two components Z_R and Z_I are not necessary statistically independent because $z\bar{z}$ is not necessary a product of a function of Z_R and a function of Z_I , as in the Gaussian case. However, to derive the first and second characteristic functions for circular complex random variables, which are needed to compute moments and cumulants for circular complex random variables, an assumption of statistical independence is made for the variables z and \bar{z} , which is solely for calculation purposes, (Ikelle, 2010). So, to define the first characteristic function, two complex variables ω and $\bar{\omega}$ are defined as $\omega = u + iv$ and $\bar{\omega} = u - iv$, where u and v are real numbers. Then, the first characteristic function is given by:

$$\begin{aligned}\Phi_{z,\bar{z}}(\omega, \bar{\omega}) &= E \left[\exp \left\{ \frac{1}{2} i(\omega\bar{z} + \bar{\omega}z) \right\} \right] \\ &= \int_D \exp \left\{ \frac{1}{2} i(\omega\bar{z} + \bar{\omega}z) \right\} g_{z,\bar{z}}(z, \bar{z}) d_z d_{\bar{z}}\end{aligned}\tag{B.28}$$

where $g_{z,\bar{z}}(z, \bar{z})$ is the joint probability density function of z and \bar{z} . The argument of the exponential function can also be defined as follows:

$$\frac{1}{2}(\omega\bar{z} + \bar{\omega}z) = uz_R + vz_I\tag{B.29}$$

Then, the first characteristic function can be expressed as:

$$\Phi_{z,\bar{z}}(\omega, \bar{\omega}) = E[\exp\{i(uz_R + vz_I)\}] \quad (\text{B.30})$$

The first characteristic function of the complex variables ω and $\bar{\omega}$ is the extension to the two dimensional complex plane, of the characteristic function for the real random variables defined for the real variables. Thus, the integral of the expectation is calculated on all subsets of z and \bar{z} generated by z_R and z_I . Also, when z is real, it is obvious that the first characteristic function for complex random variables (B.29) is reduced to the classical first characteristic function for real random variable (B.4). The second characteristic function, which is used to compute cumulants, is the natural log of the first characteristic function, as for the real case.

$$\Psi_{z,\bar{z}}(\omega, \bar{\omega}) = \ln \Phi_{z,\bar{z}}(\omega, \bar{\omega}) \quad (\text{B.31})$$

For example, the first characteristics function for a circular Gaussian complex random variable, with a probability density function given by (B.25) (Ikelle, 2010), is as follow:

$$\Phi_{z,\bar{z}}(\omega, \bar{\omega}) = \exp\left(-\frac{\sigma^2 \omega \bar{\omega}}{2}\right) \quad (\text{B.32})$$

While the second characteristics function for the same distribution is as follow:

$$\Psi_{z,\bar{z}}(\omega, \bar{\omega}) = \ln\left(\exp\left(-\frac{\sigma^2 \omega \bar{\omega}}{2}\right)\right) = -\frac{\sigma^2 \omega \bar{\omega}}{2} \quad (\text{B.33})$$

For circular complex random variables, the first and second characteristic functions are functions of the product of $\omega\bar{\omega}$, as shown by the first and second characteristic functions for the circular Gaussian complex random variable (B.31 and B.32). So, a general definition of circular complex random variables is that its probability density function is a function of a product of $z\bar{z}$, or alternatively, its first and second characteristic functions is a product of $\omega\bar{\omega}$.

Similarly, as was for the real random variables, moments and cumulants, for the complex random variables, are defined as the coefficients of the Taylor expansion of the first and second characteristic functions respectively, with respect to its variables ω and $\bar{\omega}$. Thus, any moments of order n is found by differentiating the first characteristic function n times and set $\omega = \bar{\omega} = 0$. So, the n^{th} order moment for the complex random variable Z is found by:

$$\begin{aligned} m_{p,n-p}^{(Z)} &= \text{Mon}[Z^p \bar{Z}^{n-p}] = E[Z^p \bar{Z}^{n-p}] \\ &= \frac{2^n}{i^2} \left. \frac{\partial^n \Phi_{Z,\bar{Z}}(\omega, \bar{\omega})}{\partial \bar{\omega}^p \partial \omega^{n-p}} \right|_{\omega = \bar{\omega} = 0} \end{aligned} \quad (\text{B.34})$$

The n^{th} order moment of a complex random variable does not only depend on n , but also on an arbitrary integer p that can take the value between 0 and n . Thus, there are $n + 1$ different moments that need to be compute in order to describe a complex random variable completely. This result comes from the fact that not only the complex random variable Z is considered, but also its complex conjugate \bar{Z} . For example at order 2, there

are three moments: $m_{0,2}^{(Z)} = E[\bar{Z}^2]$, $m_{1,1}^{(Z)} = E[Z\bar{Z}]$ and $m_{2,0}^{(Z)} = E[Z^2]$. For a real random variable, these three moments are the same, but they are different for a complex random variable. However, for circular complex random variable all moments are null except the ones that have the same power in Z and \bar{Z} . Thus, the only non-null moments are the ones of the following form:

$$m_{p,p}^{(Z)} = \text{Mon}[Z^p \bar{Z}^p] = E[Z^p \bar{Z}^p] \quad (\text{B.35})$$

This results comes from the fact that the characteristic function of circular complex random variables is a product of $\omega \bar{\omega}$. In other words, when all moments $m_{p,q}$ with $p \neq q$ are zero, the Taylor expansion for the characteristic function reduces to a function of $\omega \bar{\omega}$ only.

Computing cumulants for complex random variables are done similarly, except the second characteristic function is used instead. So, the Taylor expansion of $\Psi_{z,\bar{z}}(\omega, \bar{\omega})$ with respect to ω and $\bar{\omega}$ lead to the followings:

$$c_{p,n-p}^{(Z)} = \text{Cum}[Z^p \bar{Z}^{n-p}] = \frac{2^n}{i^2} \frac{\partial^n \Psi_{z,\bar{z}}(\omega, \bar{\omega})}{\partial \bar{\omega}^p \partial \omega^{n-p}} \Big|_{\omega = \bar{\omega} = 0} \quad (\text{B.36})$$

Similarly to moments for complex random variables, cumulants, at the n^{th} order, have $n + 1$ cumulants because the second characteristic function is a product of $\omega \bar{\omega}$.

Considering circular complex random variables, then the only nonzero cumulants are the ones of the same power.

$$c_{p,p}^{(Z)} = \text{Cum}[Z^p \bar{Z}^p] = \text{E}[Z^p \bar{Z}^p] \quad (\text{B.37})$$

For example, computing the second order cumulants from the second characteristic function, for the circular Gaussian complex random variable (B.32), show that $c_{0,2}^{(Z)} = c_{2,0}^{(Z)} = 0$. The only nonzero cumulant for circular Gaussian complex random variable is $c_{1,1}^{(Z)} = \text{variance} (\sigma^2)$. Cumulants, for $n > 2$, are always zero because of the Gaussianity, which is a result discussed and shown previously. One important aspect of circular complex random variables is that the first- and third-order cumulants, which are equivalent to the mean and skewness respectively, are always zero. The second- and fourth-order cumulants which are equivalent to the variance and kurtosis respectively are real numbers; despite the fact the random variables are complex-valued.

Joint complex random variables

In this section, I extend the definitions of moments and cumulants to joint moments and joint cumulants of complex random variables, just as was done for the real random variables. I will particularly focus on the second- and fourth-order statistics for complex random variables.

Joint moments

As shown previously, moments are derived from the first characteristic function.

So, the first characteristic function for I joint complex random variables is defined as follow:

$$\begin{aligned}
 \Phi(\omega_1, \omega_2, \dots, \omega_I, \bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_I) \\
 &= E \left\{ \exp \left[\frac{1}{2} i (\omega_1 \bar{z}_1 + \omega_2 \bar{z}_2 + \dots + \omega_I \bar{z}_I + \bar{\omega}_1 z_1 \right. \right. \\
 &\quad \left. \left. + \bar{\omega}_2 z_2 + \dots + \bar{\omega}_I z_I) \right] \right\}
 \end{aligned} \tag{B.38}$$

Where $\omega_1, \omega_2, \dots, \omega_I$ are complex variables. So, the joint moment of order $r = k_1 + k_2 + \dots + k_I$ for I joint complex random variables Z_1, Z_2, \dots, Z_I is defined as follows:

$$\begin{aligned}
 &\text{Mon}[Z_1^{p_1}, Z_2^{p_2}, \dots, Z_I^{p_I}, \bar{Z}_1^{k_1-p_1}, \bar{Z}_2^{k_2-p_2}, \dots, \bar{Z}_I^{k_I-p_I}] \\
 &= E[Z_1^{p_1}, Z_2^{p_2}, \dots, Z_I^{p_I}, \bar{Z}_1^{k_1-p_1}, \bar{Z}_2^{k_2-p_2}, \dots, \bar{Z}_I^{k_I-p_I}] \\
 &= \frac{2^r}{i^r} \frac{\partial^r \Phi_{z, \bar{z}}(\omega_1, \omega_2, \dots, \omega_I, \bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_I)}{\partial \bar{\omega}_1^{p_1} \partial \bar{\omega}_2^{p_2} \dots \partial \bar{\omega}_I^{p_I} \partial \omega_1^{k_1-p_1} \partial \omega_2^{k_2-p_2} \dots \partial \omega_I^{k_I-p_I}} \Bigg|_{\omega_1=\omega_2=\dots=\omega_I=\bar{\omega}_1=\bar{\omega}_2=\dots=\bar{\omega}_I=0}
 \end{aligned} \tag{B.39}$$

Where p_i is an integer that can vary between 0 and k_i .

Joint cumulants

Joint cumulants are derived similarly as joint moments except the second characteristic function is used instead.

$$\begin{aligned} \psi(\omega_1, \omega_2, \dots, \omega_I, \bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_I) \\ = \ln \Phi(\omega_1, \omega_2, \dots, \omega_I, \bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_I) \end{aligned} \quad (\text{B.40})$$

So, the joint cumulants of order $r = k_1 + k_2 + \dots + k_I$ for I joint complex random variables Z_1, Z_2, \dots, Z_I are defined as follows:

$$\begin{aligned} \text{Cum}[Z_1^{p_1}, Z_2^{p_2}, \dots, Z_I^{p_I}, \bar{Z}_1^{k_1-p_1}, \bar{Z}_2^{k_2-p_2}, \dots, \bar{Z}_I^{k_I-p_I}] \\ = \frac{2^r}{i^r} \frac{\partial^r \psi(\omega_1, \omega_2, \dots, \omega_I, \bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_I)}{\partial \bar{\omega}_1^{p_1} \partial \bar{\omega}_2^{p_2} \dots \partial \bar{\omega}_I^{p_I} \partial \omega_1^{k_1-p_1} \partial \omega_2^{k_2-p_2} \dots \partial \omega_I^{k_I-p_I}} \Bigg|_{\omega_1=\omega_2=\dots=\omega_I=\bar{\omega}_1=\bar{\omega}_2=\dots=\bar{\omega}_I=0} \end{aligned} \quad (\text{B.41})$$

When $p_1 = p_2 = \dots = p_I = 0$ in (B.40) and (B.41), they reduce to the formulas of joint moments and joint cumulants for real random variables.

Second-order joint cumulants

The second-order cumulant is one of the tools used derive the virtual seismic array. Therefore, I limit the formulas of joint cumulants (B.40) to the second-order joint cumulants ($r = 2$), and for the class of circular complex random variables, which are characterized as zero-mean. If I consider a complex random vector made of two circular

complex random variables $\mathbf{Z} = [Z_1, Z_2]$. The second-order joint cumulants are equal to the second-order joint moments.

$$\begin{aligned} \text{Cum}[Z_i, \bar{Z}_j] &= \text{Mon}[Z_i, \bar{Z}_j] = E[Z_i, \bar{Z}_j], \\ &\text{where } i, j = 1, 2. \end{aligned} \quad (\text{B.42})$$

The second-order joint cumulants for two circular complex random variables consist of four joint cumulants that can be grouped in a square matrix known as the covariance matrix denoted as $C_{\mathbf{Z}}^{(2)}$ and defined as follows:

$$C_{\mathbf{Z}}^{(2)} = \begin{pmatrix} E[Z_1, \bar{Z}_1] & E[Z_1, \bar{Z}_2] \\ E[Z_2, \bar{Z}_1] & E[Z_2, \bar{Z}_2] \end{pmatrix} = \begin{pmatrix} \sigma_{Z_1}^2 & E[Z_1, \bar{Z}_2] \\ E[Z_2, \bar{Z}_1] & \sigma_{Z_2}^2 \end{pmatrix} \quad (\text{B.43})$$

The covariance matrix $C_{\mathbf{Z}}^{(2)}$ is Hermitian, since the covariance matrix, in a general term, is symmetric as shown for the covariance matrix of real random variables. A Hermitian matrix is symmetric and conjugated. Also, the diagonal joint cumulants give the variances of the circular complex random variables, which are real numbers, and the off-diagonal measure the correlation between the complex random variables. Any other joint cumulants are null since the complex random variables are assumed to be circular. Actually, one of the classic ways to test for circularity is to examine that all other joint cumulants are null. Extending the definition of the covariance matrix $C_{\mathbf{Z}}^{(2)}$ to a complex random vector made of I circular complex random variables $\mathbf{Z} = [Z_1, Z_2, \dots, Z_I]$

produces a covariance matrix made of $I \times I$ joint cumulants, which they can be grouped as follows:

$$\mathbf{C}_Z^{(2)} = E[\mathbf{Z}\mathbf{Z}^H] = \begin{pmatrix} E[Z_1, \bar{Z}_1] & \cdots & E[Z_1, \bar{Z}_I] \\ \vdots & \ddots & \vdots \\ E[Z_I, \bar{Z}_1] & \cdots & E[Z_I, \bar{Z}_I] \end{pmatrix} \quad (\text{B.44})$$

Where \mathbf{Z}^H stands for the Hermitian vector \mathbf{Z} .

In many situations the probability density function is not known, but samples are generally available. Using the samples of \mathbf{Z} and $\bar{\mathbf{Z}}$, the joint expectation of is estimated by averaging their samples using the following formula:

$$E[Z, \bar{Z}] = \frac{1}{N} \sum_{i=1}^N Z_i \bar{Z}_i, \quad \text{where } i, j = 1, 2, \dots, N \quad (\text{B.45})$$

One of the properties, of the covariance matrices of complex random variable that is needed for the derivation of the virtual seismic array, is the multilinearity. If a complex random vector \mathbf{Z} , made of I complex random variables, is linearly related to another complex random vector $\hat{\mathbf{Z}}$ by a matrix \mathbf{V} made of $I \times I$ elements such that $\mathbf{Z} = \mathbf{V}\hat{\mathbf{Z}}$. Then, the second-order joint cumulants of \mathbf{Z} are related to second-order joint cumulants of $\hat{\mathbf{Z}}$, as follows:

$$\text{Cum}[Z_i, \bar{Z}_j] = \sum_{k=1}^I \sum_{l=1}^I v_{ik} \bar{v}_{jl} \text{Cum}[\dot{Z}_k, \bar{\dot{Z}}_l] \quad (\text{B.46})$$

Where v_{ik} are the elements of \mathbf{V} . In covariance notation, it is expressed as follows:

$$\mathbf{C}_Z^{(2)} = \mathbf{V} \mathbf{C}_Z^{(2)} \mathbf{V}^H \quad (\text{B.47})$$

Fourth-order joint cumulants

For circular complex random variables, all joint moments and joint cumulants of the third-order are null as a result of the inequality between the conjugated and non-conjugated terms. Actually, the only nonzero terms for all even joint moments and joint cumulants are the ones that have equal terms of the conjugated and non-conjugated terms. For example, the fourth-order joint cumulant $\text{Cum}[Z_1, Z_2, Z_3, \bar{Z}_4]$ is null because the inequality of the conjugated and non-conjugated terms, while the joint cumulant $\text{Cum}[Z_1, Z_2, \bar{Z}_3, \bar{Z}_4]$ is not zero because the conjugated equal non-conjugated terms. Therefore, while the fourth-order joint cumulants can be grouped in a fourth-order tensor, which is an array of four indices, denoted by $\mathbf{C}_Z^{(4)}$, the only nonzero terms for a complex random vector, which has I circular complex random variables, are the ones of the following form:

$$\begin{aligned}
\text{Cum}[Z_i, Z_j, \bar{Z}_k, \bar{Z}_l] &= E[Z_i, Z_j, \bar{Z}_k, \bar{Z}_l] - E[Z_i, \bar{Z}_k]E[Z_j, \bar{Z}_l] \\
&\quad - E[Z_i, \bar{Z}_l]E[Z_j, \bar{Z}_k],
\end{aligned} \tag{B.48}$$

where

i, j, k and l vary from 1 to l

The fourth-order joint cumulant tensor is an array of I^4 elements that can be grouped into $I^2 \times I^2$ matrix or I^2 matrices. In other words, the fourth-order joint cumulant tensor can be rearranged into I^2 matrices; similar to the covariance matrices. This arrangement is a fundamental to deriving the virtual seismic array. Another important property of the fourth-order joint cumulant tensor is the multilinearity property, which is similar to the multilinearity property of the covariance matrices. Again, If a complex random vector \mathbf{Z} , made of I complex random variables, is linearly related to another complex random vector $\hat{\mathbf{Z}}$ by a matrix \mathbf{V} made of $I \times I$ elements such that $\mathbf{Z} = \mathbf{V}\hat{\mathbf{Z}}$. Then, the fourth-order joint cumulants of \mathbf{Z} are related to fourth-order joint cumulants of $\hat{\mathbf{Z}}$, as follows:

$$\begin{aligned}
& \text{Cum}[Z_i, Z_j, \bar{Z}_k, \bar{Z}_l] \\
&= \sum_{p=1}^I \sum_{q=1}^I \sum_{r=1}^I \sum_{s=1}^I v_{ip} v_{jq} \bar{v}_{kr} \bar{v}_{ls} \text{Cum}[\acute{Z}_p, \acute{Z}_q, \bar{\acute{Z}}_r, \bar{\acute{Z}}_s] \tag{B.49}
\end{aligned}$$

In covariance notation, it is expressed as follows:

$$\mathbf{C}_Z^{(4)} = (\mathbf{V} \otimes \bar{\mathbf{V}}) \mathbf{C}_Z^{(4)} (\mathbf{V} \otimes \bar{\mathbf{V}})^H \tag{B.50}$$

Where \otimes is the Kronecker product (see Appendix C).

APPENDIX C

KRONECKER PRODUCT

If \mathbf{A} is an $M \times N$ matrix and \mathbf{B} is a $P \times Q$ matrix, then the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is the $MP \times NQ$ block matrix, which has the following form:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix} \quad (\text{C.1})$$

More explicitly,

$\mathbf{A} \otimes \mathbf{B}$

$$= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1q} & \cdots & \cdots & a_{1n}b_{11} & a_{1n}b_{12} & \cdots & a_{1n}b_{1q} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2q} & \cdots & \cdots & a_{1n}b_{21} & a_{1n}b_{22} & \cdots & a_{1n}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & a_{11}b_{p2} & \cdots & a_{11}b_{pq} & \cdots & \cdots & a_{1n}b_{p1} & a_{1n}b_{p2} & \cdots & a_{1n}b_{pq} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \cdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \cdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1q} & \cdots & \cdots & a_{mn}b_{11} & a_{mn}b_{12} & \cdots & a_{mn}b_{1q} \\ a_{m1}b_{21} & a_{m1}b_{22} & \ddots & a_{m1}b_{2q} & \cdots & \cdots & a_{mn}b_{21} & a_{mn}b_{22} & \cdots & a_{mn}b_{2q} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{p1} & a_{m1}b_{p2} & \cdots & a_{m1}b_{pq} & \cdots & \cdots & a_{mn}b_{p1} & a_{mn}b_{p2} & \cdots & a_{mn}b_{pq} \end{pmatrix}$$

Here are two examples:

$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 0 & 6 \\ 2 & 1 & 4 & 2 \\ 0 & 9 & 0 & 3 \\ 6 & 3 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\ a_{31}b_{11} & a_{31}b_{12} & a_{31}b_{13} & a_{32}b_{11} & a_{32}b_{12} & a_{32}b_{13} \\ a_{31}b_{21} & a_{31}b_{22} & a_{31}b_{23} & a_{32}b_{21} & a_{32}b_{22} & a_{32}b_{23} \end{pmatrix}$$

APPENDIX D**FINAL EXAM PRESENTATION**

**The concept of virtual arrays in
seismic data acquisition
(Thesis defense)**

*Ibrahim Alhukail
Texas A&M University
May 31, 2012*

Thesis Objective

To improve seismic array response by constructing artificial sensors from the real sensors

“virtual sensors”

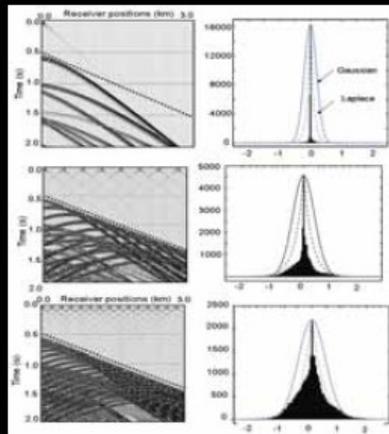
The combination of real and virtual sensors is called:

“virtual seismic array”

Key characteristic:

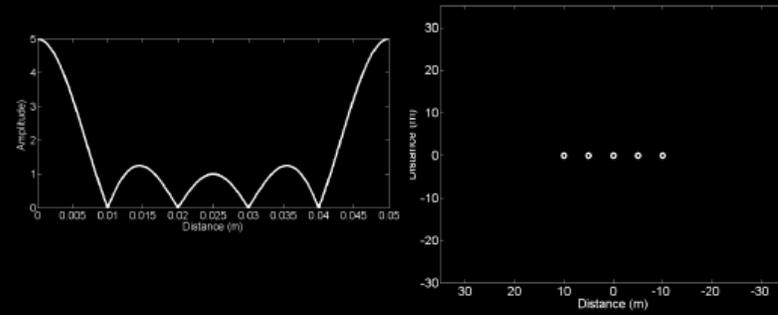
“Seismic data are non-Gaussian”

Seismic data: nonGaussian

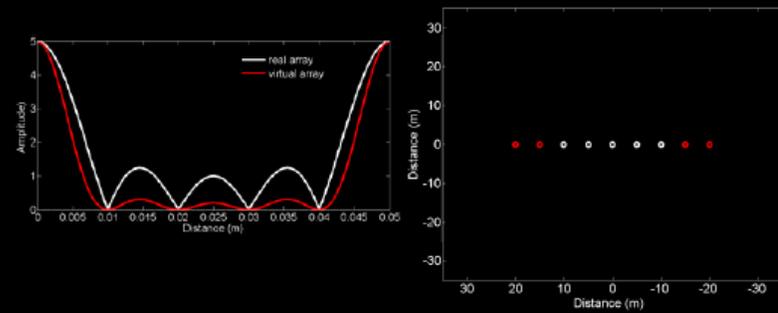


(Ikelle, 2010)

Thesis Objective



Thesis Objective



Outline

- Virtual seismic array
 - Review of statistics of second- and fourth-order statistics (nonGaussianity).
 - Mathematics of virtual seismic arrays
 - Virtual array examples
- Modeling array data
 - Theoretical comparison of staggered-grid and pseudo-spectral finite-difference techniques
 - Numerical comparisons
 - Numerical example (groundroll)
- Conclusions and recommendations

Second-order statistics (Gaussianity)

Zero-mean random variables: $\{X_1, X_2, \dots, X_I\}$

$$\text{Cum}[X_i] = \text{E}[X_i] = 0$$

$$\text{Cum}[X_i, X_j] = \text{E}[X_i X_j]$$

**Covariance
matrix**

$I \times I$

$$\mathbf{C}_X^{(2)} = \begin{pmatrix} \text{E}[X_1 X_1] & \dots & \text{E}[X_1 X_I] \\ \vdots & \ddots & \vdots \\ \text{E}[X_I X_1] & \dots & \text{E}[X_I X_I] \end{pmatrix}$$

Higher-order statistics (nonGaussianity)

$$\text{Cum}[X_i, X_j, X_k] = E[X_i X_j X_k] = 0 \text{ (symmetry)}$$

$$\begin{aligned} \text{Cum}[X_i, X_j, X_k, X_l] &= E[X_i X_j X_k X_l] - E[X_i X_j]E[X_k X_l] \\ &\quad - E[X_i X_k]E[X_j X_l] - E[X_i X_l]E[X_j X_k] \end{aligned}$$

Fourth-rank
cumulant tensor



$C_X^{(4)}$

$$I \times I \times I \times I$$

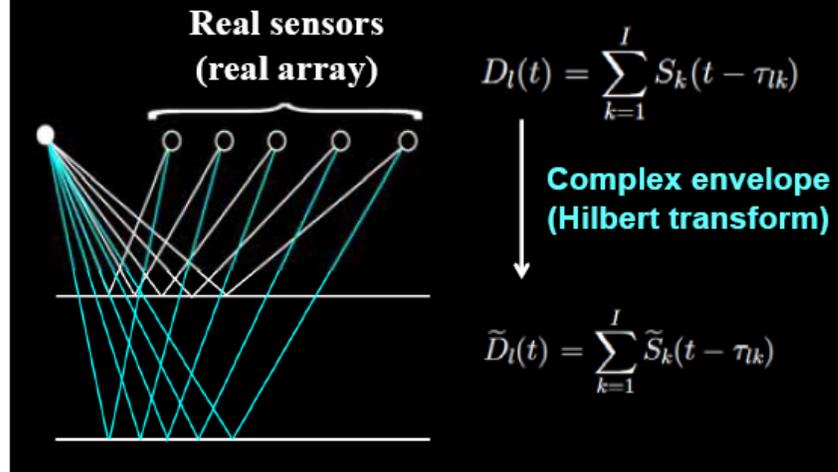
or

$$I^2 \times I^2$$

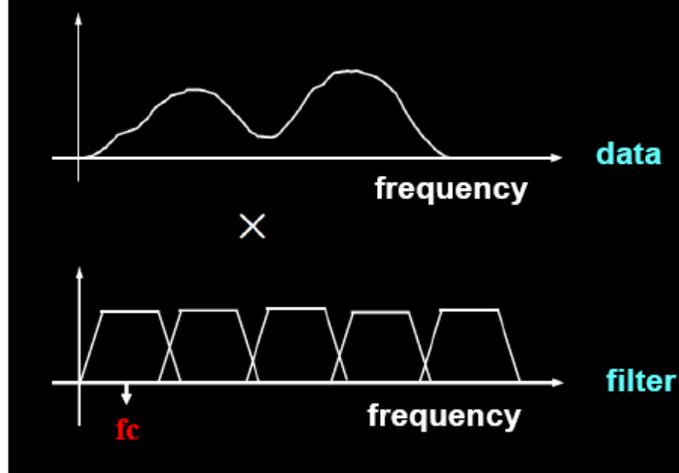
Outline

- Virtual seismic array
 - Review of statistics of second- and fourth-order statistics (nonGaussianity).
 - Mathematics of virtual seismic arrays
 - Virtual array examples
- Modeling array data
 - Theoretical comparison of staggered-grid and pseudo-spectral finite-difference techniques
 - Numerical comparisons
 - Numerical example (groundroll)
- Conclusions and recommendations

Array response



Narrow-band signal decomposition



Narrow-band signal decomposition

So we can carry on our processing with narrow-band signals and assume that this process is repeated for all the narrow-band components of the original signals.

For one narrow-band component:

$$\tilde{D}_l(t) = \sum_{k=1}^I \exp[2\pi i f_c \tau_k] \tilde{S}_k(t)$$

⇒ We now have a **linear relationship**.

Narrow-band signal decomposition

Using the inter-sensor phase shift θ_k , which is can be as follow:

$$\tilde{D}_l \sum_{k=1}^K \exp[i(l-1)\theta_k] \tilde{S}_k(t)$$

where

$$\theta_k = \omega_c \frac{(l-1) \Delta x \sin \alpha_k}{V}$$

Narrow-band signal: Array response

$$\tilde{\mathbf{D}}(t) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_L)] \tilde{\mathbf{S}}(t) = \mathbf{A} \tilde{\mathbf{S}}(t)$$

where $\tilde{\mathbf{D}}(t)$ describes an L -dimension vector of the array responses,

$\tilde{\mathbf{S}}(t)$ represents a L -dimension vector of the signal responses and \mathbf{A} represents the mixing matrix, whose size is $L \times L$.

$$\mathbf{a}(\theta_k) = [1, \exp(-i\theta_k), \dots, \exp(-i(L-1)\theta_k)]^T$$

Narrow-band signal: Array response

l -the components of steering vectors for areal arrays:

$$a_l(\theta_k) = \exp \left[i \frac{2\pi}{\lambda} [x_l \cos(\theta_k) + y_l \sin(\theta_k)] \right]$$

(x_l, y_l) : Coordinate of the l sensor of the array.

Covariance vs. cumulant tensor (quadricovariance):

$$\mathbf{C}_D^{(2)} = \sum_{k=1}^I \mathbf{C}_S^{(2)}(k, k) \mathbf{a}(\theta_k) \mathbf{a}^H(\theta_k)$$

$$\mathbf{C}_D^{(4)} = \sum_{k=1}^I \mathbf{C}_S^{(4)}(k, k, k, k) [\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)] [\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)]^H$$

- Covariance and quadricovariance have the same algebraic structure.
- The vector $[\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)]$ is a steering vector as the vector $\mathbf{a}(\theta_k)$

Real and virtual arrays:

L real sensors :	L^2 real and virtual sensors :
$\mathbf{a}(\theta_k)$	$[\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)]$

Real and virtual arrays: (areal case):

Components of steering vectors for real arrays:

$$a_l(\theta_k) = \exp \left[i \frac{2\pi}{\lambda} [x_l \cos(\theta_k) + y_l \sin(\theta_k)] \right]$$

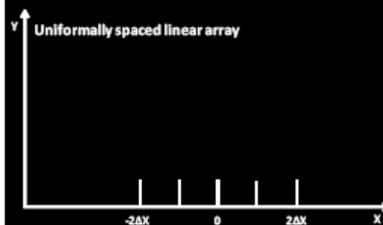
Components of steering vectors for virtual arrays:

$$[\mathbf{a}(\theta_k) \otimes \mathbf{a}^*(\theta_k)]_{r,q} = \exp \left[i \frac{2\pi}{\lambda} [(x_r - x_q) \cos(\theta_k) + (y_r - y_q) \sin(\theta_k)] \right]$$

$(x_r - x_q, y_r - y_q)$: Coordinate of the r,q -sensor

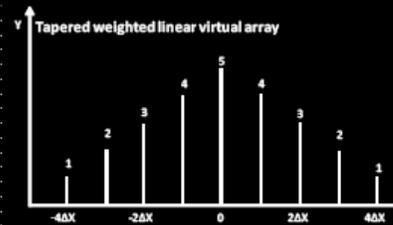
Example: linear array (weights)

Real array



$$(x_l, 0) = ([l - 1]\Delta x, 0)$$

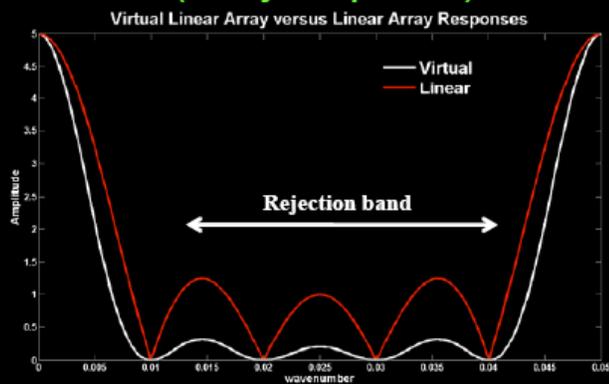
Virtual array



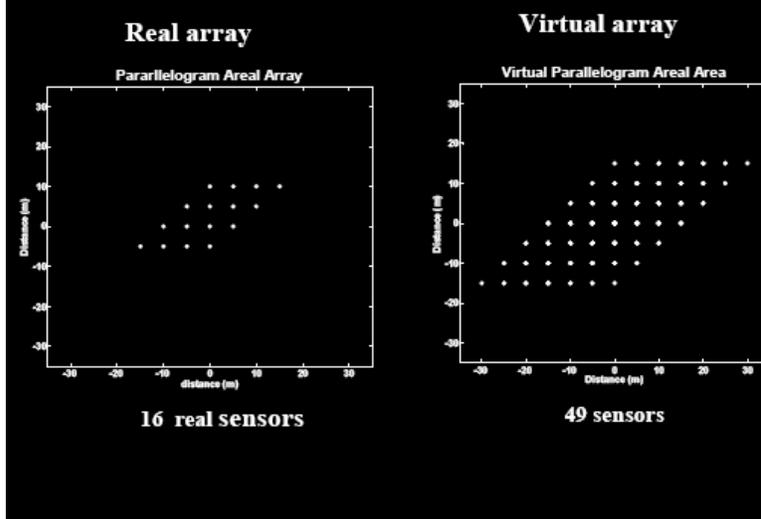
$$(x_r - x_q, 0) = ([r - q]\Delta x, 0)$$

- $2L-1$ virtual sensors (9 virtual sensors in this case)
- The r,q -sensor has a weights of $L-|r-q|$

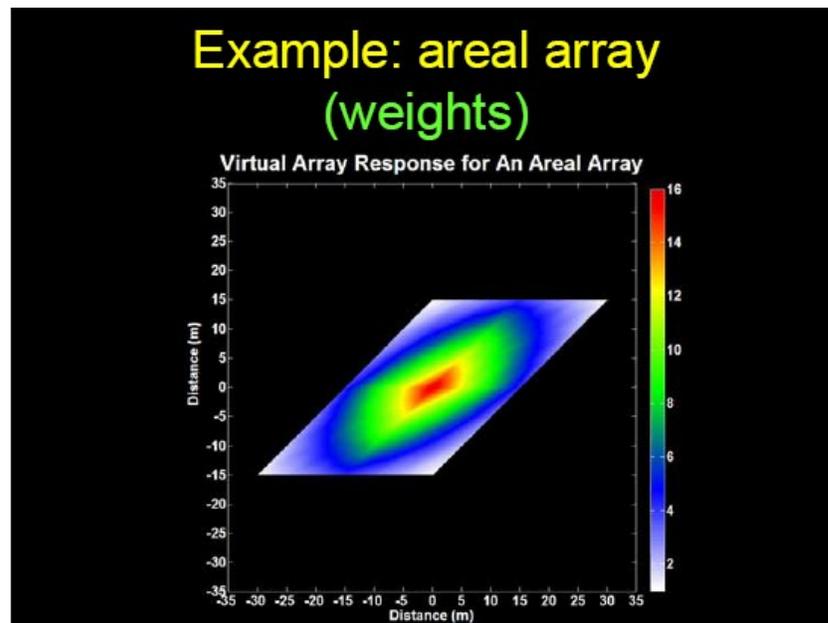
Example: linear array (array response)



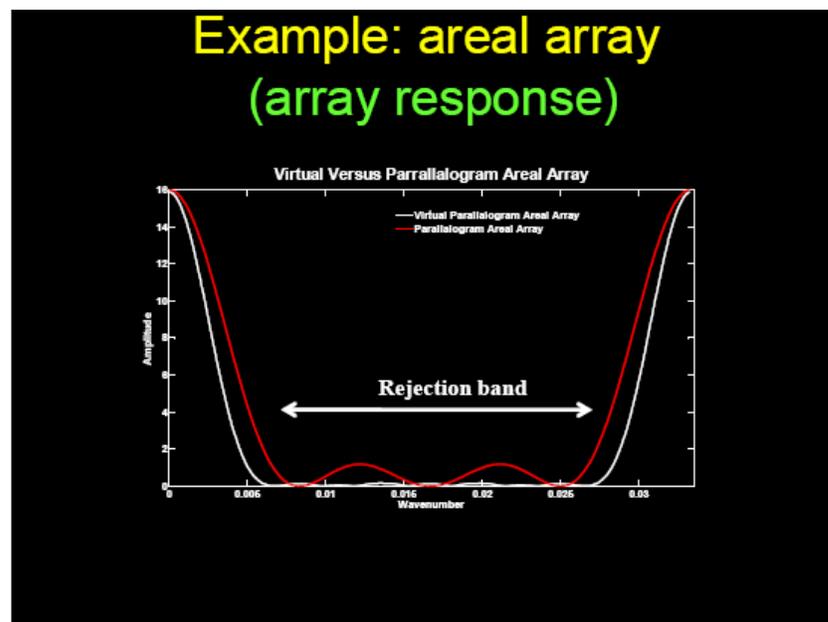
Example: areal array



Example: areal array (weights)



Example: areal array (array response)



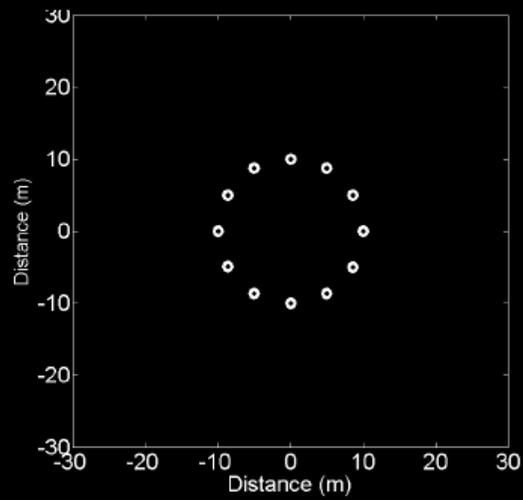
Outline

- **Virtual seismic array**
 - Review of statistics of second- and fourth-order statistics (nonGaussianity).
 - Mathematics of virtual seismic arrays
 - Virtual array examples
- **Modeling array data**
 - Theoretical comparison of staggered-grid and pseudo-spectral finite-difference techniques
 - Numerical comparisons
 - Numerical example (groundroll)
- **Conclusions and recommendations**

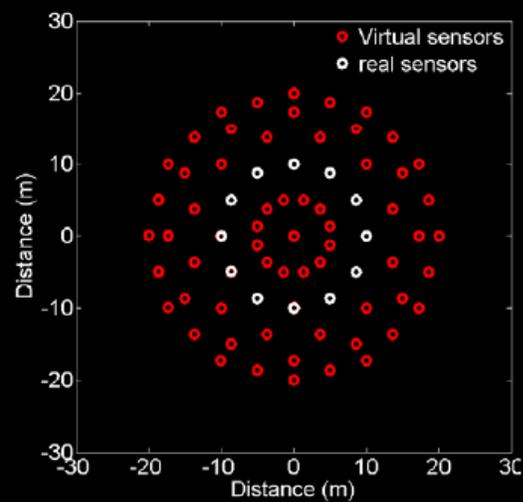
In what cases virtual sensors can't be created?

- **When my signal is Gaussian.**
- **My array consists of only one sensor.**

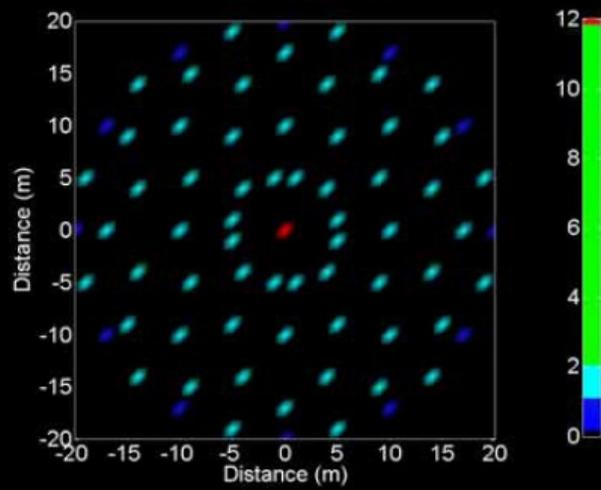
In what cases I create no multiplicity?



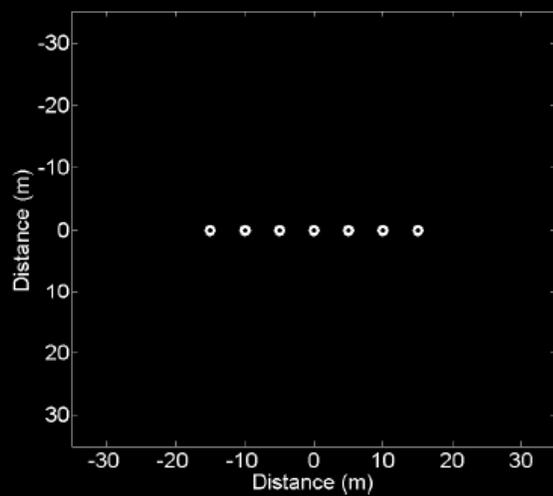
In what cases I create no multiplicity?



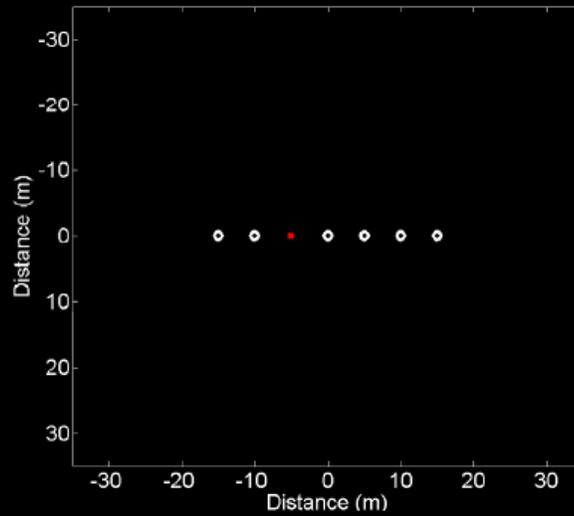
Virtual array benefits: (weights)



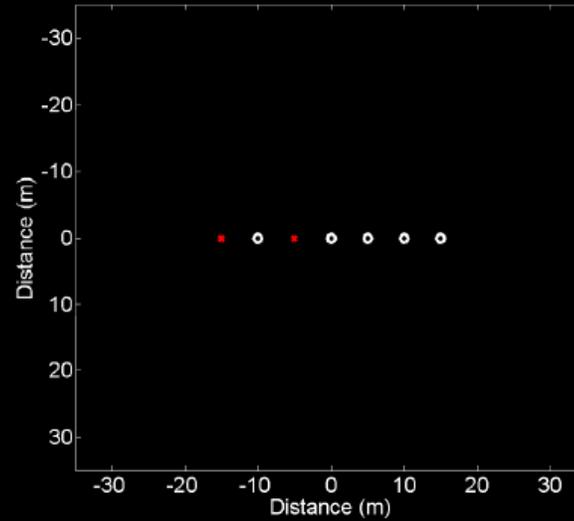
Virtual array benefits: (replace malfunctioning sensors)



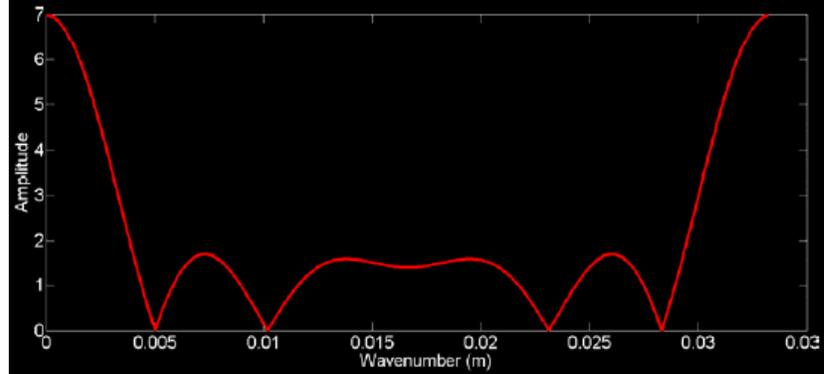
Virtual array benefits:
(replace malfunctioning sensors)



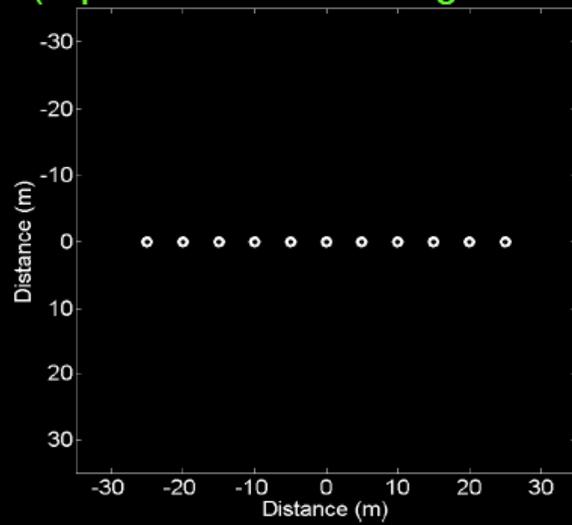
Virtual array benefits:
(replace malfunctioning sensors)



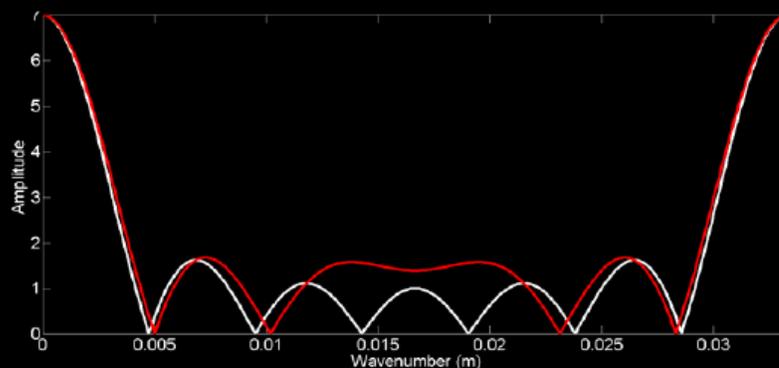
Virtual array benefits:
(replace malfunctioning sensors)



Virtual array benefits:
(replace malfunctioning sensors)



Virtual array benefits: (replace malfunctioning sensors)



Outline

- Virtual seismic array
 - Review of statistics of second- and fourth-order statistics (nonGaussianity).
 - Mathematics of virtual seismic arrays
 - Virtual array examples
- Modeling array data
 - Theoretical comparison of staggered-grid and pseudo-spectral finite-difference techniques
 - Numerical comparisons
 - Numerical example (groundroll)
- Conclusions and recommendations

Modeling array data

Finite-difference modeling tool:

- Array modeling requires very fine sampling, which makes current techniques (staggered-grid finite-difference) cost prohibitive.
- Investigated an alternative finite-difference technique to the staggered-grid finite-difference (pseudo-spectral finite-difference), that I can use to generate seismic data at the array level.

Wavenumber-based finite differences (Pseudo-spectral method)

- KOSLOF, RESHEF and LOEWENTHAL (1984)
- It is an alternative approach for solving the elastic wave equations.
- Spatial derivatives are computed with the use of the Fast Fourier Transform
- Temporal derivatives are calculated with second order differencing.

Spatial derivatives

Fourier transform reduces
differentiation to multiplication

$$\frac{\partial f(x)}{\partial x} = \hat{F}^{-1}[-ik\hat{F}(k)]$$

$$\frac{\partial^2 f(x)}{\partial x^2} = \hat{F}^{-1}[k^2\hat{F}(k)]$$

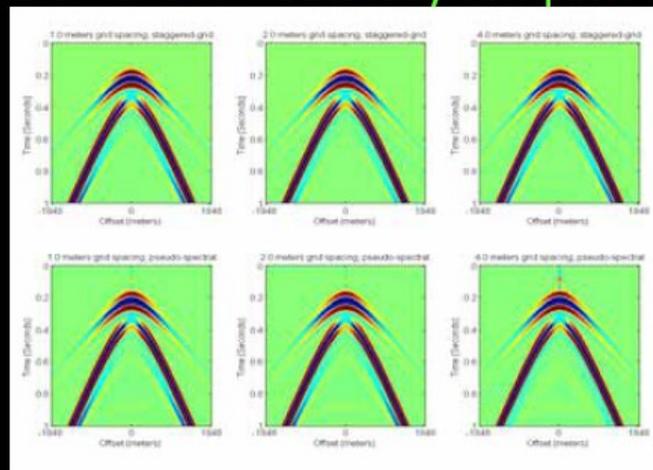
Staggered-grid and pseudo-spectral finite-difference (2D): Theoretical comparison

Theoretical comparison	Staggered-grid finite-difference	Pseudo-spectral finite-difference
Spatial derivatives	Fourth-order difference operator	Fast Fourier-transform
Temporal derivatives	Second-order difference operator	Second-order difference operator
Stability condition	$\Delta t < 0.606 \frac{\Delta x}{V_{max}}$	$\Delta t < 0.4502 \frac{\Delta x}{V_{max}}$
Grid-dispersion	$\Delta x < \frac{V_{min}}{5f_{max}}$	$\Delta x < \frac{V_{min}}{2f_{max}}$

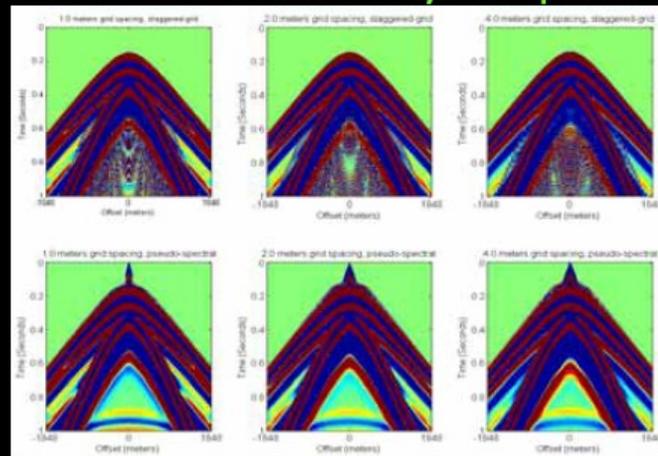
Staggered-grid and pseudo-spectral finite-difference: accuracy comparison

Grid-spacing	4 meters	2 meters	1 meters
Size of the model	(1024X512)	(2048X1024)	(4096X2048)
Timestep	5.0×10^{-4}	2.5×10^{-4}	1.0×10^{-4}
Number of timesteps	2000	4000	10000
Number of grid-points (absorbing boundaries)	100	200	400
Total record time	1 second	1 second	1 second

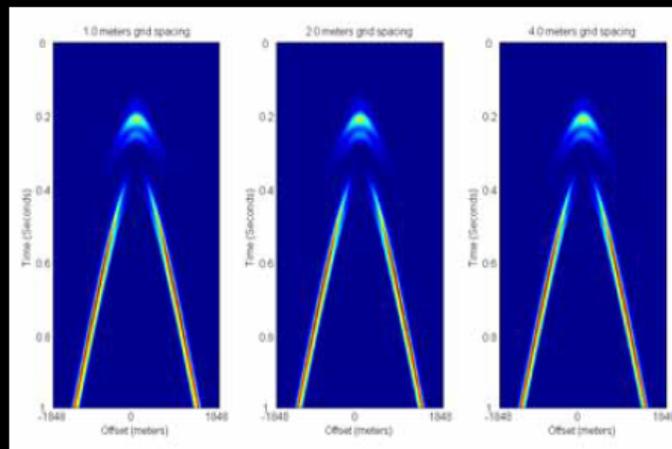
Staggered-grid and pseudo-spectral finite-difference: accuracy comparison



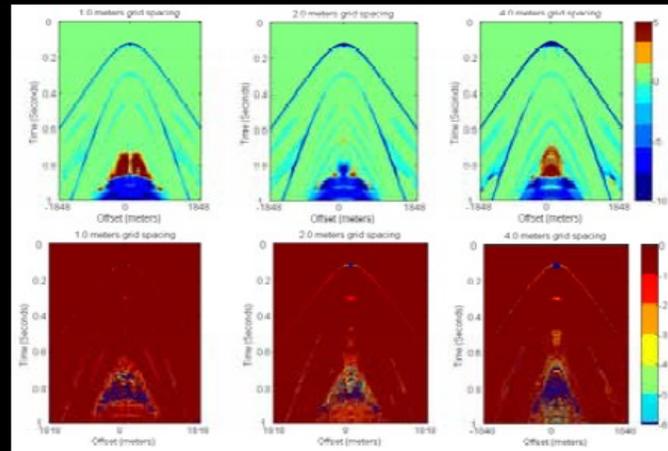
Staggered-grid and pseudo-spectral finite-difference: accuracy comparison



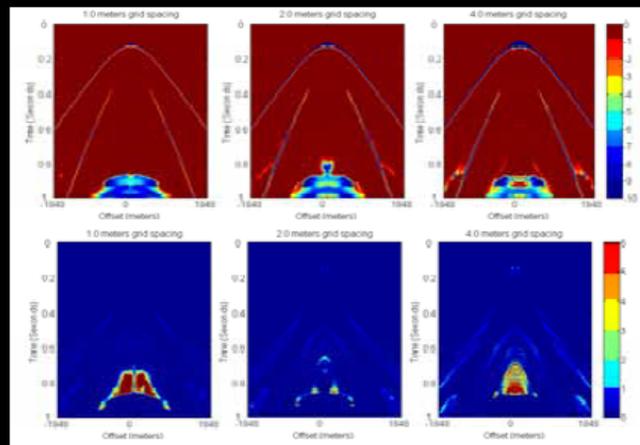
Staggered-grid and pseudo-spectral finite-difference: crosscorrelations



Staggered-grid and pseudo-spectral finite-difference: normalized crosscorrelations



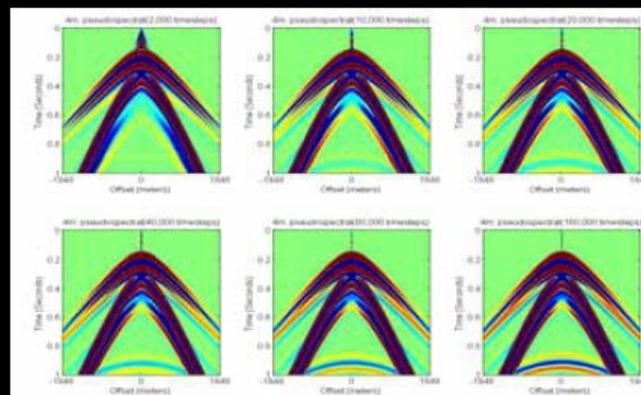
Staggered-grid and pseudo-spectral finite-difference: normalized crosscorrelations



Pseudo-spectral finite-difference: accuracy comparison (timesteps)

Shot-record	Timestep (millisecond)	Number of timesteps	Length of shot records
1	0.5	2,000	One-second long
2	0.1	10,000	One-second long
3	0.05	20,000	One-second long
4	0.025	40,000	One-second long
5	0.0125	80,000	One-second long
6	0.00625	160,000	One-second long

Pseudo-spectral finite-difference: accuracy comparison (timesteps)



Staggered-grid and pseudo-spectral finite-difference: efficiency comparison Example III

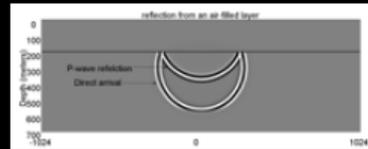
Examples III		Staggered-grid finite-difference				Pseudo-spectral finite-difference			
model size (km)	(2.5X2.5)	Model size	Δx (m)	Δt (ms)	#of timesteps	Model size	Δx (m)	Δt (ms)	#of timesteps
Record length	2 seconds								
Vmin	300m/s								
Vmax	5000m/s	(10,000X10,000)	0.25	0.025	80,000	(2000X2000)	1.25	0.1	20,000
Fmax	100 Hz								

Outline

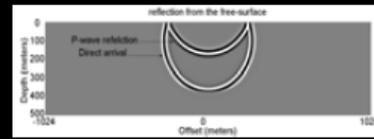
- Virtual seismic array
 - Review of statistics of second- and fourth-order statistics (nonGaussianity).
 - Mathematics of virtual seismic arrays
 - Virtual array examples
- Modeling array data
 - Theoretical comparison of staggered-grid and pseudo-spectral finite-difference techniques
 - Numerical comparisons
 - Numerical example (groundroll)
- Conclusions and recommendations

Pseudo-spectral finite-difference: numerical example (air-filled layer)

Air-filled layer (free-
surface)

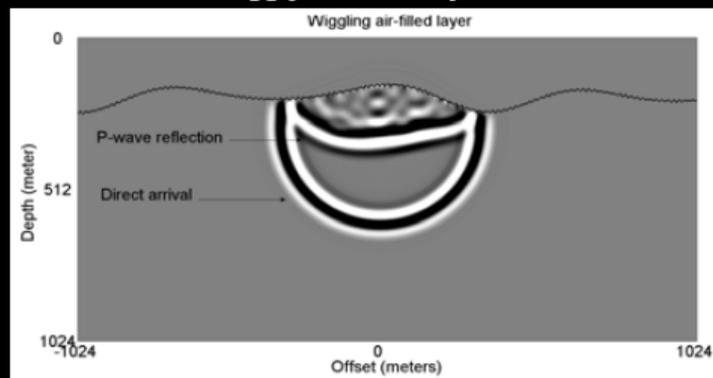


Free-surface

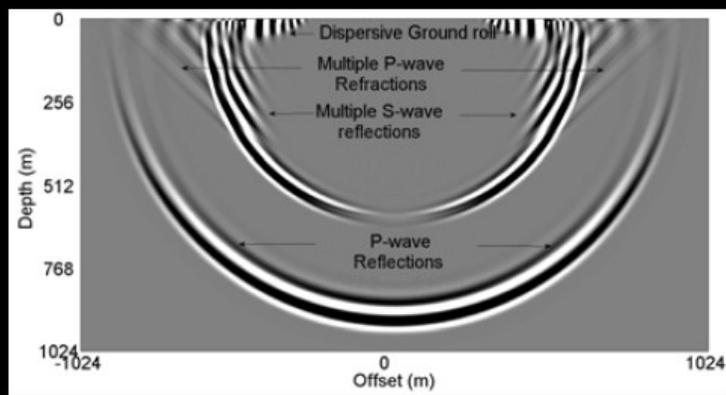


Pseudo-spectral finite-difference: numerical example (air-filled layer)

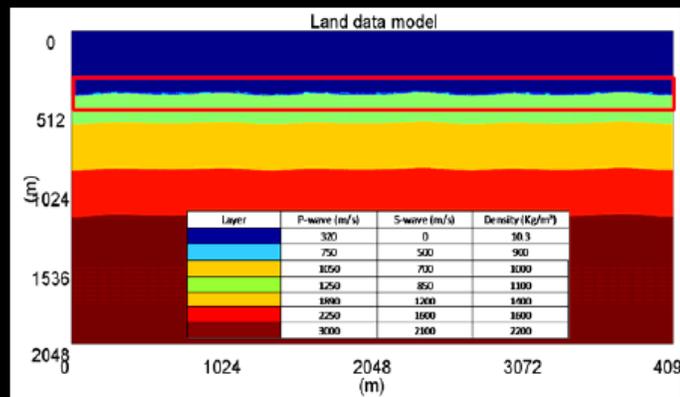
Wiggly air-filled layer



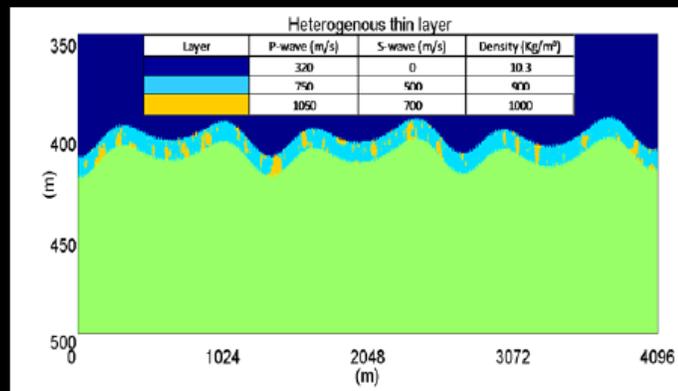
Pseudo-spectral finite-difference: dispersive groundroll



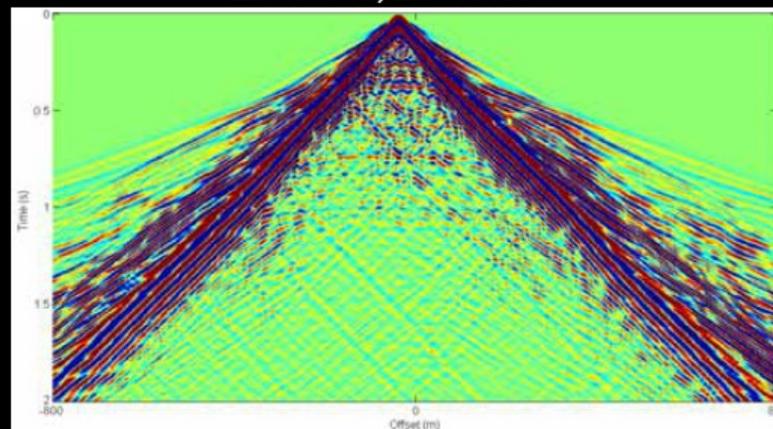
Pseudo-spectral finite-difference: land data model



Pseudo-spectral finite-difference: land data model



Pseudo-spectral finite-difference: land data model (shot record) Finite-difference model size (2048X1024) $\Delta t = .05\text{ms}$, $\Delta x = 1\text{meter}$



Conclusion

- Because spatial derivatives computed by the Fourier transform has an infinite order of accuracy, the pseudo-spectral finite-difference only requires two grid-points per wavelength to sample the wavefields, versus 8 grid-points per wavelength for a fourth-order difference operator or 16 grid-points per wavelength for a second-order difference operator, in the staggered-grid finite-difference, to achieve comparable accuracy.
- Therefore, the pseudo-spectral finite-difference is much more efficient than the staggered-grid finite-difference, particularly for numerical modeling of array data; where the grid-spacing is in the order of one meter or less.

Conclusion

- The key assumption is that seismic data are non-Gaussian; hence its fourth-order cross-cumulants of the sensors signals are nonzero.
- Additional sensors can be constructed that are artificial by using the relationship between covariance matrices and the fourth-order cross-cumulants.
- For example, for an equally weighted, linear array of five sensors, a tapered-weighted virtual array of nine sensors can be made.

Recommendations of future work

- We are in the process of investigating how this concept of seismic virtual arrays can be integrated into the existing seismic acquisition technology.
- One of the key assumptions in the derivation of the virtual seismic array is that array responses are narrowband. However, seismic data are wideband signals.
- The future plans will include a description of how a wideband signal can be decomposed into narrowband signals.
- In particular, we focus on the filterbank technique for decomposing a wideband signal into a group of narrowband signals.

Thank you